



Investigating Simple and Complex Mechanisms in Texture Segregation Using the Speed–Accuracy Tradeoff Method

ANNE SUTTER,*‡ NORMA GRAHAM†

Received 4 November 1994; in revised form 23 January 1995

Several recent models of texture segregation have proposed two mechanisms: *simple*, linear channels (first-order, Fourier mechanisms) and *complex* channels (second-order, non-Fourier mechanisms). We used the speed–accuracy tradeoff (SAT) method to examine the *time course* of texture segregation processing in simple and complex channels. The stimuli were texture patterns designed to segregate primarily as a result of activity in one set of channels or the other. We presented subjects with textures that were checked or striped arrangements of either Gaussian-blob or Gabor-patch elements. Subjects were required to identify the orientation of a rectangular texture region embedded in a background field of a different texture. A range of contrasts and a control task were used to equate visibility of the Gabor and Gaussian textures. SAT functions were obtained by requiring subjects to respond within 200 msec after an auditory cue. We found that when segregation depended primarily on simple channels, performance was faster than when it depended primarily on complex channels: the 75% correct level was reached 100–200 msec sooner and this extra speed was reflected both in smaller delay and higher rate parameters.

Texture perception Processing dynamics Simple Complex

INTRODUCTION

Preattentive texture segregation has frequently been defined as the process by which adjacent areas of differing local spatial structure are perceived immediately and effortlessly to be distinct regions in an image. For years, investigators have used the words “immediate” and “effortless” to describe an important feature of texture segregation (e.g. Julesz, 1975; Beck, 1982; Bergen, 1991; Ben-Av, Sagi & Braun, 1992; Wolfe, 1992). However, these words have referred to a phenomenological attribute of the perception of texture rather than to an empirically determined characteristic of the processes underlying the perception. Partly because texture segregation seems ‘immediate and effortless’, investigators have proposed theories of texture processing consisting predominantly of stages occurring relatively early in vision, usually linear filters and some early nonlinearities (Beck, Sutter & Ivry, 1987; Bergen & Landy, 1991; Bovik, Clark & Geisler, 1987; Caelli, 1988; Chubb & Sperling, 1988; Clark, Bovik & Geisler, 1987;

Fogel & Sagi, 1989; Graham, Beck & Sutter, 1992; Landy & Bergen, 1991; Malik & Perona, 1989a, b; Nothdurft, 1985a, b; Sagi, 1990; Sutter, Beck & Graham, 1989; Turner, 1986; Victor & Conte, 1987, 1989a, b, 1991; Victor, 1988; Wilson, Ferrara & Yo, 1992; Wilson & Richards, 1992). Nearly all experimental tests of these models have used information about completed segregation processing, i.e. information about the final percept. Here we report explorations of this type of model based on investigations of the time course of texture segregation processing.

Simple and complex channels

Many models of texture and motion perception propose mechanisms that include two stages of spatial-frequency and orientation selective linear filters separated by a rectification or similar nonlinearity (e.g. Chubb & Sperling, 1988; Sperling, 1989; Sutter *et al.*, 1989; Graham *et al.*, 1992; Wilson *et al.*, 1992; Wilson & Richards, 1992; Graham, Sutter & Venkatesan, 1993). We call our version of this mechanism a *complex channel*; others have referred to this kind of non-linear mechanism as a *non-Fourier* or *second-order* mechanism. The first-stage filter of the example shown in Fig. 1 is tuned to a high spatial frequency and horizontal orientation. [That the first-stage filters are orientation and spatial-frequency selective has been shown by Graham

*Department of Psychology, Loyola University, 6525 N. Sheridan Road, Chicago, IL 60626, U.S.A.

†Department of Psychology, Columbia University, New York, NY 10027, U.S.A.

‡To whom all correspondence should be addressed [Email asutter@luc.edu].

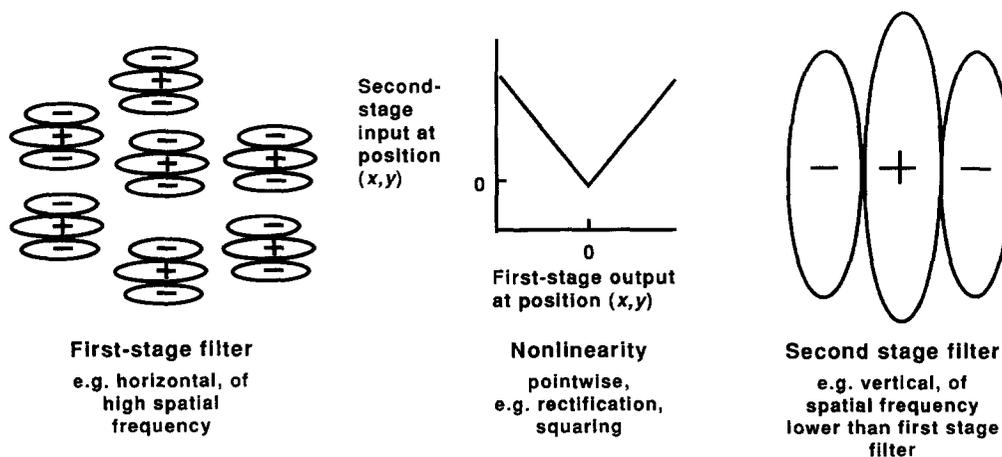


FIGURE 1. Diagram of complex channel.

et al. (1993)]. The outputs of the first-stage filter are rectified or undergo a similar nonlinearity, and are passed to the second-stage linear filter which is depicted in Fig. 1 as tuned to a low frequency and vertical orientation. Other combinations of first- and second-stage filters are assumed to exist. Complex channels can be thought of as detecting low spatial frequency arrangements of high spatial frequency elements.

In our model, information about the outputs of single-stage filters is available to segregation decision processes in addition to and independently of the outputs of the second-stage filters. We refer to such single-stage filters, considered by themselves, as *simple channels*. Others have referred to this kind of simple linear mechanism as a *Fourier* or *first-order* mechanism. Whether the single-stage filters are distinct filters or also serve as the first stage in the complex channels is not known.

Gaussian and Gabor-patch Embedded-Rectangle textures

In the experiment reported here, we used patterns composed of either Gabor patches or Gaussian blobs to examine the time course of segregation processing, and specifically, to address the question of whether or not processing dynamics differ for patterns that are segregated primarily by simple channels compared to those segregated primarily by complex channels.

Examples of the patterns that we used appear in Fig. 2. These patterns contain "element arrangement textures" (first used by Beck, Prazdny & Rosenfeld, 1983). The ones used in this study were composed either of Gaussian blobs of equal but opposite sign-of-contrast, or of Gabor-patch elements that were luminance-balanced with the background. These elements were arranged into either checked or striped regions, with a rectangular patch of one type of texture embedded in a background field of the other texture. According to our model, segregation of the embedded rectangle in these patterns is due to within-channel or within-filter differences in the pooled responses to the checked vs the striped regions (see Sutter *et al.*, 1989; Graham *et al.*, 1992 for details of the model). Also, according to the

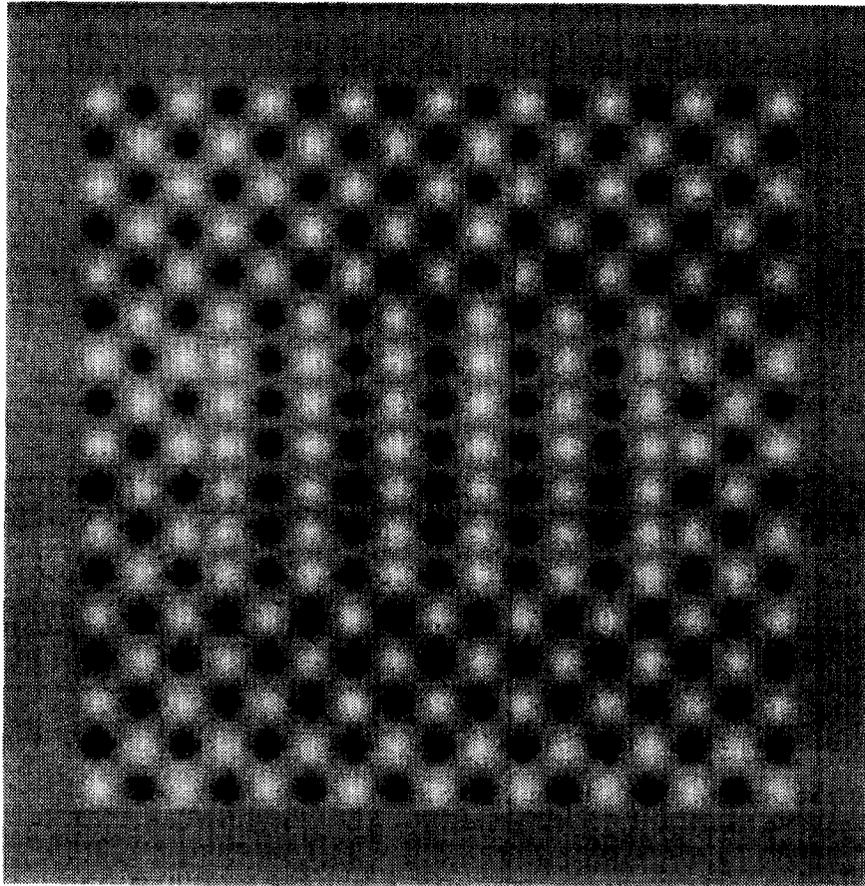
model, patterns composed of Gaussian blobs are segregated primarily as a result of activity in simple channels (one stage of linear filtering), whereas patterns composed of Gabor patches are segregated primarily as a result of activity in complex channels (two stages of linear filtering separated by a rectification or similar nonlinearity).

Figure 3(a) illustrates why patterns composed of Gaussian blobs of opposite sign of contrast segregate mainly because of the activity they produce in simple channels. Consider the top row of Fig. 3(a), which shows the response of a simple channel tuned to the fundamental frequency of the striped Gaussian-blob texture. The first box shows a small area of the striped region with a receptive field from this simple channel filter. The next box shows the output of this filter: note that there is a lot of modulated activity in response to the stripes. On the other hand, this filter will be unresponsive to the checked texture: thus this simple channel effectively discriminates between the two regions.

Now consider the second row of Fig. 3(a), which shows the response of a particular complex channel to the striped Gaussian-blob texture. This complex channel has a first-stage filter of high spatial frequency and oblique orientation, and a second-stage filter of low spatial frequency and vertical orientation, with a rectification (half-wave here) between the two stages. Notice, as represented in the second box, that the frequency of the first-stage filter is too high for the filter to signal the presence of the blobs, thus this complex channel does not do a better job of signaling the difference between the checked and striped textures than does the simple channel illustrated in the top row.

