Is the straddle effect in contrast perception limited to second-order spatial vision?

Norma V. Graham

Department of Psychology, Columbia University, New York, NY, USA ✉

S. Sabina Wolfson

Department of Psychology, Columbia University, New York, NY, USA ✉

Previous work on the straddle effect in contrast perception (Foley, 2011; Graham & Wolfson, 2007; Wolfson & Graham, 2007, 2009) has used visual patterns and observer tasks of the type known as spatially second-order. After adaptation of about 1 s to a grid of Gabor patches all at one contrast, a second-order test pattern composed of two different test contrasts can be easy or difficult to perceive correctly. When the two test contrasts are both a bit less (or both a bit greater) than the adapt contrast, observers perform very well. However, when the two test contrasts straddle the adapt contrast (i.e., one of the test contrasts is greater than the adapt contrast and the other is less), performance drops dramatically. To explain this drop in performance—the straddle effect—we have suggested a contrast-comparison process. We began to wonder: Are second-order patterns necessary for the straddle effect? Here we show that the answer is "no". We demonstrate the straddle effect using spatially first-order visual patterns and several different observer tasks. We also see the effect of contrast normalization using first-order visual patterns here, analogous to our prior findings with second-order visual patterns. We did find one difference between first- and second-order tasks: Performance in the first-order tasks was slightly lower. This slightly lower performance may be due to slightly greater memory load. For many visual scenes, the important quantity in human contrast processing may not be monotonic with physical contrast but may be something more like the unsigned difference between current contrast and recent average contrast.

only as the result of eye movements. Thus it is important to know how the contrast at one moment affects the perception of contrast in the next.

Some years ago we came across an effect of previous contrast on current perception of contrast in second-order spatial patterns (Graham & Wolfson, 2007; Wolfson & Graham, 2007) that we now call the *straddle effect*. We have continued to study this effect (Graham, 2011; Graham & Wolfson, 2013; Wolfson & Graham, 2009). Foley (2011) replicated the straddle effect and studied it further. Kachinsky, Smith, and Pokorny (2003) have presented results we think may be closely related to the straddle effect (see also Pokorny, 2011). The straddle effect's relationship to previously reported effects in adaptation and masking literature is discussed by Wolfson and Graham (2009) and Foley (2011).

All the published studies of the straddle effect use visual stimuli in which the two contrasts to be discriminated are presented in different spatial locations of a single pattern; in other words, they are presented to the visual system simultaneously. Thus the observer's performance in all the published straddle-effect studies could reflect the action of so-called second-order spatial-vision pathways, and we originally thought it did. (The term "second-order" as is generally used in the spatial-vision literature is discussed many places; e.g., Graham, 2011; Landy, 2013.) The main point of the studies reported here is to ask whether the straddle effect in contrast discrimination is limited to second-order spatial-vision pathways or not.

Introduction

Humans spend most of their waking hours looking at regions filled with texture or pattern (regions of greater than 0% contrast), and very little time looking at blank, unpatterned areas (0% contrast). Further, this visual contrast is changing rather rapidly in time, if

Replication of the second-order orientation-identification experiment

In order to provide the necessary background to go on to the main point of this article, this section will

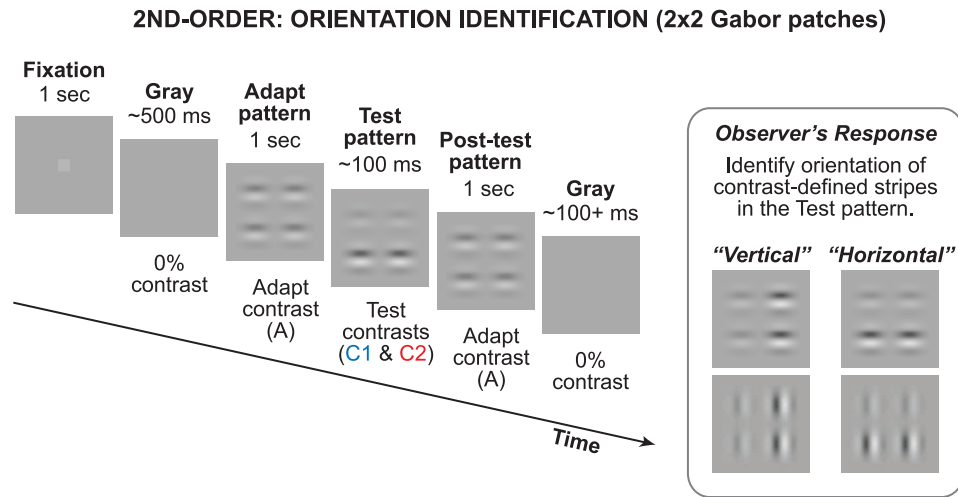**2ND-ORDER: ORIENTATION IDENTIFICATION (2x2 Gabor patches)**