Figure 3(b) illustrates why patterns composed of Gabor patches that are luminance-balanced with respect to the background segregate mainly because of the activity they produce in complex channels. Consider the first row of Fig. 3(b), which shows the response of a

(a)



(b)

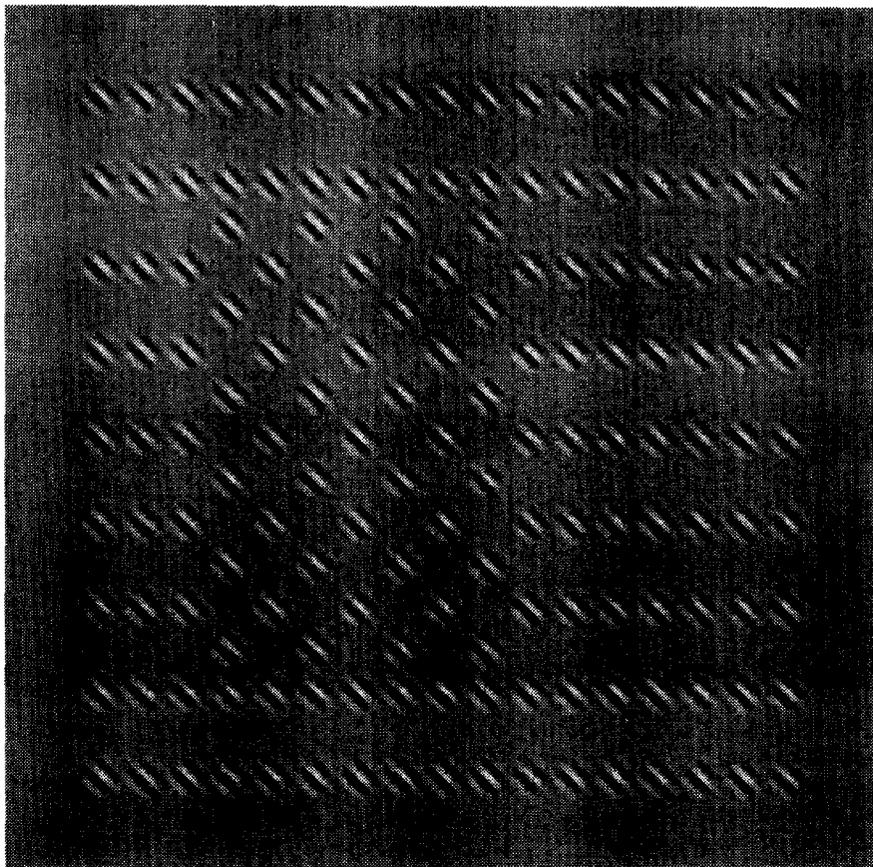
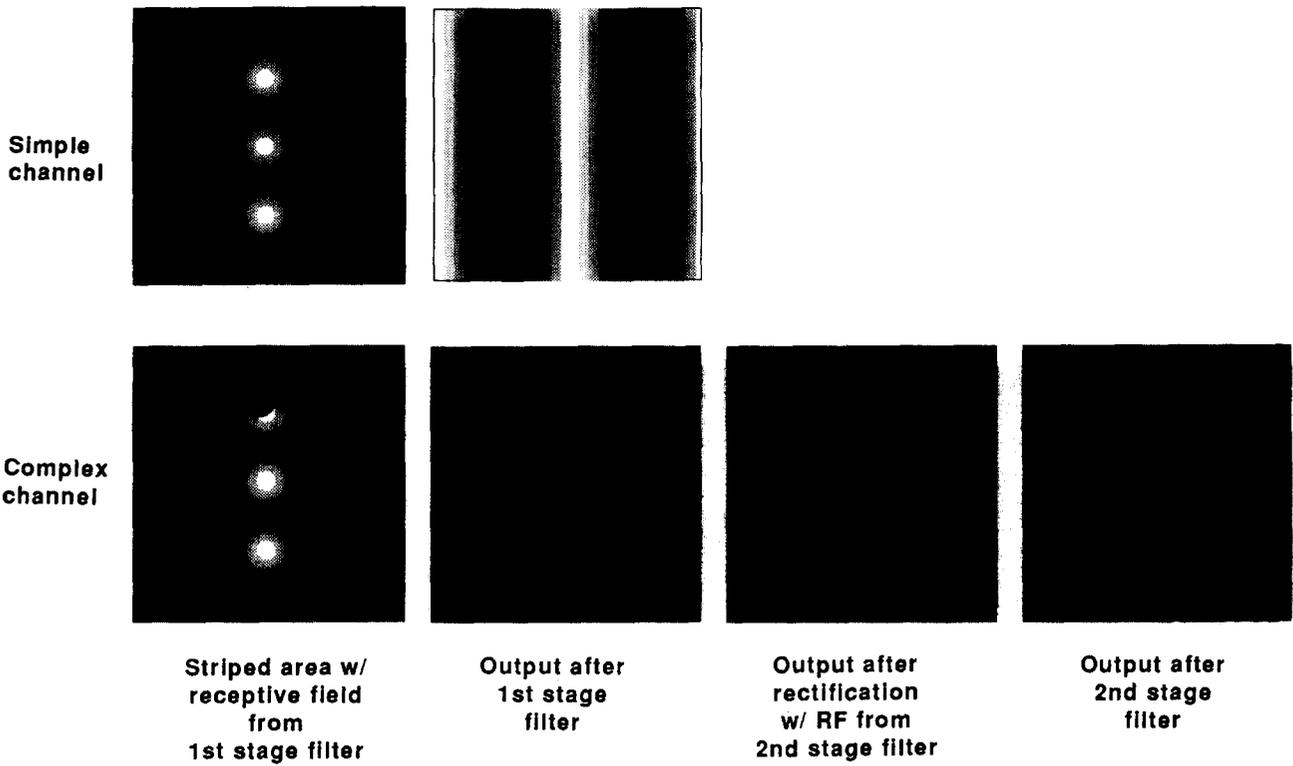


FIGURE 2. Two examples of Embedded-Rectangle patterns. (a) A pattern composed of Gaussian blobs of equal-but-opposite sign-of-contrast. (b) A pattern composed of Gabor patches luminance-balanced with the background. Reproduction will have distorted the images somewhat.

(a) Gaussian-blob embedded-rectangle textures are segregated primarily by simple channels



(b) Gabor-patch embedded-rectangle textures are segregated primarily by complex channels

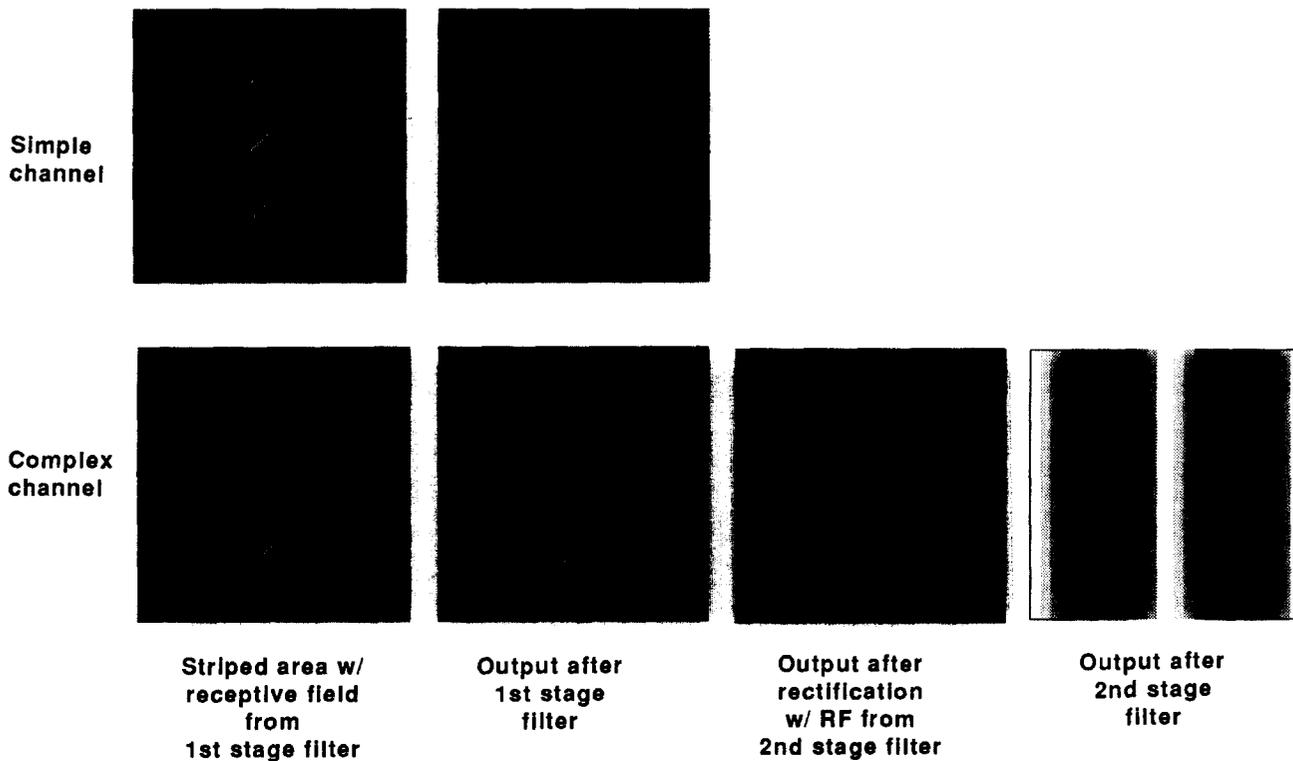


FIGURE 3. Diagram of responses of simple and complex channels to the striped portion of textures composed of Gaussian blobs (a), and textures composed of Gabor patches (b).

simple channel tuned to the fundamental frequency of the striped Gabor-patch texture. Notice that there is no response from this simple channel because its receptive fields are too large to detect these luminance-balanced patches. It so happens that *no* simple channel or single-stage filter will show a difference between the spatially-pooled responses to the checked and striped regions of these Embedded-Rectangle Gabor-patch textures. (As will be discussed later, however, a rectangle of texture containing Gabor-patch elements can be segregated from a *blank* surround by simple channels.)

Now consider the bottom row of Fig. 3(b), which shows the response of a complex channel consisting of a first-stage filter tuned to a high spatial frequency and oblique orientation, followed by a rectification and a second-stage filter tuned to the fundamental frequency and orientation of the striped Gabor-patch texture. The small receptive fields of the first-stage filter are of a size appropriate to detect the Gabor patches. There will be activity in both the checked and striped regions in the outputs of this first-stage filter, but after rectification, a second-stage filter (tuned to the fundamental frequency of the striped texture) will respond strongly to the striped region, but not to the checked region.

The speed-accuracy tradeoff paradigm

In this paper we investigate differences in the time course of texture segregation processing for Gaussian-blob (simple channel) and Gabor-patch (complex channel) patterns using the method of cued response to generate speed-accuracy tradeoff (SAT) functions.

SAT functions have been used for years by investigators of cognitive processes such as semantic and recognition memory (Reed, 1973; Doshier, 1976; Wickelgren, Corbett & Doshier, 1980; for overviews of the SAT method see Wickelgren, 1977 and Doshier, 1979). However, they have seldom been used to study visual or perceptual processing, and the few published studies we know of (e.g. Schouten & Bekker, 1967; Lappin & Disch, 1972; Wandell, 1977) have not investigated texture segregation or any related pattern discrimination processes. An SAT function gives some accuracy measure as a function of some response-time measure. The basic method for obtaining SAT functions requires the subject to make responses during several different time intervals through the manipulation of response cues, deadlines, payoffs, etc. By causing the subject to respond at different points in time during the processing of a stimulus, one can obtain data regarding the amount of information available to decision processes at those points in time.

At the very least, the SAT method gives the investigator a good chance of discovering differences in the processing of various stimuli which may be obscured when using other response methods. The advantage of causing subjects to respond during specific intervals in time is that they can be forced to make a response before they would ordinarily choose to in, for example, simple reaction time or the usual forced-choice discrimination experiments. Forcing a response during particular time

intervals yields a measure of performance in a task before the final percept has been established, that is, at various stages of completion of processing. In this way, one may discover differences between two texture stimuli that would produce equivalent performance if responses were delayed until completion of processing. This information may then be used to test models of these processes.

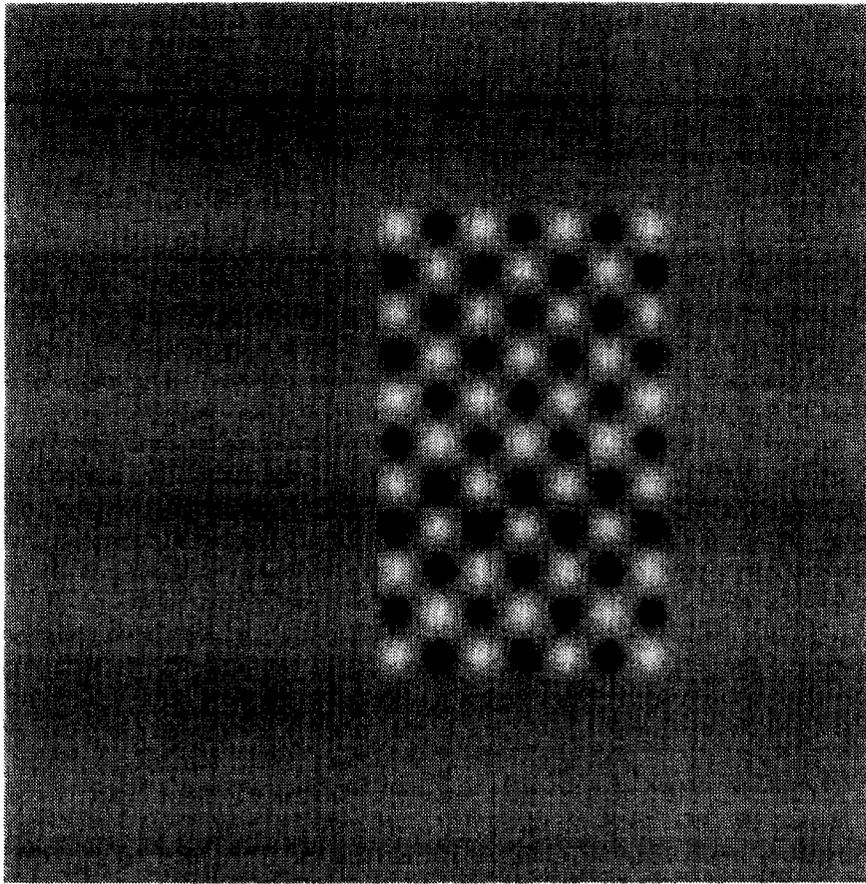
We measured subjects' ability to discriminate between two orientations (vertical and horizontal) of a rectangular patch of texture (either checked or striped arrangement) which was embedded in a background field of the other arrangement. The patterns were composed of either Gaussian blobs (for simple channels) or Gabor patches (for complex channels). We investigated the time course of processing for these two types of patterns by requiring the subjects to respond within 200 msec after an auditory cue which could occur at several different times (response cue lags) after the onset of the stimulus (which was 50 msec in duration).

Other investigators have chosen to control the processing of a stimulus by manipulating the duration of the stimulus and/or by using post-stimulus masking (see e.g. Arsenault & Wilkinson, 1993; Ben-Av *et al.*, 1992; Caelli & Julesz, 1978; Nothdurft, 1985b; Bergen & Julesz, 1983b). The effects of these manipulations are then determined by measurements of the outcome of processing, that is, of the final percept. One difference between those methods and the cued-response method used here is that by masking or varying the duration of the stimulus one limits the amount of information available for processing while allowing processing to proceed to its conclusion before taking a measurement. The SAT method, on the other hand, measures performance at various stages of completion of processing. We chose not to mask the stimulus, as other investigators have, because although it may indeed limit processing of the original stimulus, it does so by interference and replacement with a new stimulus (the mask), which might have its own time-dependent effects on processing of the original stimulus. By choosing not to use a post-stimulus mask, we allowed the possibility that on the trials with longer cue delays, subjects could use information in iconic storage, but this is part of the natural process of perception that we are investigating.

Rectangle-Only patterns—control for visibility

In order to be able to compare directly the processing dynamics for the Gaussian-blob and Gabor-patch patterns, we needed some way of equating the visibility of the two types of pattern. Simply equating the physical contrasts of the elements would not work because the spatial frequency composition of the two types of elements is different—the Gabor patches contain higher spatial frequencies than the Gaussian blobs. At equal physical contrasts, and especially at low contrasts, the Gabor-patch textures were frequently harder to see than the Gaussian-blob textures. To equate the Gaussian and Gabor patterns for visibility, we ran a control condition in which we presented subjects with blocks of

(a)



(b)

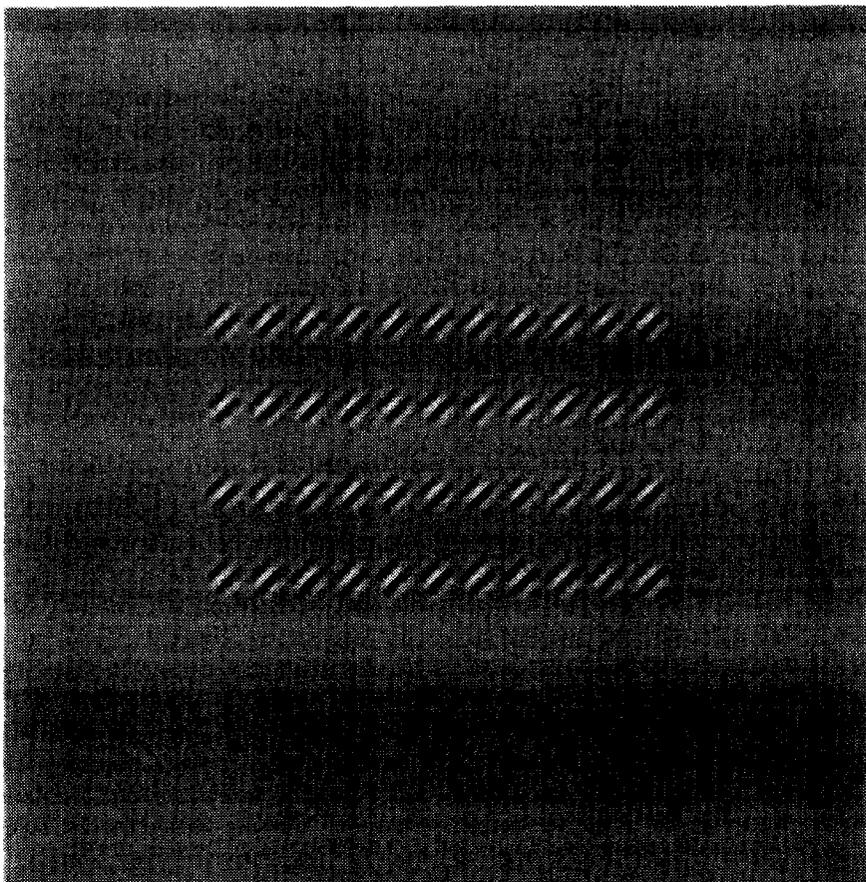


FIGURE 4. Two examples of Rectangle-Only patterns. (a) A pattern composed of Gaussian blobs of equal-but-opposite sign-of-contrast. (b) A pattern composed of Gabor patches luminance-balanced with the background. Reproduction will have distorted the images somewhat.

Rectangle-Only patterns. Examples of these patterns are presented in Fig. 4. The Rectangle-Only patterns were identical to the Embedded-Rectangle patterns except that they contained no background texture. Detection of the orientation of these rectangles without a background should be accomplished by simple channels, or one stage of filtering, precisely because there are no background elements. No segregation of regions of different element arrangements is necessary in these patterns, and any channel sensitive to the individual elements will signal the presence of the rectangle. We presented the patterns at several different contrasts (six for the Gabor and six others for the Gaussian patterns) in order to find contrasts at which the Gaussian and Gabor Rectangle-Only patterns produced identical, or nearly identical, performance across the entire measured time course. Once a pair of "matched contrasts" has been found for the Rectangle-Only control, one asks whether differences exist between the time courses for the Gaussian and Gabor Embedded-Rectangle patterns at these same contrasts. Any such differences may then be taken as reflecting differences in the operation of segregation mechanisms rather than differences in the visibility of the two types of patterns.

By matching performance of the Gabor and Gaussian patterns in the Rectangle-Only condition, we also controlled for the tendency of high spatial frequency stimuli (in this case the Gabor-patch patterns) to produce longer reaction times than stimuli of lower spatial frequency (in this case the Gaussian-blob patterns). Longer response times to higher spatial frequency stimuli have been found by several investigators (Breitmeyer, 1975; Tolhurst, 1975; Gish, Shulman, Sheehy & Leibowitz, 1986). Such a difference between high and low spatial frequency response times would potentially manifest itself in this experiment (everything else being equal) as a lower proportion of correct responses to the Gabor patterns than to the Gaussian-blob patterns at any given response cue lag. Matching the two types of pattern for performance in the Rectangle-Only condition controls for this difference between response times to high and low spatial frequency stimuli.

Patterns and channels

The Rectangle-Only patterns not only served to control for visibility but also to provide further information about the dynamics of processing of simple channels. In summary: we used four kinds of patterns, one of which stimulates complex channels and three of which stimulate simple channels.

The Embedded-Rectangle Gabor-patch pattern primarily stimulates complex channels (as explained earlier, see Fig. 3). The complex channels under investigation have first-stage filters that are sensitive to the spatial frequency and orientation of the Gabor patch elements (this spatial frequency was 8 c/deg and the orientation oblique). They have second-stage filters that are sensitive to the spatial frequency and orientation characterizing the arrangement of elements. (The spatial frequency equals the fundamental frequency of the pattern, which

was 1.5 c/deg in the striped region and $\sqrt{2}$ higher in the checked region; the orientation was vertical, horizontal or oblique depending on whether a vertically-striped, horizontally-striped or checked region was under consideration.)

The Rectangle-Only Gabor-patch pattern stimulates simple channels at the same spatial frequency and orientation as the first stage of the complex channels under investigation (i.e. at the spatial frequency and orientation of the Gabor patches).

The two Gaussian-blob patterns (Rectangle-Only and Embedded-Rectangle) both stimulate simple channels at the same spatial frequency and orientation as the second stage of the complex channels under investigation (i.e. at the spatial frequency and orientation characterizing the arrangement).

Thus we can compare low spatial frequency simple channels (the Embedded-Rectangle and Rectangle-Only Gaussian-blob patterns) with high spatial frequency simple channels (the Rectangle-Only Gabor-patch patterns) and the complex channel of high spatial frequency first stage and low spatial frequency second stage (Embedded-Rectangle Gabor-patch) can be compared to both simple channels.

METHOD AND PROCEDURES

Apparatus

The stimuli were presented on a standard Apple monochrome monitor. Stimulus generation, experimental control, and luminance linearization of the monitor were accomplished using a Macintosh IICI with Pascal programs based on software generously provided by Hugh Wilson.

Subjects

Three young adults participated in the experiment. One was an author (AS) and the other two were naive (CS and JH). The subjects had normal (AS and CS) or corrected-to-normal (JH) vision.

Stimuli

Examples of the stimuli appear in Figs 2 and 4. The patterns were composed of either Gaussian-blob or Gabor-patch elements. The elements were arranged into striped or checked texture regions, and appeared in either Embedded-Rectangle (Fig. 2) or Rectangle-Only (Fig. 4) patterns.

The arrangements of the elements. Forty-eight Embedded-Rectangle patterns were constructed, consisting of a field of either checks or stripes (17×17 elements, 6.12×6.12 deg) in which was embedded a rectangular patch (7×11 elements, 2.52×3.96 deg) of the other type of texture. The rectangular texture patch was oriented either vertically or horizontally. It occurred at one of three different locations in the background texture field—near the top, at the middle, or near the bottom for horizontal patches, and near the right edge, at the middle, or near the left edge for vertical patches. There

were always at least three rows or columns of background elements between the embedded-rectangle and the nearest edge of the pattern. There were four assignments of checked and striped arrangements into the rectangle and the background, namely: horizontally-striped rectangle with checked background; vertically-striped rectangle with checked background; checked rectangle with horizontally-striped background; and checked rectangle with vertically-striped background. Thus, the 48 Embedded-Rectangle patterns included all possible combinations of two rectangle orientations, three locations, two element types (Gaussian blob or Gabor patch) and four striped/checked assignments.

Thirty-six Rectangle-Only patterns were constructed, consisting of a rectangular patch (7×11 elements, 2.52×3.96 deg) of either checks or stripes. The rectangular texture patch was oriented either vertically or horizontally, and occurred at one of three different locations; top, middle, or bottom for horizontal patches, and right, middle, and left for vertical patches. The positions of these rectangles corresponded exactly with the positions of the rectangles in the Embedded-Rectangle patterns. The 36 patterns included all possible combinations of two rectangle orientations, three locations, two element types (Gaussian blob or Gabor patch) and three arrangements in the rectangular region (checked, horizontally striped, and vertically striped).

The elements. The Gaussian-blob patterns [Figs 2(a) and 4(a)] contained two types of elements which were of equal but opposite sign of contrast. The Gaussian blob elements had a concentric Gaussian window with a half-width half-height of 8 pixels (0.167 deg at viewing distance of 0.91 m). The elements were truncated at 16 pixels to avoid overlapping with adjacent elements. The center-to-center spacing between adjacent elements was 16 pixels (0.33 deg); thus the fundamental frequency of the striped texture was 1.5 c/deg (one period of the striped texture consisted of two elements, i.e. 32 pixels). These patterns are opposite-sign-of-contrast patterns in the terminology of Graham *et al.* (1992).

The Gabor patch patterns [Figs 2(b) and 4(b)] were composed of one type of element having a concentric Gaussian window with a half-width half-height of 8 pixels (0.167 deg at a viewing distance of 0.91 m). Each element was truncated at 16 pixels. The oscillation of the Gabor patch was in sine phase with respect to the Gaussian window so that the space-average luminance across the Gabor patch was equal to the background luminance. The spatial frequency of the Gabor patch elements was 8 c/deg (a period of six pixels), and the orientation of the patch was 45 deg oblique. The center-to-center spacing between adjacent elements was 32 pixels (0.67 deg); thus the fundamental frequency of the striped texture was 1.5 c/deg. The patterns composed of these Gabor patch elements may be thought of as patterns composed of two element types (like the Gaussian-blob patterns) but with the contrast of the second type of element set to zero. Thus, these patterns are one-element-only patterns in the terminology of Graham *et al.* (1992).