Figure 1. The stages of a typical trial in the second-order orientation-identification experiment. The trial starts with a fixation screen, followed by a plain gray screen, followed by an adapt pattern and then a test pattern; this is then followed by a posttest pattern and a final gray screen. The adapt pattern is composed of four Gabor patches all at the same contrast (adapt contrast A). The test pattern is composed of four Gabor patches with two different contrasts: test contrasts C1 and C2. The posttest pattern is identical to the adapt pattern. In each of these patterns, the four Gabor patches are arranged in a 2 × 2 grid. The final gray screen remains on until the observer responds. Auditory feedback as to correctness of the response is given immediately after the response. The observer starts a new trial with another key press. Between trials, the screen remains gray. Approximate durations are shown in the figure; the exact values for every experiment are shown in Table A1. (The observer is not allowed to respond until at least 100 ms have passed, hence the marked duration of 100+ ms on the diagram for the final gray screen.) Contrast differences in the gray-level images are exaggerated to increase their salience. The spatial layout and dimensions of the pattern are shown in the left column of Figure 13. The box on the right-hand side of the figure illustrates the response the observer is instructed to make. The bottom two test patterns are composed of vertical Gabor patches, and the top two are composed of horizontal Gabor patches. The left-hand two test patterns have contrast-defined stripes that run vertically, and the right-hand two have contrast-defined stripes that run horizontally. The observer responds by indicating the orientation of the contrast-defined stripes. The labels "Vertical" and "Horizontal" above the columns give the correct responses. See Appendix A for more details of the methods and procedures.

briefly describe a replication of the second-order orientation-identification experiment we have previously published. The experimental procedure producing the straddle effect and then the effect itself will be illustrated using data from two new observers. (These two observers also took part in almost all the experiments presented later in this article.) The processes we hypothesized to explain these results are then briefly reviewed at the end of this section. All the explanatory material in this section is presented in greater depth in our previous publications (see, in particular, Wolfson & Graham, 2009).

Figure 1 shows the spatial characteristics of the stimulus at each stage in a typical trial of the second-order orientation-identification experiment. The *test pattern* contains four Gabor patches arranged in a 2 × 2 grid; two are at one contrast (C1) and two are at another (C2). Test contrasts C1 and C2 vary from trial to trial. The test pattern is presented between the *adapt pattern* and the *posttest pattern*. The adapt and posttest patterns are entirely identical. The Gabor patches in the adapt and posttest patterns are always at 50% contrast. All the Gabor patches throughout a single trial (all the patches in the adapt, test, and posttest patterns) have

the same spatial frequency and orientation and occupy the same spatial positions.

The observer is required to say whether the two different contrasts in the test pattern define horizontal stripes or vertical stripes (see inset at right of Figure 1 for examples of patterns and the correct response associated with each). The orientation of the contrast-defined stripes is technically a second-order orientation; thus this task is a second-order orientation-identification task. One might also describe this task as identifying the global orientation of contrast differences (i.e., the orientation of the contrast-defined stripes) while ignoring as irrelevant the local orientation (i.e., the orientation of the bars in individual Gabor patches).

Replacing the posttest pattern with a gray field does not affect the results materially as far as the questions and conclusions of the present article go. (Some results with a gray field as the posttest pattern are shown in the supplementary material to Wolfson & Graham, 2009.) Keeping the adapt and posttest patterns identical to each other has an advantage in that, in many situations, it allows the hypothesized processes to be disentangled more easily in interpreting results.

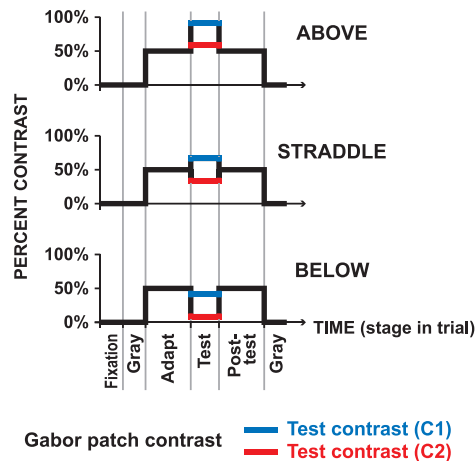**2ND-ORDER: ORIENTATION IDENTIFICATION (2x2 Gabor patches)**

Figure 2. The contrasts throughout three different kinds of trials are shown here as a function of time (stage in the trial). During the test pattern, the blue line gives test contrast C1 and the red line gives test contrast C2. The thick black lines show the single contrast when all four patches have the same contrast (50% during the adapt, 50% during the posttest, and 0% during the gray stages). The three trials (from top to bottom) illustrate an Above, a Straddle, and a Below test pattern.

Contrasts are plotted for three kinds of trials in Figure 2. During the test pattern (labeled "Test" on the horizontal axis), two Gabor patches have one contrast (blue line, labeled C1) and the other two Gabor patches have another contrast (red line, labeled C2). At other times, all four Gabor patches have the same contrast (thick black line).

In the figure's top row, both test contrasts are *above* the adapt contrast. We will refer to this case as an Above test pattern. In the middle row, the two test contrasts *straddle* the adapt contrast; that is, one test contrast is above and the other is below the adapt contrast. We refer to this case as a Straddle test pattern. In the bottom row, both the test contrasts are *below* the adapt contrast, and we refer to this case as a Below test pattern.

Details of the spatial and temporal characteristics of the stimuli and procedure used in the second-order orientation-identification experiment (and all the other experiments presented in this article) are described in Appendix A, and the values of the parameters are listed in Table A1.

Aside on terminology

Our experimental procedure might be called *short-term adaptation* to visual contrast. Others may prefer other terms for this procedure—e.g., a masking procedure (Foley, 2011). See further discussion in Wolfson and Graham (2009).

Second-order orientation-identification results: Straddle effect and Weber-law behavior

The five rows in Figure 3 show performance for five test patterns. The adapt pattern is the same for all rows and has Gabor patches of 50% contrast. The contrasts in the test pattern vary from very high (top row) to very low (bottom row), but the difference between the two test contrasts is held constant, at $|C1 - C2| = 10\%$ in this case. We refer to such a set of patterns as a *constant-difference series.*

The third row in Figure 3 shows the test pattern which has contrasts of 45% and 55%, contrasts that straddle the adapt contrast of 50%. Performance on this Straddle test pattern is poor (71% and 59% correct, where chance is 50%).