Luminance and contrast. The background luminance was 17.6 ft-L. Stimulus patterns were presented at six different contrasts. For the Gaussian blob patterns, the contrasts were 0.04, 0.08, 0.12, 0.15, 0.33 and 0.67. The contrasts for the Gabor patch patterns were 0.18, 0.26, 0.33, 0.41, 0.82 and 1.0.

Procedures

Subjects were seated 0.91 m from the CRT screen. They wore their normal vision correction if needed. Head movements were unrestrained, but subjects were instructed to avoid making them. The room was dimly illuminated by a partially obscured lamp on the floor behind the subjects.

A trial. Figure 5 illustrates the procedure for a single trial. The subject initiated a trial by pressing a key, after which the following sequence of events occurred: A fixation "X" (10% contrast, 0.17 deg wide \times 0.33 deg high) appeared on the screen for 1 sec and then was replaced by one stimulus pattern which was presented for 50 msec with abrupt onset and offset. After a variable interval of time following the onset of the stimulus (the *cue lag*, which was 50, 100, 150, 200, 250, 300, 500 or 800 msec), the subject received an auditory cue (a 50 msec beep). The subject was then required to respond within 200 msec after the cue onset by pressing one of two keyboard buttons depending on whether the texture rectangle was oriented horizontally or vertically. Immediately after the subject's response, auditory feedback indicated a correct response, an incorrect response, or a missed response deadline (a response occurring more than 200 msec after the cue onset). The subject was then free to initiate the next trial. Trials on which the response deadline was missed were re-randomized into the sequence of trials. During the periods between stimuli, the screen remained blank at the mean luminance of the stimulus patterns.

Structure of blocks and sessions. The experiment was run in sessions, each session consisting of two blocks of 192 trials. Embedded-Rectangle and Rectangle-Only patterns were presented in different sessions. Within each block, the 192 trials represented a complete crossing of two types of pattern element (Gabor patch or

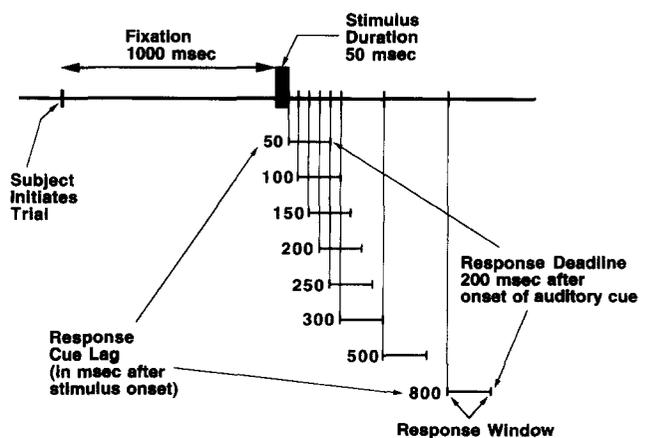


FIGURE 5. The cued response trial procedure.

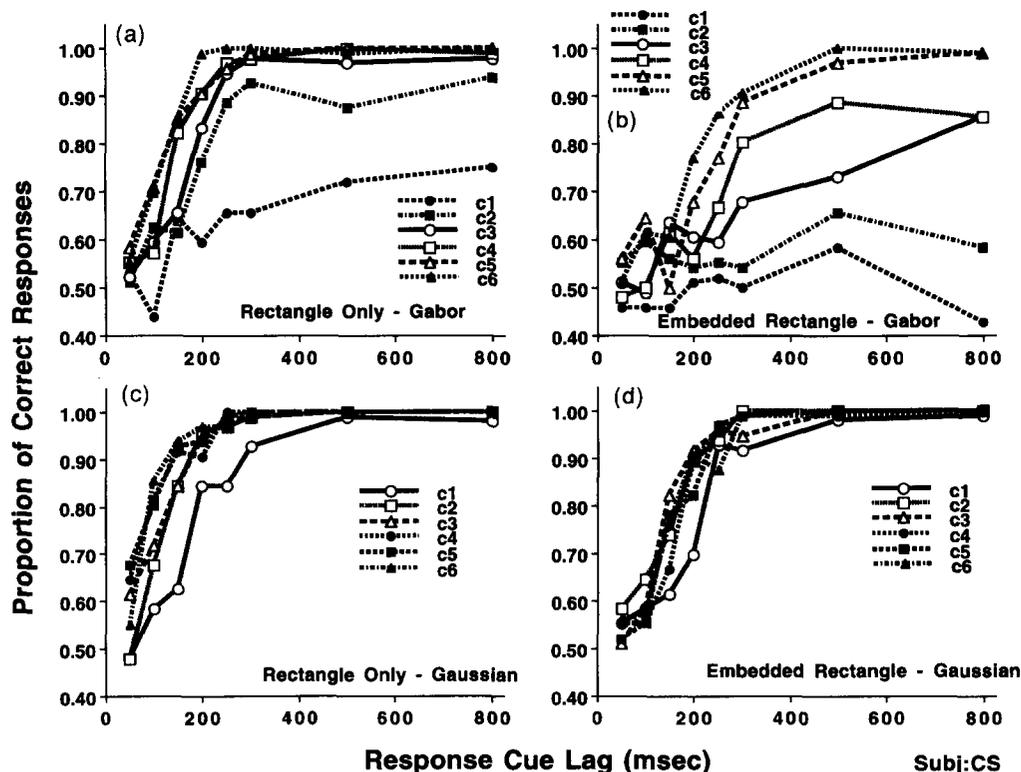


FIGURE 6. Proportion of correct responses as a function of cue lag (msec) for subject CS. (a) Results for Rectangle-Only patterns composed of Gabor patches. (b) Results for Embedded-Rectangle patterns composed of Gabor patches. (c) Results for Rectangle-Only patterns composed of Gaussian blobs. (d) Results for Embedded-Rectangle patterns composed of Gaussian blobs. Each panel contains six SAT functions (curves) representing results for the six different levels of contrast. Each point represents the proportion of correct responses in 100 trials. Open symbols indicate the functions for which contrast matches were obtained. These contrast-matched functions are replotted in Fig. 9.

Gaussian blob), two rectangle orientations (vertical or horizontal), six possible contrasts and eight cue lags. The other variables (three rectangle locations and four background/rectangle texture assignments) were not completely crossed with the other factors within each block, but combinations of these variables were counter-balanced within each block to avoid possible cues based on them. Trials were randomly intermixed.

Subjects completed between one and four sessions per day. Each session took 30–40 min to complete, and subjects took a break of at least 15 min between sessions. Each subject completed 25 sessions in each of the Embedded-Rectangle and Rectangle-Only conditions, subject AS over the course of 6 weeks, CS over the course of 5 weeks, and JH over the course of 14 weeks.

RESULTS

Figures 6–8 present SAT functions (proportion of correct responses as a function of cue lag) obtained from the three subjects (CS, AS and JS respectively). Results for patterns composed of Gabor-patch elements are presented in the top halves of Figs 6–8 and results for the Gaussian blob patterns are presented in the bottom halves of Figs 6–8. The right and left halves of Figs 6–8 present data for the Rectangle-Only and Embedded-Rectangle patterns respectively. Each panel contains six

SAT functions (curves) representing results for the six different levels of contrast as a function of the delay to the cue (cue lag). Each point in these figures represents the proportion of correct responses in 100 trials collapsed across all other variables not explicitly plotted (i.e. rectangle orientation, rectangle location and checked/striped assignment).

Several general characteristics of the results are easily seen in Figs 6–8. One feature shared by all of the curves is a positive, generally monotonic relationship between response accuracy and response cue lag. The form of these curves is characteristic of SAT functions in general (see Wickelgren, 1977). Second, there is a general tendency for performance to improve as pattern contrast increases, everything else being equal.

A quick inspection of the four panels of each figure reveals that performance is generally best for the patterns composed of Gaussian blobs (bottom halves of Figs 6–8). In addition, performance is also frequently good for Rectangle. Only patterns composed of Gabor patches (top left panels of Figs 6–8) especially when the contrast of the patterns is high. On the other hand, performance is generally worse for Embedded-Rectangle patterns composed of Gabor patches (top right panels of Figs 6–8). Thus, even before contrast-matching to equate visibility of the Gabor and Gaussian patterns, there are indications that patterns gated primarily through activity in simple channels (the Gabor and

Gaussian Rectangle-Only patterns, and Gaussian Embedded-Rectangle patterns) are segregated faster and better than the patterns segregated primarily through activity in complex channels (the Gabor Embedded-Rectangle patterns).

An estimate of the "immediacy" of texture segregation can also be found in Figs 6–8. For the fastest conditions involving a region segregation (lower right panels) discriminating the orientation of the Gaussian blob Embedded-Rectangle at high contrast, performance did not reach 75% correct until the cue signaling the observer to respond came at least 120 msec (for AS and CS) or 180 msec after stimulus onset (for JH). Asymptotic performance was not reached until the cue came at least 250–300 msec after stimulus onset. (The actual times at which the response was initiated by the observer may have been even later. The buttons tended to be pushed 150–200 msec after the cue.) For the slower conditions involving a region segregation (lower contrast Gaussian-blob Embedded-Rectangle, all contrasts Gabor-patches Embedded-Rectangle), performance did not reach 75% correct until the response cue came many hundreds of milliseconds after stimulus onset.

Also visible in these figures (although perhaps easier to see in the later curve-fitting results) is the following comparison: the times for even the fastest region segregation (the Embedded-Rectangle Gaussian-blob) are 50–100 msec longer than those for the corresponding

Gaussian-blob Rectangle-Only conditions in which the observer did not have to segregate two regions of texture. See Discussion for the theoretical implications of this comparison. Whether these times correspond to one's ideas of "immediate" may depend on one's history.

Match contrast conditions

Recall that in order to equate the visibility of Gaussian and Gabor patterns we presented the patterns at several different contrasts (six for Gabor and six others for Gaussian patterns). This allows us to match the contrast levels producing roughly equal performance in the Gaussian and Gabor Rectangle-Only patterns. In other words, contrast levels for the two types of pattern (Gabor and Gaussian Rectangle-Only) were considered a match (i.e. equally visible) if they produced SAT functions that juxtaposed well. The matching of performance between pairs of contrast conditions was done by eye, without knowledge of the corresponding performance for Embedded-Rectangle patterns of the same contrasts. Matches were determined for each subject separately and are shown in Table 1. Three matched-contrast conditions were obtained for subject CS, four for subject AS, and two for subject JH. As can be seen from Table 1, the contrasts of the Gabor-patch patterns had to be much higher than the contrasts of the Gaussian-blob patterns in order for the two types of patterns to produce equal performance.

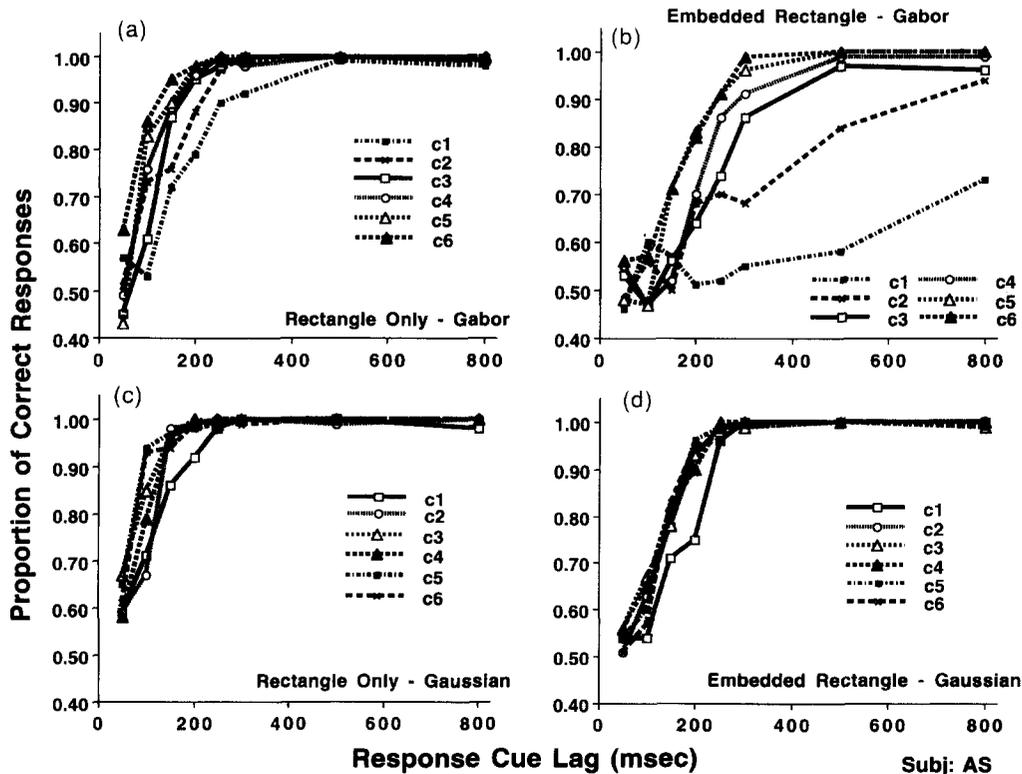


FIGURE 7. Proportion of correct responses as a function of cue lag (msec) for subject AS. (a) Results for Rectangle-Only patterns composed of Gabor patches. (b) Results for Embedded-Rectangle patterns composed of Gabor patches. (c) Results for Rectangle-Only patterns composed of Gaussian blobs. (d) Results for Embedded-Rectangle patterns composed of Gaussian blobs. Each panel contains six SAT functions (curves) representing results for the six different levels of contrast. Each point represents the proportion of correct responses in 100 trials. Open symbols indicate the functions for which contrast matches were obtained. These contrast-matched functions are replotted in Fig. 10.