Performance on the test patterns in the second and fourth rows—where both test contrasts are on the same side of, and quite near, the adapt contrast—is very close to perfect (98% to 100%). We refer to the low performance on the Straddle test pattern compared to its near neighbors in a constant-difference series as the straddle effect.

Performance on the test patterns in the first and fifth rows—where both test contrasts are on the same side of the adapt contrast but far from it; that is, Far-Below and Far-Above test patterns—is poor (59% to 65%).

Figure 4 shows results from the two observers shown in Figure 3, but now for many more test-contrast combinations. Look first at the left-hand panels. Percentage correct is on the vertical axis and average test contrast (i.e., the average of the two test contrasts) is on the horizontal axis. The value of the adapt contrast is 50%, indicated by the red asterisk at the middle of the horizontal axis. Each curve shows performance for a constant-difference series (e.g., the difference $|C1 - C2|$ between the two test contrasts is always 5% for the lowest curves).

In Figure 4, the right-hand column shows performance as d' values rather than percentage correct. We show these d' values here (which we have not done in our previous publications on the straddle effect) in order to better compare the results in the second-order orientation-identification experiment with results from, for example, the first-order experiments presented later in this article.

As is generally true in computing d' values, there is a problem if performance is 100% or 0% (i.e., the proportion is either 1 or 0): In these cases, the d' value computed straightforwardly from the experimental results will be infinite. To correct for this problem we
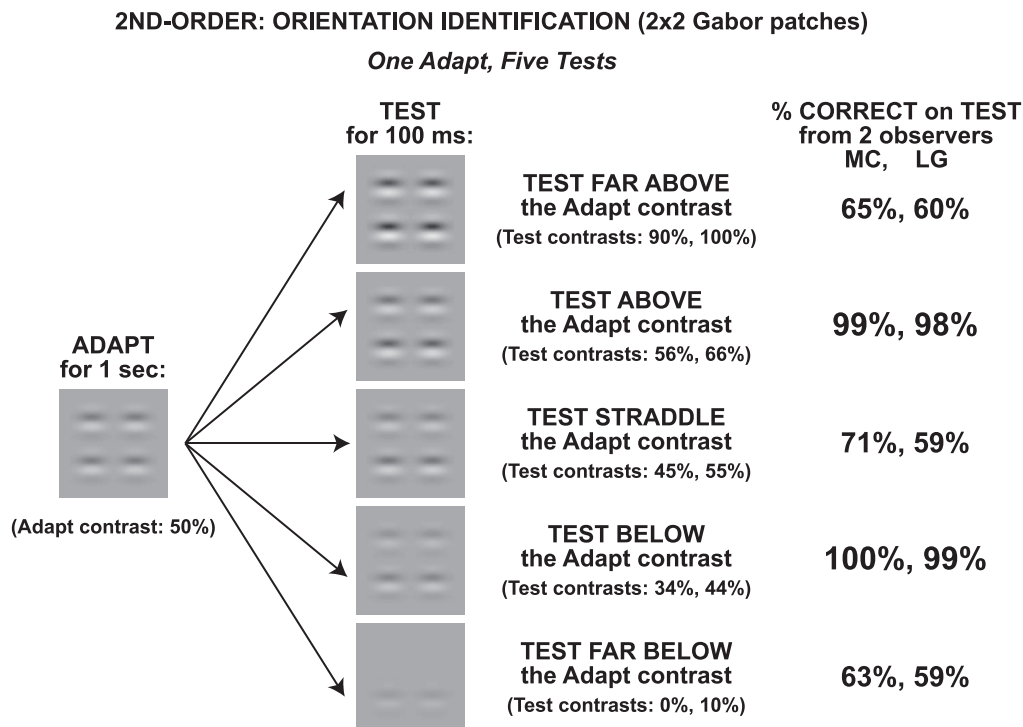
Figure 3. Performance of observers MC and LG in the second-order orientation-identification experiment for five test patterns that form a constant-difference series. For further description, see the text. The test patterns in the gray-level images here will not appear to the reader the way the actual test patterns did to the observers in the experiment because the reader will not have looked at the 50% adapt pattern immediately before the test pattern.

used a rule to truncate the computed infinite values to finite values that can be plotted. The truncation rule we used allows sample size (the number of trials entering into any plotted point) to affect the truncated value. Truncated data points in this and subsequent figures are shown as open symbols. See Appendix B for details of the d' computations.

In summary, each plotted curve shows that the results for a constant-difference series of patterns is at a local minimum at (or near) the perfect Straddle pattern (where the average test contrast equals the adapt contrast). Each curve then reaches a local maximum on either side (at points where the average test contrast is either above or below the adapt contrast). Each curve then drops again as the average test contrast moves further away from the adapt contrast in either direction, forming two tails on each curve.

These tails conform quite well to Weber's law, as we have shown previously (Graham & Sutter, 2000; Wolfson & Graham, 2009). The dimension on which this version of Weber's law holds is the absolute value of the difference between the adapt contrast and the test contrast.

The results for observers LG and MC in Figures 3 and 4 replicate the results of the observers for whom we

have previously published results (see, e.g., the results in figure 13 of Wolfson & Graham, 2009).

Proposed explanation of the second-order orientation-identification results: Contrast comparison and contrast normalization

The curve in Figure 5 is an idealized representation of the observer performances shown in Figure 4. The kind of curve in Figure 5 has been called a *butterfly curve* (Hochberg, 1978, p. 240, although in a very different perceptual-adaptation context, involving bathwater temperature). The terms *contrast normalization* and *contrast comparison* that label regions of the curve are given to the processes discussed in this subsection.