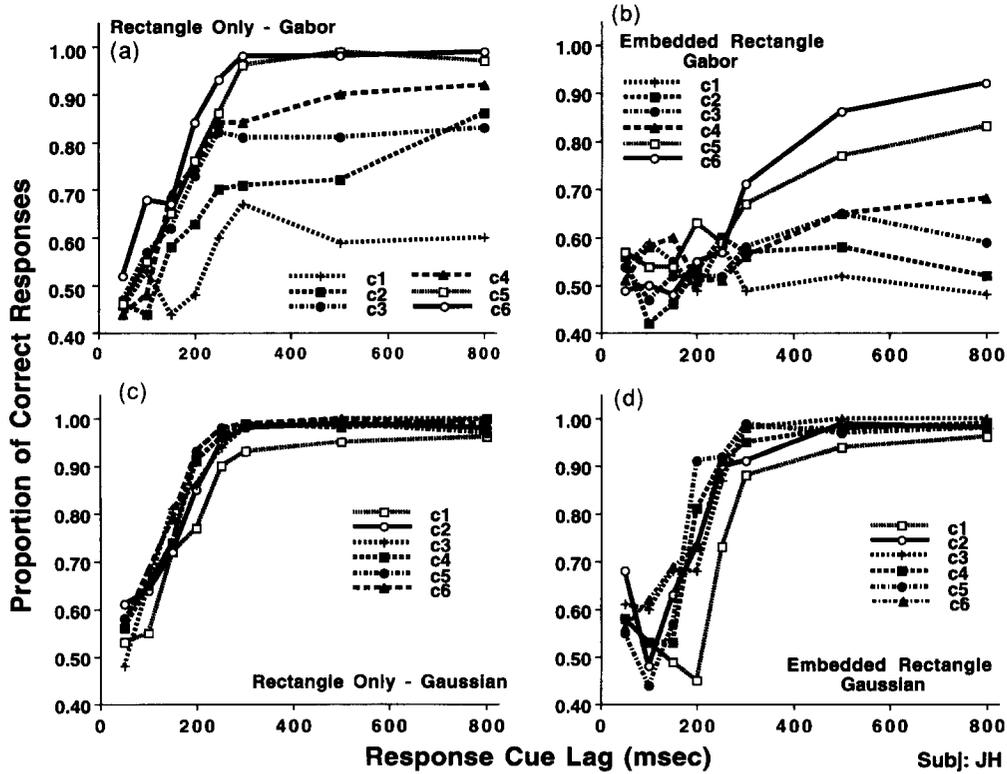


FIGURE 8. Proportion of correct responses as a function of cue lag (msec) for subject JH. (a) Results for Rectangle-Only patterns composed of Gabor patches. (b) Results for Embedded-Rectangle patterns composed of Gabor patches. (c) Results for Rectangle-Only patterns composed of Gaussian blobs. (d) Results for Embedded-Rectangle patterns composed of Gaussian blobs. Each panel contains six SAT functions (curves) representing results for the six different levels of contrast. Each point represents the proportion of correct responses in 100 trials. Open symbols indicate the functions for which contrast matches were obtained. These contrast-matched functions are replotted in Fig. 11.

Figures 9–11 present the matched-visibility SAT functions for subjects CS, AS and JH respectively. Each panel presents one matched-contrast condition. Here the horizontal axis shows response time rather than cue lag (as in Figs 6–8). Each of these response times equals the average response time for 100 trials per point at a given cue lag, and most of them are 150–185 msec longer than their respective cue lags. Because subjects were required to wait for the auditory cue before responding, and were then required to respond within 200 msec, these response times are not the same as typical reaction time measures (which are usually not constrained by response cues or

deadlines). We present the results (and the curve fitting to follow) in this form because processing is more easily thought of in terms of these response times than in terms of cue lag times.

The four curves in each panel of Figs 9–11 are the SAT functions for the contrast-matched Gabor and Gaussian Rectangle-Only patterns (thin lines), and the corresponding Gabor and Gaussian Embedded-Rectangle patterns (thick lines). The thin dashed line shows performance for the Gabor-patch Rectangle-Only patterns, and the thin solid line shows performance for the Gaussian-blob Rectangle-Only patterns. It can be seen that for each matched visibility condition displayed in Figs 9–11, these two curves superimpose or are very close to each other, indicating that the chosen contrast matches were good ones.

Simple vs complex channel patterns

The most important feature of Figs 9–11 are the locations, relative to the other three curves in each panel, of the SAT curves for the Gabor-patch Embedded-Rectangle patterns (thick dashed lines), which presumably reflect primarily the action of complex channels. (Recall that the Gaussian-blob Embedded-Rectangle patterns are segregated primarily by simple channels, just as the Rectangle-Only patterns are.) In every panel of Figs 9–11 performance is worst for the patterns reflecting the action of complex channels: it takes longer to reach a given level of accuracy, and in

TABLE 1. Contrasts of pairs of Rectangle-Only Gabor-patch and Rectangle-Only Gaussian-blob patterns that were matched for equal performance

Subject	Match	Contrast	
		Gaussian	Gabor
JH	1	0.04	0.82
	2	0.08	1.00
CS	1	0.04	0.33
	2	0.08	0.41
	3	0.12	0.82
AS	1	0.04	0.33
	2	0.08	0.41
	3	0.12	0.82
	4	0.15	1.00

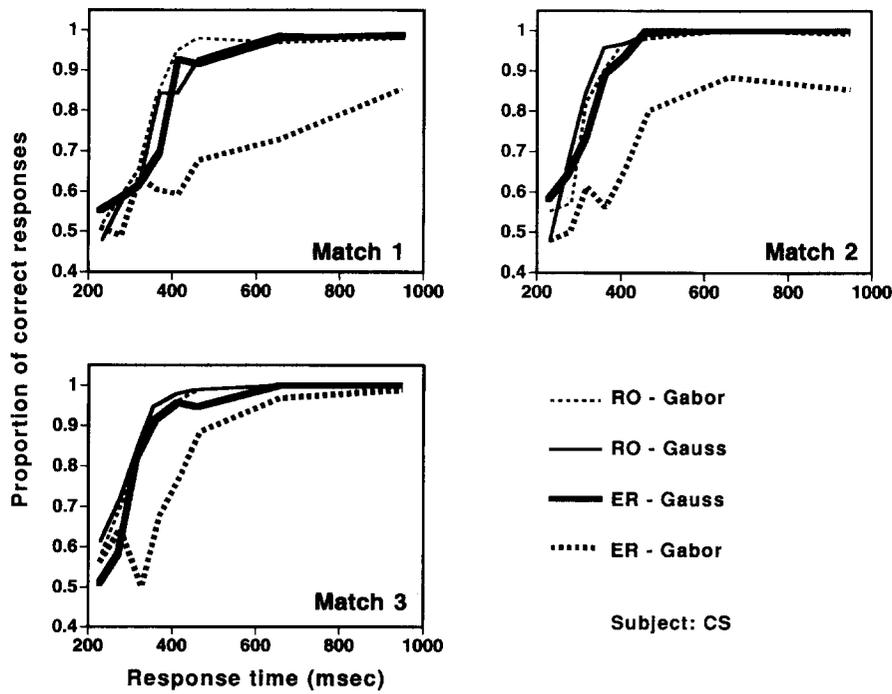


FIGURE 9. Three matched-visibility conditions obtained for subject CS (see Table 1). Proportion of correct responses is plotted as a function of response time. *Thin lines*: curves for SAT functions for the contrast-matched Gabor (dashed line) and Gaussian (solid line) Rectangle-Only patterns. *Thick lines*: curves for corresponding Gabor (dashed line) and Gaussian (solid line) Embedded-Rectangle patterns. Curves are replotted from Fig. 6.

some panels, performance never reaches an asymptote or even a high level of accuracy even after 900 msec. One other trend is also visible—the performance on the Embedded-Rectangle Gaussian-blob patterns (solid

thick curve) seems to be worse than performance on both Rectangle-Only patterns. These comparisons of simple vs complex channel action will be discussed further.

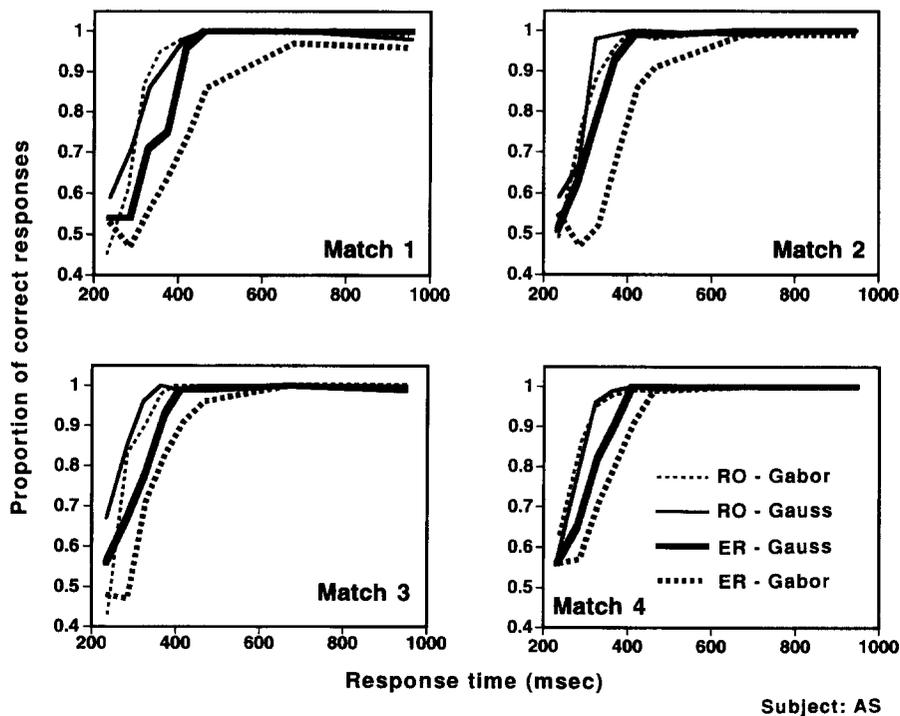


FIGURE 10. Four matched-visibility conditions obtained for subject AS (see Table 1). Proportion of correct responses is plotted as a function of response time. *Thin lines*: curves for SAT functions for the contrast-matched Gabor (dashed line) and Gaussian (solid line) Rectangle-Only patterns. *Thick lines*: curves for corresponding Gabor (dashed line) and Gaussian (solid line) Embedded-Rectangle patterns. Curves are replotted from Fig. 7.