The straddle effect can be explained by assuming that the relevant part of the visual system cannot always discriminate an increase in contrast from a decrease of the same magnitude. It is as if the contrast of a new stimulus is always compared to the average recent contrast at that same location. Then the magnitude of the change is registered and sent upstream, but the sign of the contrast change is lost or at least imperfectly retained.

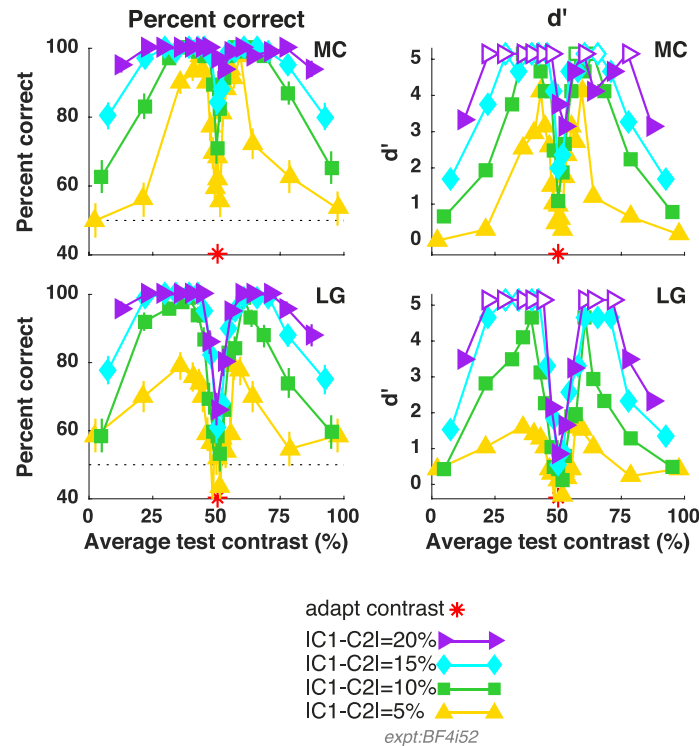## 2ND-ORDER: ORIENTATION IDENTIFICATION (2x2 Gabor patches)

Figure 4. Performance in the second-order orientation-identification experiment (using a 2 × 2 Gabor-patch grid) plotted as percentage correct (left panels) or d' (right panels). Results are shown for observers MC (top row) and LG (bottom row). The horizontal axis shows the average test contrast. The adapt contrast was 50% and is marked by a red asterisk on the horizontal axis. The difference $|C1 - C2|$ between test contrasts distinguishes the curves. The patterns on a single curve form a constant-difference series. Results for the smallest test-contrast difference we used (5%) are in the lowest curve (yellow upright triangle symbols); results for the highest test-contrast difference (20%) are in the upper curve (purple right-pointing triangle symbols). There were about 100 trials per point. The open symbols in the right panels are cases where the computed d' values were truncated to correct for the problems in computing d' values associated with 0% and 100% performance. (See Appendix B for further description of this truncation method.) The error bars in the left panels show ± 1 standard error across blocks. The experiment from which these results came also included patterns of test-contrast differences larger than 20%, but for the sake of visual clarity those results are not shown here.

More specifically, we have proposed a *contrast-comparison process* (Graham & Wolfson, 2007; Wolfson & Graham, 2007). This is a rapid adaptation process in which a comparison level is continually updated at each spatial position:

1. The comparison level equals the recent (much less than 1 s) weighted average of contrast at that spatial position.
2. The comparison level is subtracted from the current input contrast.
3. The magnitude of the difference is sent upstream, but information about the sign of that difference is lost or at least degraded—that is, a full-wave or partial rectification occurs.

To have some intuition into why this hypothesized contrast-comparison process can explain the straddle effect, consider a simple numerical example. Here one stimulus (called the earlier stimulus) is immediately replaced by a second stimulus (called the later stimulus) that is identical except for contrast. For simplicity in this example, assume that the earlier and later stimuli are single Gabor patches identical in all characteristics except contrast. Assume also that the earlier stimulus has been present long enough that the comparison level equals the contrast of that earlier stimulus. Now consider two possible contrasts for the later stimulus:

- The later stimulus has a contrast 3 arbitrary units above that of the earlier stimulus. Then (see Figure 6) the contrast-comparison process's output to this later stimulus will be +3 arbitrary output units.
- The later stimulus has a contrast 3 arbitrary units below that of the earlier stimulus. Then (see Figure 6), the contrast-comparison process's output to this later stimulus will also be +3 arbitrary output units.

Thus, any process subsequent to this contrast-comparison process could not tell the difference between the
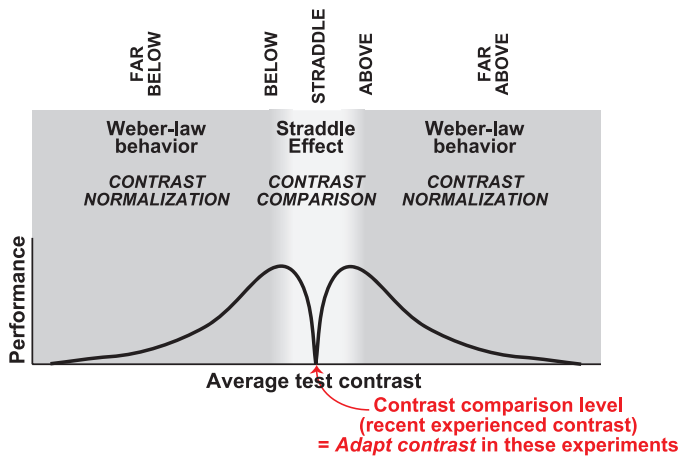
Figure 5. Idealized version of results in second-order orientation-identification experiments (e.g., Figure 4). We have suggested that the drop in performance in the center—the straddle effect—is the result of a shifting, rectifying contrast-comparison process (a process previously nicknamed "Buffy"). The contrast-comparison level shifts to equal the recently experienced contrast (which in these experiments is the adapt contrast). We have also suggested that the drops in performance at the two ends (which obey Weber's law) are the result of a contrast-normalization process. (Later in this article we show similar results for first-order experiments.)
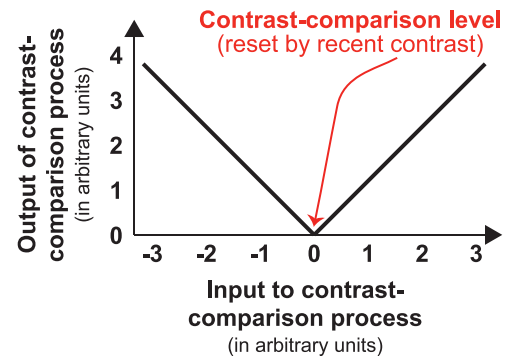

Figure 6. Diagram of the action of a shifting, rectifying contrast-comparison process. The input to this process—on the horizontal axis—reflects the local contrast at each position in the visual field, perhaps something like the size of a single Gabor patch. This input is compared to a comparison level indicated by the red arrow. The comparison level adapts over time to equal the recently experienced contrast. The output from this process is an unsigned measure of the difference between the current contrast and the recent average contrast. A more complete picture of the input/output functions from the contrast-comparison process is published in Graham (2011, figure 17).

Above and Below later stimuli, since both have an output of +3 arbitrary units. (This hypothesized contrast-comparison process was originally nicknamed "Buffy" for reasons described in Graham & Wolfson, 2007. Explanations that will not work to explain the straddle effect are considered in Wolfson & Graham, 2009.)

The Weber-law behavior, on the other hand, cannot be explained by the contrast-comparison process, but it is consistent with a gain-control process of the normalization type that has been suggested in many context both behavioral and physiological (see the earlier references in Graham, 2011; more recently, see also Carandini & Heeger, 2011; Solomon & Kohn, 2014).

In our situation it is the rectified signal from the contrast-comparison process that becomes the input to the proposed contrast-normalization process. The labels in Figure 5 identify the region where the hypothesized contrast-comparison process dominates and the region where the hypothesized contrast-normalization process dominates, although both processes are assumed to operate over the whole range. Such a combination model applies to our results qualitatively (Wolfson & Graham, 2009) and quantitatively (Graham, 2011, figures 16–18).

Foley (2011) has suggested a different model that he showed worked well in predicting the straddle effect in his experimental results. His model incorporated two processes. One, called the V-response, explained the straddle effect and had effects much like those of the

contrast-comparison process of our model. The second process, called the S-response, produced a monotonic S-shaped response like that in many previous models. While both Foley's model and ours postulate two processes and predict the occurrence of the straddle effect, there are differences between the two models' predictions. We are not going to attempt to test between these two models (or any others) in the current article.

Foley also has a very nice discussion of the phenomenology. He asked two of his observers to report systematically and in considerable detail the perceived appearances of the patterns. We have never done as systematic or thorough a job of collecting observers' perceptions; however, our observers' occasional informal reports—as well as our own experiences—agree with Foley's descriptions. His work makes it clear once again that the qualities of perceived appearance do not necessarily correlate in any simple way with discrimination performances.

Aim of this article

The straddle effect in contrast discrimination has been shown only for second-order spatial vision. But nothing of what we postulated in our model (or Foley in his) requires that it be limited to second-order pathways: The straddle effect is happening at every spatial position in these explanations. It could thus act the same way within the pathways that process first-order spatial patterns. In other words, the rectification-like behavior does not

necessarily require that the two contrasts to be discriminated occur in the visual field simultaneously.

The current article asks whether the straddle effect in contrast discrimination and the accompanying Weber-law behavior occur in first-order spatial vision (when only one test contrast is present at a time)—and argues that the answer is "yes".

First-order (same–different, two-interval) spatial-vision experiment

To ask whether the straddle effect can occur in first-order spatial vision, we need a task—using only first-order spatial patterns—that could potentially show a straddle effect. To be able to see such a straddle effect, one needs to measure the discrimination between two test contrasts after adaptation to a third contrast (the adapt contrast). To be able to see a straddle effect in first-order vision, the two test contrasts to be discriminated should not be presented simultaneously. If they are presented simultaneously, the performance on the task is too likely to be influenced by second-order spatial vision.

Generalizing from first- and second-order terminology to comparison across time or space

Instead of framing the questions of this article in terms of second- versus first-order vision, one can generalize the terminology. Note that in our second-order experiments, the two test contrasts to be discriminated occurred simultaneously at two different positions in space; in our first-order experiments, the two tests contrasts to be discriminated occur at two different times. Thus this article is asking some of the many possible questions about whether comparison of contrasts across time behaves the same as comparison of contrasts across space.

The same–different two-interval task using first-order spatial patterns

The task we chose to study first-order vision is a same–different two-interval task, so the two test contrasts can appear at different points in time by appearing in two different intervals. The first-order experiment is illustrated in the top half of Figure 7; in the bottom half of Figure 7 we have repeated Figure 1, showing the second-order experiment (in order to illustrate the relationship between the two experiments). In the top half of the figure, all four Gabor patches at any moment in time have the same contrast.