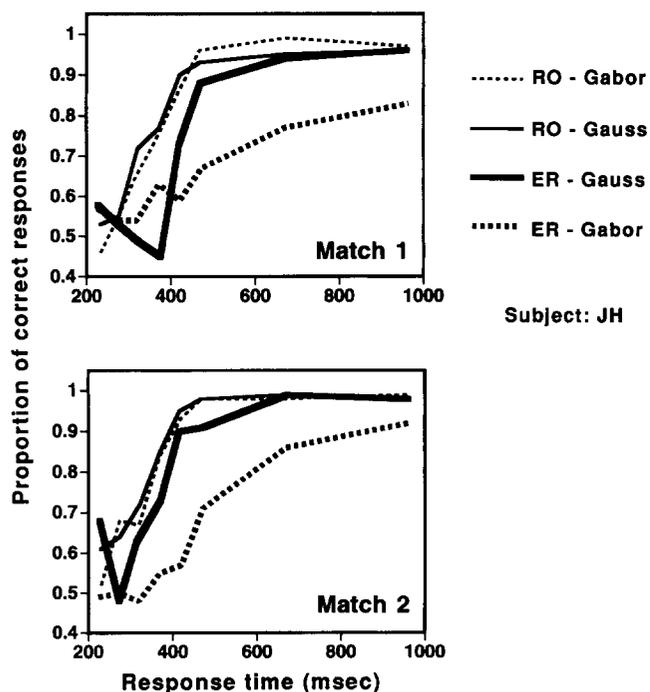


FIGURE 11. Two matched-visibility conditions obtained for subject JH (see Table 1). Proportion of correct responses is plotted as a function of response time. *Thin lines*: curves for SAT functions for the contrast-matched Gabor (dashed line) and Gaussian (solid line) Rectangle-Only patterns. *Thick lines*: curves for corresponding Gabor (dashed line) and Gaussian (solid line) Embedded-Rectangle patterns. Curves are replotted from Fig. 8

Curve fitting

To quantify the differences among conditions, we fit a function to each of the SAT functions in Figs 9–11 (to the proportion of correct responses as a function of response time). At this point, we do not strongly suggest an interpretation of this function in terms of any underlying process, but use it primarily as an objective and concise means of summarizing the data. The function we used is a delay followed by an exponential approach to a limit (see inset of Fig. 12). It is horizontal at chance performance (probability correct = 0.50) for a period of time (called *delay*), and then the function rises exponentially to a maximum value (called R_{max}). We used the algebraic form:

$$F(t) = 0.5 \text{ for } t \leq \text{delay}$$

$$= R_{max} \cdot [1 - e^{-\alpha \cdot (t - t_0)}] \text{ for } t \geq \text{delay} \quad (1)$$

where α and t_0 are the rate and intercept parameters respectively, of the exponential function describing the rise from chance to maximal value. The value of delay (the time at which the exponential function crosses 0.50) is related to the other three parameters:

$$\text{delay} = t_0 - \frac{\ln\left(1 - \frac{0.5}{R_{max}}\right)}{\alpha}$$

The function F was fit separately to the 36 different data sets (nine matched-contrast sets, each containing four conditions) in Figs 9–11. The function $F(t)$ turned out

overall to be an excellent description of our data, as shown in the example in Fig. 12, which displays the best fits (lines) and the proportion of correct responses (symbols) for the third matched-contrast condition for subject CS (Fig. 9, Match 3). See the Appendix for further discussion of the curve-fitting procedure and the goodness of the fits.

Quick-Weibull fits. In addition to fitting the above function, we fit a Quick-Weibull function to the 36 data sets. The fitted Quick-Weibull functions were extremely similar to the functions we report here and the goodness of fit was equally good. However, the actual values of the parameters were less clearly related to features of the data (in particular to the delay evidenced before performance rises from chance). Therefore, we will restrict our discussion to the fits produced by the above function which incorporate a delay followed by a rise.

The rate, delay and asymptote parameters

Figure 13 summarizes the result of curve-fitting by showing, for the fitted functions $F(t)$, the response time at which performance reaches 75% accuracy [$F(t) = 0.75$] for all conditions (horizontal axes of Fig. 13) and subjects (different panels) for the matched-contrast sets (different curves) shown earlier in Figs 9–11. Notice that in all the curves, the time for the Embedded-Rectangle task with the Gabor-patch patterns (the condition involving complex channels) was the slowest, the time for the Embedded-Rectangle task with the Gaussian pattern was the second longest, while the times for the Rectangle-Only conditions were very much the same (as expected since they had been chosen as such in the process of matching for visibility).

Furthermore, Fig. 13 shows a general effect of contrast, with performance for patterns of relatively lower contrasts taking longer to reach 75% accuracy than similar patterns of higher contrast. This contrast effect

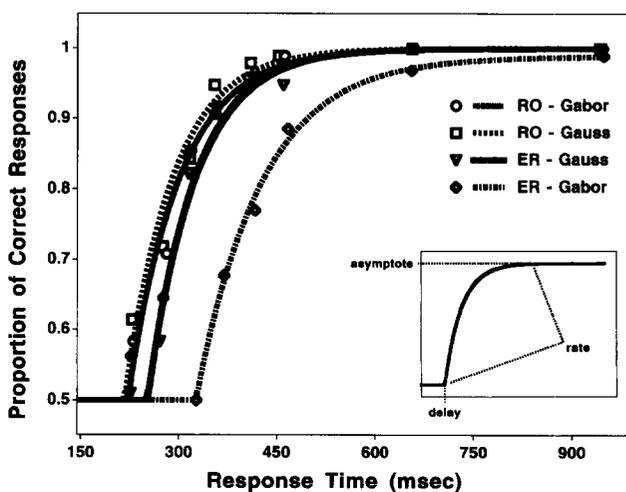


FIGURE 12. Example of functions of form of equation (1), fit to the set of four matched-contrast curves for subject CS, contrast Match 3 (Fig. 9, Match 3). *Symbols*: data points from subject CS, Match 3. *Lines*: functions fitted to the data points. *Inset*: a delayed-exponential function of the form of equation (1) with parameters marked.

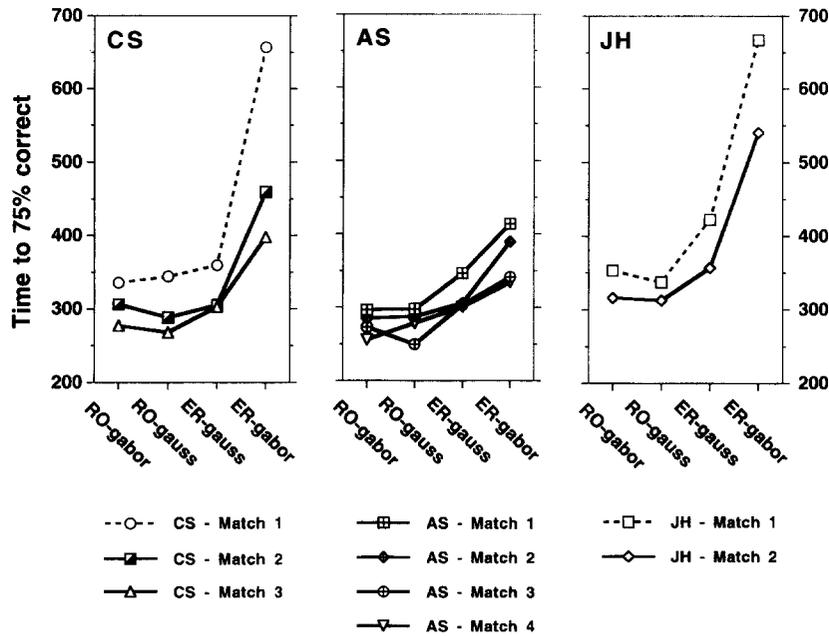


FIGURE 13. Results of curve fitting. Response time at which performance reaches 75% accuracy [$F(t) = 0.75$] for all stimulus classes (horizontal axes) and subjects (different panels) for the matched-contrast sets (different curves) shown in Figs 9–11.

is most noticeable for the naive subjects (JH and CS).

Figure 14 displays, in separate panels, the rate (α), delay, and asymptote (R_{max}) parameters, as a function of the stimulus class, for the 36 fitted functions $F(t)$. The nine curves in each panel represent the nine matched-contrast conditions. It can be seen in the first panel [Fig. 14(a)] that for all nine matches, the rate (α) parameters for Embedded-Rectangle Gabor patterns

(involving complex channels) are consistently lower than rates for patterns involving only simple channels (Rectangle-Only Gabor, Rectangle-Only Gaussian, and Embedded-Rectangle Gaussian patterns). The two dashed curves are somewhat misleading, however, as discussed below.

Figure 14(b) displays the delay parameters for all nine matched-contrast conditions. The delay parameter marks the response time at which each fitted function

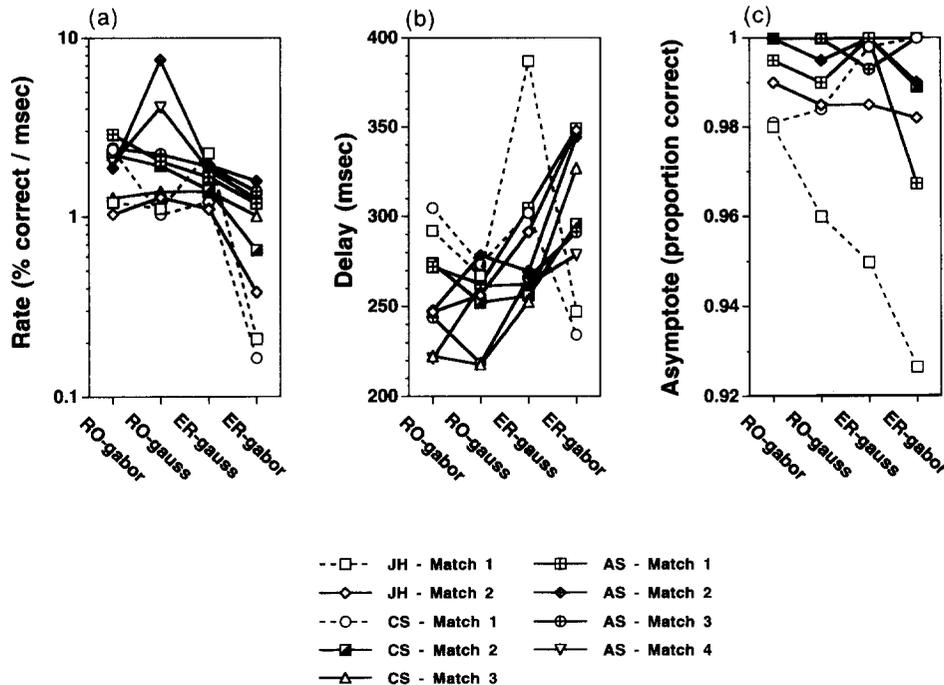


FIGURE 14. Results of curve fitting. Values of (a) the rate (α) (b) the *delay* and (c) the asymptote (R_{max}) parameters, as a function of the stimulus class, for the 36 fitted functions $F(t)$. The nine curves in each panel represent the nine matched-contrast conditions depicted in Figs 9–11.

departs from 50% accuracy. Consider first the seven solid curves: the delay is longest for Embedded-Rectangle Gabor (complex) patterns; the delays for the three other patterns are similar although there is a tendency for the Embedded-Rectangle Gaussian delay to be longer than those for the two Rectangle-Only patterns as discussed further below.

However, two matched-contrast sets (represented by dotted lines in Figs 13 and 14) do not show the same results for the delay parameter: the delay parameters for the Embedded-Rectangle Gabor (complex) patterns from these two matched-contrast sets are much lower than for the other seven matched-contrast sets. These are the two matched sets at the lowest contrast for subjects JH and CS. Looking back at the SAT curves for these two matched sets, one sees that the SAT function for the Embedded-Rectangle Gabor patterns is very shallow as it rises out of the 50% chance baseline (thick, dashed lines in Fig. 9, Match 1 panel, and in Fig. 11, Match 1 panel). With a shallow function, accurate estimation of the delay and rate parameters is not possible because the effects of these parameters interact so greatly. The effect of shortening the delay parameter can be counteracted by decreasing the rate parameter so that the resulting function (and hence its fit to the data) will remain much the same. Consistent with this, notice that not only were the fitted delay parameters unusually short for those two functions, but also that the fitted rate parameters were unusually low [see the rightmost points of the dashed curves in Fig. 14(b) and (c)]. Increasing the rate parameter and decreasing the delay parameter over wide ranges would still produce good fits to these two functions. Some of these fits would make the two dashed curves in Fig. 14(b, c) like the solid curves. Since any of a wide range of parameters fit these two SAT curves equally well, however, these two conditions are not informative about the delay and rate parameters which is why they are shown as dashed lines in Fig. 14. Note, however, that all acceptable fits will predict much the same response-time-to-75%-correct (plotted on the vertical axis of Fig. 13) since the 75% point is on a steeper part of the SAT function. Hence, although the curves are also shown as dashed in Fig. 13 for uniformity's sake, they can be considered as seriously as the solid lines.

Figure 14(c) displays the asymptote (R_{\max}) parameters for all nine matched-contrast conditions. The asymptote parameter marks the proportion of correct responses, $F(t)$, at which performance levels off. Notice that even though an asymptote may be determined by curve-fitting, it is not necessarily true that this asymptote is reached by 1000 msec (the maximum response time allowed in this experiment). Most of the fitted asymptotes are above 0.98.

DISCUSSION

In the experiment reported here, we explored the time course of texture segregation processes in simple (Fourier, first-order) and complex (non-Fourier, second-

order) channels. Using the method of cued response, we obtained speed-accuracy tradeoff SAT functions for patterns segregated primarily through activity in one type of channel or the other (after choosing contrasts so that the patterns were matched for visibility). The SAT functions (proportion accuracy vs response time) were well described by a function combining a delay with an exponential approach to a limit.

Simple vs complex channels

The SAT functions for the Embedded-Rectangle Gabor-patch patterns (segregated primarily by complex channels) were in general slower than those for the other three kinds of pattern (segregated primarily by simple channels). The time-to-75%-correct was longer by 100–200 msec (Fig. 13) depending on observer and contrast. This lengthening seemed to come both from a slower rate [Fig. 14(a)] and a lengthening delay [Fig. 14(b)].

Other information about the dynamics of simple vs complex channels

Many investigators of texture and motion perception (e.g. Chubb & Sperling, 1988; Sperling, 1989; Wilson *et al.*, 1992; Yo & Wilson, 1992; Lin & Wilson, 1994) have invoked similar nonlinear, two-stage processing mechanisms in their work. Several of these investigators have also found evidence that processing involving non-Fourier (complex, second-order) mechanisms may take longer than processing involving Fourier (simple, first-order) mechanisms. Lin and Wilson (1994) compared pattern discrimination performance with Fourier and non-Fourier stimuli (Gaussian windowed, sixth derivative of Gaussians (D6s) and cosine gratings contrast-modulated by D6s). They found that orientation and spatial frequency discriminations with Fourier stimuli were unaffected by changes in stimulus duration (33.3, 100 or 500 msec). However, performance with the non-Fourier stimuli was poorer at the shorter stimulus durations.

Wilson *et al.* (1992) and Yo and Wilson (1992) investigated the perception of motion in two-dimensional patterns composed of cosine gratings (plaids). They presented observers with patterns in which motion was signalled in two different directions, one direction resulting from processing by a Fourier mechanism and the other direction resulting from processing by a non-Fourier mechanism. At short stimulus durations (less than 60 msec), subjects perceived motion in a direction consistent with processing by Fourier (first-order, simple) mechanisms. At longer stimulus durations (around 140 msec), the perceived direction of motion was consistent with processing by non-Fourier (second-order, complex) mechanisms. Similarly, Derrington and his colleagues (Derrington, Badcock & Holroyd, 1992; Derrington, Badcock & Henning, 1993) investigated the perceived direction of motion of various spatial patterns. They found that the temporal resolution of second-order (non-Fourier, complex) motion-detection mechanisms was poorer than that of first-order (Fourier, simple)

mechanisms. They determined that second-order (non-Fourier, complex) mechanisms did not contribute to perceived direction of motion when stimulus durations were less than 100 msec.

One difference between our results and those of the investigations described above is that we found evidence that even at short stimulus durations (all of our patterns were presented for 50 msec) complex channels (second-order, non-Fourier processes) led, albeit slowly, to reasonably accurate performance in most cases. In the motion literature, it appears that a minimum stimulus duration of about 100 msec is necessary for second-order processes to have any influence on perception. This difference between motion and texture perception is not too surprising considering the time-dependent nature of motion.

Possible explanation of a methodological problem

When we initially tried to replicate some experiments reported by Graham *et al.* (1992), we were unable to fully replicate the effects in those experiments which suggested complex channels. (These were rating experiments in which the observer simply had to indicate, after each trial, the extent to which the regions appeared to segregate perceptually.) After many months of tedious comparisons, the following seemed to be the crucial difference: in Graham *et al.*'s experiments, due to limitations in the equipment, the subjects had had to write down their responses after each trial, thereby introducing a 1- or 2-sec delay between the stimulus and the actual response. In the attempted replication, the procedure was entirely computerized and the subjects just pushed a button, which they did much faster than they could have written a response. Indeed, in the automated procedure, they raced through trials four or five times faster than in the Graham *et al.* experiments. We then tried requiring a 1-sec delay between the end of the stimulus and the subject's button push. This enforced delay re-established the signature of the complex channels where it had been weak. Several possible reasons for this effect were briefly discussed in Graham *et al.* (1993). The reason that now appears most likely to us is supported by the SAT study here. Perhaps observers who can go as fast as they want (as they were initially allowed to do in the attempted replication) answer as soon as any information is available. Then, according to the results of this SAT experiment, they will tend to answer on the basis of simple-channel output before the complex-channel output is fully available. Thus any effect of complex channels will be attenuated in their responses. When a 1-sec delay is enforced before responding, however, they will have both simple- and complex-channel outputs available at the time of their response and effects of complex channels will be visible.

Edge effects

It is very unlikely that edge effects (information particular to the boundaries between the texture regions) have contributed in any significant way to the results

reported here. Sutter *et al.* (1989) described the results of experiments in which they employed stimulus patterns very similar to the ones presented in the present study, i.e. checked-vs-stripped texture regions. Sutter *et al.* presented the results and a discussion of extensive filtering of their patterns, and over a wide range of spatial frequencies and orientations found no responses at the boundaries sufficient to support segregation of the regions.

Embedded-Rectangle vs Rectangle-Only

Three SAT functions in each matched set presumably reflect the action of simple channels (those for the Embedded-Rectangle Gaussian-blob, Rectangle-Only Gaussian-blob, and Rectangle-Only Gabor-patch patterns). The functions for the two Rectangle-Only patterns were quite similar to one another, as expected since the matching was based on performance in that condition. However, the function for the Embedded-Rectangle Gaussian-blob pattern seems to be slightly slower than either of the Rectangle-Only cases. In particular, the response time to 75% correct is longer in nine out of nine cases (Fig. 13); the fitted delay parameter is longer in six of the seven reliable cases [Fig. 14(b)] and the fitted rate parameter is slower in six of the seven reliable cases [Fig. 14(a)].

Within the framework of our current model, this difference between the two Gaussian-blob patterns (Embedded-Rectangle and Rectangle-Only) presents a problem since both ought to be primarily segregated by the same simple channels (the simple channels sensitive at the fundamental frequency of the pattern). To understand this problem, consider a concrete example: a vertically-stripped rectangle of Gaussian blobs embedded in a checkerboard background (Embedded-Rectangle) compared to a vertically-stripped rectangle of Gaussian blobs surrounded by a blank screen (Rectangle-Only). The overall orientation of the rectangle is irrelevant so it will be ignored. In the Rectangle-Only version, one set of simple channels (those sensitive to a vertical orientation and to the fundamental frequency of the striped arrangement) have a large output inside the rectangle and zero output outside. All other channels have zero output everywhere. What happens in the Embedded-Rectangle case? The same set of simple channels (those sensitive to a vertical orientation and the fundamental frequency) again have a large output inside the rectangle and zero output outside [at least by current estimates of the orientation bandwidth of simple channels, e.g. Graham *et al.* (1993)]. Now there is a second set of simple channels (those sensitive to an oblique orientation and to the fundamental frequency) which are active; they have a large output outside the rectangle and zero input inside, and if anything, they ought to augment the perceived segregation of the regions. The Embedded-Rectangle Gaussian-blob segregation ought to be better, if anything, than the Rectangle-Only case, opposing the trend in the results.

There are several possible explanations of the superior segregation for the Rectangle-Only relative to

the Embedded-Rectangle Gaussian-blob patterns. Perhaps the orientation bandwidth of the simple channels operating here is substantially broader than previously estimated [Graham *et al.* (1993) were working at much higher spatial frequencies]. Then in the Embedded-Rectangle case, the channels that are sensitive to the pattern inside the rectangle would also respond outside (and vice versa) and thus their ability to signal the differences that lead to segregation would be reduced. A second possibility arises from the possible existence of cross-channel inhibition, perhaps as modelled by a normalization network that we have proposed to explain other texture results (Graham, 1991; Graham *et al.*, 1992; Graham & Sutter, 1995). Indeed, even if there were only a small subset of channels that responded somewhat to both orientations (perhaps channels tuned to an intermediate orientation), these other channels would inhibit (or "mask") the response of the channels signalling segregation. The latter scenario brings up a third possibility. Perhaps, even though the orientation bandwidth is so narrow that no channel at all responds either inside or outside the rectangle, there is inhibition that occurs with some substantial spatial spread (see e.g. Sagi, 1990). Then the fact that some channels are responding outside the rectangle (in the Embedded-Rectangle case) might reduce the responses of the channels inside the rectangle (even though they are channels sensitive to different spatial frequencies and orientations) and thus reduce perceived segregation (relative to the Rectangle-Only case). A final possibility, vague but entirely reasonable, is that the Embedded-Rectangle vs Rectangle-Only segregation comparison reveals a level of visual processing that is not well represented in the framework of our current model (one of the inevitable multitude of processes hidden in our simple decision-rule stage).

REFERENCES

- Arsenault, A. S. & Wilkinson, F. E. (1993). A temporal analysis of three texture segregation tasks. *Investigative Ophthalmology and Visual Science (Suppl.)*, 34, 1237.
- Beck, J. (1982). Texture segregation. In Beck, J. (Ed.), *Organization and representation in perception* (pp. 285–317). Hillsdale, NJ: Erlbaum.
- Beck, J., Prazdny, K. & Rosenfeld, A. (1983). A theory of textural segmentation. In Beck, J., Hope, B. & Rosenfeld, A. (Eds), *Human and machine vision* (pp. 1–38). New York: Academic Press.
- Beck, J., Sutter, A. & Ivry, R. (1987). Spatial frequency channels and perceptual grouping in texture segregation. *Computer Vision, Graphics and Image Processing*, 37, 299–325.
- Ben-Av, M. B., Sagi, D. & Braun, J. (1992). Visual attention and perceptual grouping. *Perception & Psychophysics*, 52, 277–294.
- Bergen, J. R. (1991). Theories of visual texture perception. In Regan, D. (Ed.), *Vision and visual dysfunction, Vol. 10B: Spatial vision* (pp. 114–134). New York: Macmillan.
- Bergen, J. R. & Julesz, B. (1983). Rapid discrimination of visual patterns. *IEEE Transactions on Systems, Man and Cybernetics*, 13, 857–863.
- Bergen, J. R. & Landy, M. S. (1991). Computational modeling of visual texture segregation. In Landy, M. S. & Movshon, J. A. (Eds), *Computational models of visual processing* (pp. 253–271). Cambridge, MA: MIT Press.
- Bovik, A. C., Clark, M. & Geisler, W. S. (1987). Computational texture analysis using localized spatial filtering. *Proceedings of Workshop on Computer Vision*, Miami Beach, FL, 30 November–2 December, 1987 (pp. 201–206). The IEEE Computer Society Press.
- Breitmeyer, B. G. (1975). Simple reaction time as a measure of the temporal response properties of transient and sustained channels. *Vision Research*, 15, 1411–1412.
- Caelli, T. M. (1988). An adaptive computational model for texture segmentation. *IEEE Transactions on Systems, Man and Cybernetics*, 18, 1.
- Caelli, T. M. & Julesz, B. (1978). On perceptual analyzers underlying visual texture discrimination: Part I. *Biological Cybernetics*, 28, 167–175.
- Chubb, C. & Sperling, G. (1988). Processing stages in non-Fourier motion perception. *Investigative Ophthalmology and Visual Science (Suppl.)*, 29, 266.
- Clark, M., Bovik, A. C. & Geisler, W. S. (1987). Texture segmentation using a class of narrowband filters. *Proceeding of the IEEE International Conference on Acoustics, Speech and Signal Processing*. Dallas, TX, April 6–9, 1987.
- Derrington, A. M., Badcock, D. R. & Henning, G. B. (1993). Discriminating the direction of second-order motion at short stimulus durations. *Vision Research*, 33, 1785–1794.
- Derrington, A. M., Badcock, D. R. & Holroyd, S. A. (1992). Analysis of the motion of 2-dimensional patterns: Evidence for a second-order process. *Vision Research*, 32, 699–707.
- Doshier, B. (1976). The retrieval of sentences from memory: A speed-accuracy study. *Cognitive Psychology*, 8, 291–310.
- Doshier, B. (1979). Empirical approaches to information processing: Speed-accuracy tradeoff functions or reaction time—a reply. *Acta Psychologica*, 43, 347–359.
- Fogel, I. & Sagi, D. (1989). Gabor filters as texture discriminators. *Biological Cybernetics*, 61, 103–113.
- Gish, K., Shulman, G. L., Sheehy, J. B. & Leibowitz, H. W. (1986). Reaction times to different spatial frequencies as a function of detectability. *Vision Research*, 26, 745–747.
- Graham, N. (1991). Complex channels, early local nonlinearities, and normalization in perceived texture segregation. In Landy, M. S. & Movshon, J. A. (Eds), *Computational models of visual processing*. Cambridge, MA: MIT Press.
- Graham, N. & Sutter, A. (1995). Effect of spatial scale and background luminance on the spatial and intensive nonlinearities in texture segregation. *Vision Research*. Submitted.
- Graham, N., Beck, J. & Sutter, A. (1992). Nonlinear processes in spatial-frequency channel models of perceived texture segregation: Effects of sign and amount of contrast. *Vision Research*, 32, 719–743.
- Graham, N., Sutter, A. & Venkatesan, C. (1993). Spatial-frequency- and orientation-selectivity of simple and complex channels in region segregation. *Vision Research*, 33, 1893–1911.
- Julesz, B. (1975). Experiments in the visual perception of texture. *Scientific American*, 232, 34–43.
- Landy, M. & Bergen, J. (1991). Texture segregation and orientation gradient. *Vision Research*, 31, 679–691.
- Lappin, J. S. & Disch, K. (1972). The latency operating characteristic: II. Effects of visual stimulus intensity on choice reaction time. *Journal of Experimental Psychology*, 93, 367–372.
- Lin, L. & Wilson, H. (1994). Fourier and non-Fourier pattern discrimination compared. *Investigative Ophthalmology and Visual Science (Suppl.)*, 35, 1664.
- Malik, J. & Perona, P. (1989a). *A computational model of texture perception*. Report No. UCB/CSD 89/491. Computer Science Division (EECS), University of California, Berkeley, CA 94720.
- Malik, J. & Perona, P. (1989b). Computational model of human texture perception. *Investigative Ophthalmology and Visual Science (Suppl.)*, 30, 161.
- Malik, J. & Perona, P. (1990). Preattentive texture discrimination with early vision mechanisms. *Journal of the Optical Society of America A*, 7, 923–932.
- Nothdurft, H. C. (1985a). Orientation sensitivity and texture segregation in patterns with different line orientation. *Vision Research*, 25, 551–560.