On half the trials, the test pattern in Interval 1 has a different Gabor-patch contrast than the test pattern in Interval 2 has. On the other half of the trials, the test patterns in both intervals have the same contrast. The observer's task is to indicate whether the two intervals' test contrasts are the same or different.

To clarify the relationship between the second-order experiment (bottom half of Figure 7) and the first-order experiment (upper half of Figure 7), the blue and red arrows indicate that the two test contrasts in the two intervals of the first-order experiment are identical to the two test contrasts at different spatial positions in the second-order experiment.

Since the two test contrasts in Figure 7 (upper half) appear in different intervals, the observer's response cannot be mediated by second-order spatial-vision pathways. In order to make this new first-order experiment as close as possible to the previous second-order experiment, the spatial positions, spatial frequencies, and possible orientations of the Gabor patches are identical in the two experiments. (See Appendix A for more details about the methods and procedures.)

Figure 8 shows contrast as it varies across time (stages) within a trial. The same conventions are used here as in Figure 2, except here the red and blue stand for the test contrasts in two different temporal intervals rather than in two different spatial locations. Figure 8 shows examples of six types of trials—the six that will need to be distinguished in explaining our results. The three in the left column contain different test contrasts in the two intervals, and so the correct response is for the observer to indicate "different." The three in the right column contain the same test contrast in the two intervals, and so the correct response is for the observer to indicate "same." In each row of the figure, the trials in the left and right columns have identical average test contrasts.

In the top (bottom) row, both test contrasts are above (below) the adapt contrast, and therefore the average test contrast is also above (below) the adapt contrast. The middle row shows a pure Straddle test pattern, in which the average test contrast is identical to the adapt contrast; therefore, for this Same trial (middle row, right column), both test contrasts are equal to the adapt contrast. Note also that for this Same trial, contrast stays constant through the adapt, test, and posttest stages of each interval. Thus, the only contrast transitions during the whole Same trial are the four that occur at the beginning of the adapt and the end of the posttest in each of the two intervals. In the other five types of trial in Figure 8, there are eight contrast transitions. We discuss possible implications of this fact for observer behavior in Appendix B, where our d' computations are discussed.
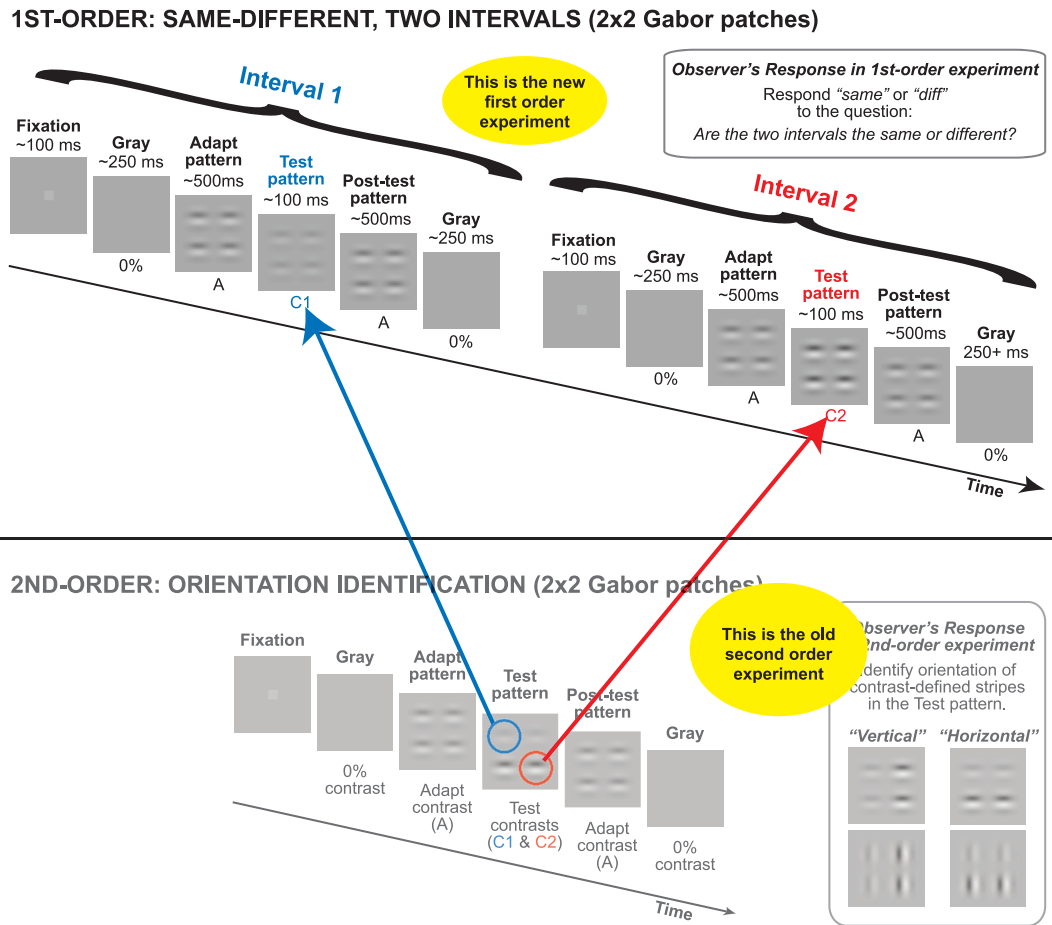
Figure 7. The top half of this figure shows a typical trial in a first-order same–different two-interval experiment. In this experiment the screen always has four Gabor patches of identical contrast present at any moment in time, but the contrast is different at different stages in the trial. Sometimes the screen is gray—that is, the contrast is 0%. The adapt contrast A is identical in both intervals, and in the example shown here, the test contrasts C1 and C2 are different. This first-order experiment in the top of the figure was derived from the second-order experiment shown in the bottom of the figure. (The second-order experiment shown in the bottom of the figure is the experiment shown in Figure 1.) The red and blue arrows may help the reader see the relationship between the two experiments. Feedback as to correctness of the response was provided. See Appendix A for more details about the methods and procedures.