- Nothdurft, H. C. (1985b). Sensitivity for structure gradient in texture discrimination tasks. *Vision Research*, 25, 1957-1968.
- Press, W. H., Flannery, B. T., Teukolsky, S. A. & Vetterling, W. T. (1986). *Numerical recipes*. Cambridge: Cambridge University Press.
- Reed, A. V. (1973). Speed-accuracy trade-off in recognition memory. *Science*, 181, 574-576.
- Sagi, D. (1990). Detection of an orientation singularity in Gabor textures: Effect of signal density and spatial frequency. *Vision Research*, 30, 1377-1388.
- Schouten, J. F. & Bekker, J. A. M. (1967). Reaction time and accuracy. *Acta Psychologica*, 27, 143-153.
- Sperling, G. (1989). Three stages and two systems of visual processing. *Spatial Vision*, 4, 183-207.
- Sperling, G. & Chubb, C. (1989). Apparent motion derived from spatial texture. *Investigative Ophthalmology and Visual Science (Suppl.)*, 30, 161.
- Sutter, A., Beck, J. & Graham, N. (1989). Contrast and spatial variables in texture segregation: Testing a simple spatial-frequency channels model. *Perception & Psychophysics*, 46, 312-332.
- Sutter, A. & Graham, N. (1992). Investigating the dynamics of preattentive texture segregation: How immediate is it? *Investigative Ophthalmology and Visual Science (Suppl.)*, 33, 956.
- Tolhurst, D. J. (1975). Reaction times in the detection of gratings by human observers: A probabilistic mechanism. *Vision Research*, 15, 1143-1149.
- Victor, J. D. (1988). Models for preattentive texture discrimination: Fourier analysis and local feature processing in a unified framework. *Spatial Vision*, 3, 263-280.
- Victor, J. D. & Conte, M. (1987). Local and long-range interactions in pattern processing. *Investigative Ophthalmology and Visual Science (Suppl.)*, 28, 362.
- Victor, J. D. & Conte, M. (1989a). Cortical interactions in texture processing: Scale and dynamics. *Visual Neuroscience*, 2, 297-313.
- Victor, J. D. & Conte, M. (1989b). What kinds of high-order correlation structure are readily visible? *Investigative Ophthalmology and Visual Science (Suppl.)*, 30, 254.
- Victor, J. & Conte, M. (1991). Spatial organization of nonlinear interactions in form perception. *Vision Research*, 31, 1457-1488.
- Wandell, B. (1977). Speed-accuracy tradeoff in visual detection: Applications of neural counting and timing. *Vision Research*, 17, 217-225.
- Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, 41, 67-85.
- Wickelgren, W. A., Corbett A. T. & Doshier, B. A. (1980). Priming and retrieval from short-term memory: A speed-accuracy tradeoff analysis. *Journal of Verbal Learning and Verbal Behavior*, 19, 387-404.
- Wilson, H. R. & Richards, W. A. (1992). Curvature and separation discrimination at texture boundaries. *Journal of the Optical Society of America A*, 9, 1653-1662.
- Wilson, H. R., Ferrera, V. P. & Yo, C. (1992). Psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience*, 9, 79-97.
- Wolfe, J. (1992). "Effortless" texture segmentation and "parallel" visual search are not the same thing. *Vision Research*, 32, 757-763.
- Yo, C. & Wilson, H. R. (1992). Perceived direction of moving two-dimensional patterns depends on duration, contrast, and eccentricity. *Vision Research*, 32, 135-147.

Acknowledgements—This research was supported by National Eye Institute grant IRO1 EYO8459. A preliminary report of this work was presented at the Annual Meeting of the Association for Research in Vision and Ophthalmology, Sarasota, FL in May 1994. The authors wish to express their gratitude to Wayne Wickelgren for many helpful discussions and for encouraging us to try the SAT paradigm.

APPENDIX

Details of the Curve-fitting Procedure

The function F [equation (1)] was fit separately to the 36 different data sets (D) corresponding to the nine matched-contrast sets depicted in Figs 9-11, each containing four conditions. To find the best fitting functions of the form of equation (1), a grid search was conducted over values of the parameter R_{max} (which was not allowed to exceed 0.9999), and the Nelder-Meade algorithm (as instantiated in MATLAB: see, e.g. Press, Flannery, Teukolsky & Vetterling, 1986) was used to find the two parameters t_0 and α from equation (1) in the text. To take into account the heterogeneity of variance in proportions, the error term that we minimized was:

$$\text{error} = \sum_{t=1}^8 \left(\frac{[D(t) - F(t)]}{\sigma_t} \right)^2 \quad (\text{A1})$$

where σ_t is the standard deviation of the data proportion $D(t)$ on the assumption that the true proportion is $F(t)$, namely:

$$\sigma_t = \sqrt{\frac{F(t) \cdot [1 - F(t)]}{n}}$$

where the eight values of t are the eight values of the average response times produced by the eight different cue lag times (50, 100, 150, 200, 250, 300, 500 and 800 msec after the stimulus onset), $D(t)$ is the proportion of correct responses given by the subject at cue delay t , $F(t)$ is the value at time t of the fitted function, and n is the number of trials per data point $D(t)$. The value of n was 100 for all the results reported here.

Generally, the goodness of fit can be assessed by considering the errors [equation (A1)] obtained for our fits. They confirm that the fits were good. If the number of trials n was large enough and if the data were indeed describable by functions of the form in equation (1) then each data proportion $D(t)$ would be normally distributed with a mean of $F(t)$ and a standard deviation of σ_t . If so, each term in equation (A1) would be the square of a standard normal deviate. Since the sum of independent squared standard normal deviates is χ^2 distributed, the error expression in equation (A1) would be χ^2 distributed. If not only was n large enough, but the parameters had been estimated by maximum likelihood methods, the number of degrees freedom (d.f.) would be 5 (d.f. = 8 - 3). This χ^2 distribution has a mean of 5, a median of 4.35, a standard deviation of 4.35, a 95% cut off of 11.07 and a 99% cut off of 15.09. The distribution of sum-squared errors [equation (A1)] for the 36 fits performed for the experimental results reported here is presented in Fig. A1. This distribution was very close to the χ^2 distribution with d.f. = 5 with perhaps a hint of being

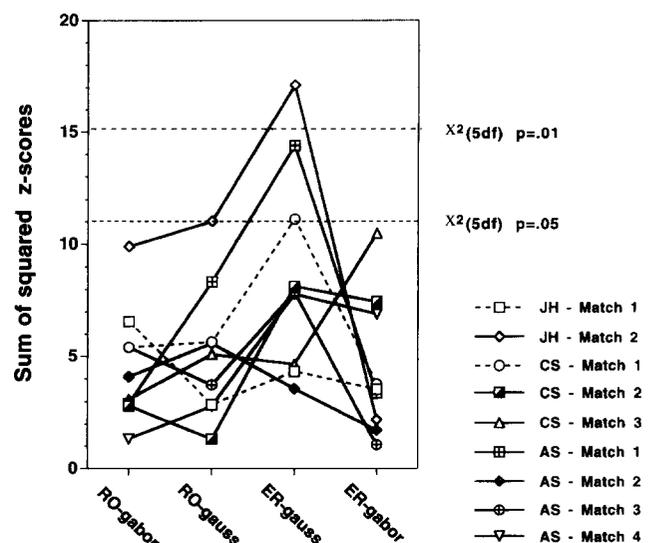


FIGURE A1. The sum-squared errors of the 36 fitted functions $F(t)$ as a function of stimulus class. χ^2 cutoffs for $\alpha = 0.01$ and $\alpha = 0.05$ are marked.

somewhat higher; in particular, for the distribution of sum-squared-errors from the experimental results the mean was 5.8, the median 4.9 and the standard deviation 3.7; also four of the 36 values were above the 95% cut off value (although two were barely above) and one was above the 99% cut off. Thus considered as a whole, the fits here are as expected from the hypothesis that the true population proportions are given by functions of the form in equation (1).

Monte Carlo simulations

However, given the extreme probabilities generated by $F(t)$ for most of our curves, the expected number of incorrect responses is much less than one at many cue lags, and the normal approximation may well be inadequate. To investigate this possibility, we ran some Monte-Carlo simulations using true probabilities $F(t)$ like those describing our data and generating the data proportions $D(t)$ randomly. (We ran seven different simulations, each composed of 1000 simulated replications of an experiment involving eight cue lags with 100 trials at each lag. We did not go so far as to fit each of the 7000 simulated sets of data, however, as the time required did not seem worth the gain.) We computed the error expression in equation (A1) from the true proportions. If n were large enough, this computed sum-squared error, which consists of eight independent terms and no parameter estimates, would be expected to be χ^2 distributed with 8 d.f. This χ^2 distribution has a mean of 8 and a standard deviation of 4. While the distributions of sum-squared errors from the simulations did turn out to have means near 8, their standard deviations were larger than 4, sometimes

substantially (by a factor of 3 or 4). The standard deviations were larger for the following reason: in the simulations, there were rare occasions on which the simulated observer made a mistake at the long cue lags even though the true probability of being correct at those lags was very high (e.g. 0.9999). When such a mistake occurs, the sum-squared errors computed from equation (A1) is very large indeed; these rare occasions lead to a somewhat bimodal distribution of sum-squared errors (most contained down near zero in a distribution that looks like a χ^2 but a few very large error terms that tend to be concentrated together) with an elevated standard deviation.

Of course, if instead of using the true proportions in equation (A1), one used fitted functions to the results and used those proportions (as we did for the data from the experiment but not for the simulations), the effect of these rare mistakes would have been attenuated (because the asymptote estimated would have been lower than the true proportions). But the direction of the effect would remain: the actual distribution of sum-squared errors would be more spread out than that given by the asymptotic theory (more spread out than the χ^2 with 5 d.f. in short). In particular, the appropriate 0.05 and 0.01 cut off would be somewhat higher. This could only make the obtained fits look even better than they already appear.

In summary, the results of the Monte Carlo simulations confirm that the obtained fits are good—in other words, the distribution of the obtained error terms is what one would expect on the hypothesis that the true functions are of the delayed exponential form F in equation (1).