A note about the duration of the adapt and posttest patterns

The duration of the adapt and posttest patterns in the second-order orientation-identification experiment was 1 s (e.g., Fig. 1). This duration was chosen to mimic the dominant duration used in our published results. However, practical considerations persuaded us to shorten the adapt and posttest durations to 500 ms in the two-interval first-order experiment (Figure 7). We did this to keep individual trials from being so long that observers complained of tedium. We have shown previously that any differences between second-order results using 500-ms versus 1,000-ms adapt and posttest durations are slight and not relevant to the conclusions of this article; one observer is published in Graham and Wolfson (2013), figure 1.8, whereas other observers are not published. (Also, there is a new kind of second-order experiment reported later in this article that uses 500-ms adapt and posttest durations.)

A note about a possible alternate response

Rather than asking observers to respond to the question "Are the two intervals the same or different?," we might have asked them to respond to the question "Which of the two intervals has the higher test contrast?" But some pilot work showed us that observers using the latter kind of response tended to perform systematically below chance on trials in which both test contrasts were below the adapt contrast. We realized that an observer's wrong answers on such trials are what is predicted by our proposed contrast-comparison process for the following reason: The output of the proposed contrast-comparison process is only the magnitude of the change from adapt to test
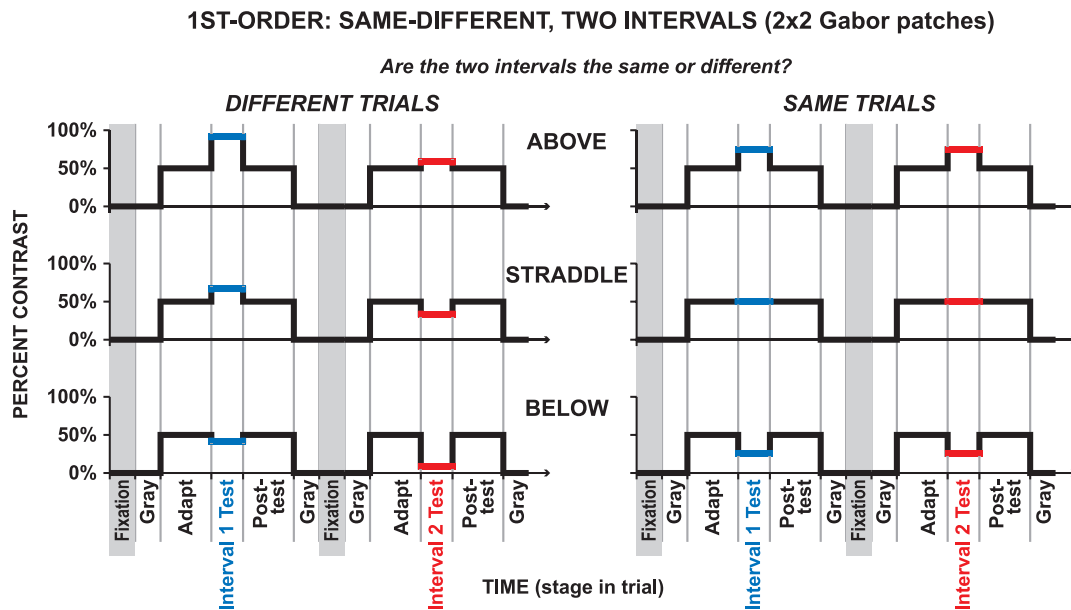
Figure 8. The contrast-versus-time profiles for six types of trial in the first-order same–different two-interval experiment. The blue lines indicate test contrast C1 and the red lines indicate test contrast C2. Thick black lines indicate contrasts of 0% (when the screen is homogeneous gray) and 50% (during the adapt and posttest patterns). The correct response is "different" for the trials in the left column and "same" for the trials in the right column. The top (bottom) row shows profiles for the cases where the test contrasts are both above (both below) the adapt contrast. The middle row shows cases where the average of the two test contrasts equals the adapt contrast. (Thus in the middle row of the right column, the two test contrasts are both equal to the adapt contrast).

(and from test to posttest); it keeps little or no track of whether the change is an increase or decrease. For the trials in which both tests contrasts are below the adapt contrast, the magnitude of the change from adapt to test is *greater* for the test contrast that is *lower*. Thus, if the observers choose the greater magnitude of change to answer the question of "Which of the two intervals has the higher test contrast?," they will choose the lower test contrast and be systematically wrong.

Foley (2011) found a similar problem when asking observers which of two simultaneously presented test patches had the higher contrast. He approached this problem by continuing to use the question but carefully instructing the observers. He describes both the problem and his solution very well, and thus we will not go into further detail here.

Results from the first-order same–different two-interval experiment

Figure 9 shows the results for five different observers (rows) with three different measures of performance (columns). The first two rows show results for the two observers who did the replication of the second-order orientation-identification experiment (MC and LG in Figure 4).

Throughout Figure 9, the large colored points show the results for Diff trials—that is, trials in which there

was a nonzero difference between the two test contrasts and so "diff" was the correct response. Results when the difference between test contrasts was 10% (20%) are shown by the green squares (purple triangles). The little black dots show the results for Same trials (in which the two test contrasts equaled one another, and so "same" was the correct answer). All the horizontal axes in the figure give the average test contrast.

In the left column of the figure, the vertical axes show the percentage of trials on which the observer responded "diff." Not surprisingly, since "diff" was an incorrect answer for the Same trials, the little black dots lie very low in the panels.

The middle and right columns of the figure show performance measured as d' computed in two different ways. We wanted to use d' values to compensate for possible kinds of response bias and thus allow comparison of this first-order (same–different two-interval) experiment to the second-order (orientation-identification) experiment. Any d' computation is the application of a detection-process model in order to calculate from experimental results to a quantity one might call true sensitivity without contamination by quantities one might call response biases. One can never be certain that the detection-process model is perfect. Indeed, one can usually be sure that it is not. Fortunately, experience has shown that the detection-process models do not need to be perfect for the
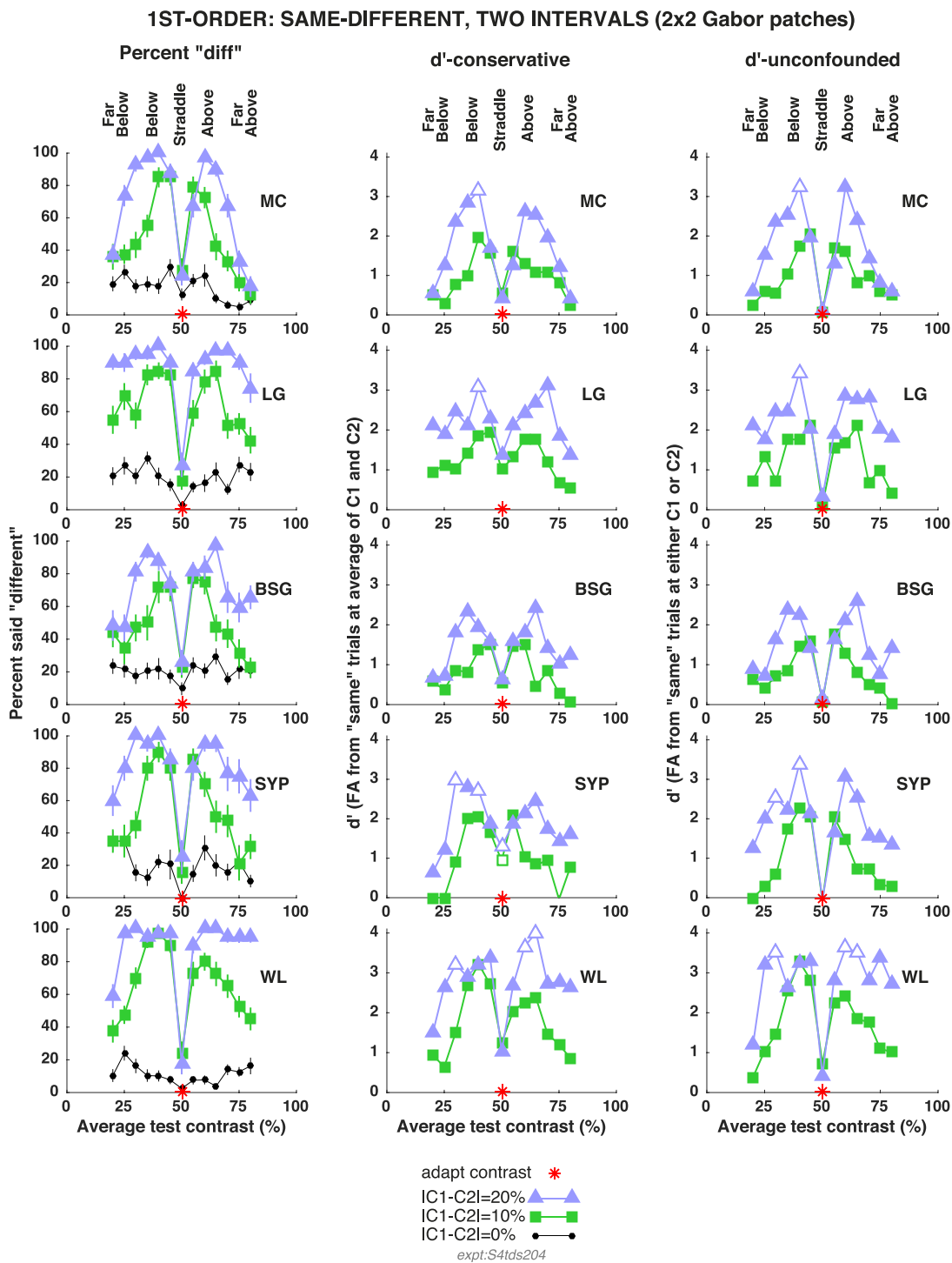
Figure 9. Results from the first-order same–different two-interval experiment (using a 2 × 2 Gabor-patch grid) with five observers (five rows). The three columns show three different measures of the observer's performance: Percentage "diff" (proportion of trials on which the observer said "diff") and the two *d'* measures (*d'*-conservative, *d'*-unconfounded). Each curve gives the results for a particular constant-difference series (10% in green squares, 20% in purple triangles, 0% in small black dots). The error bars in the percentage "diff" plots show ±1 standard error across blocks. The open symbols in the *d'* plots show values that were truncated to deal with the problems produced by 0% or 100% performance. See Appendix B for more details.
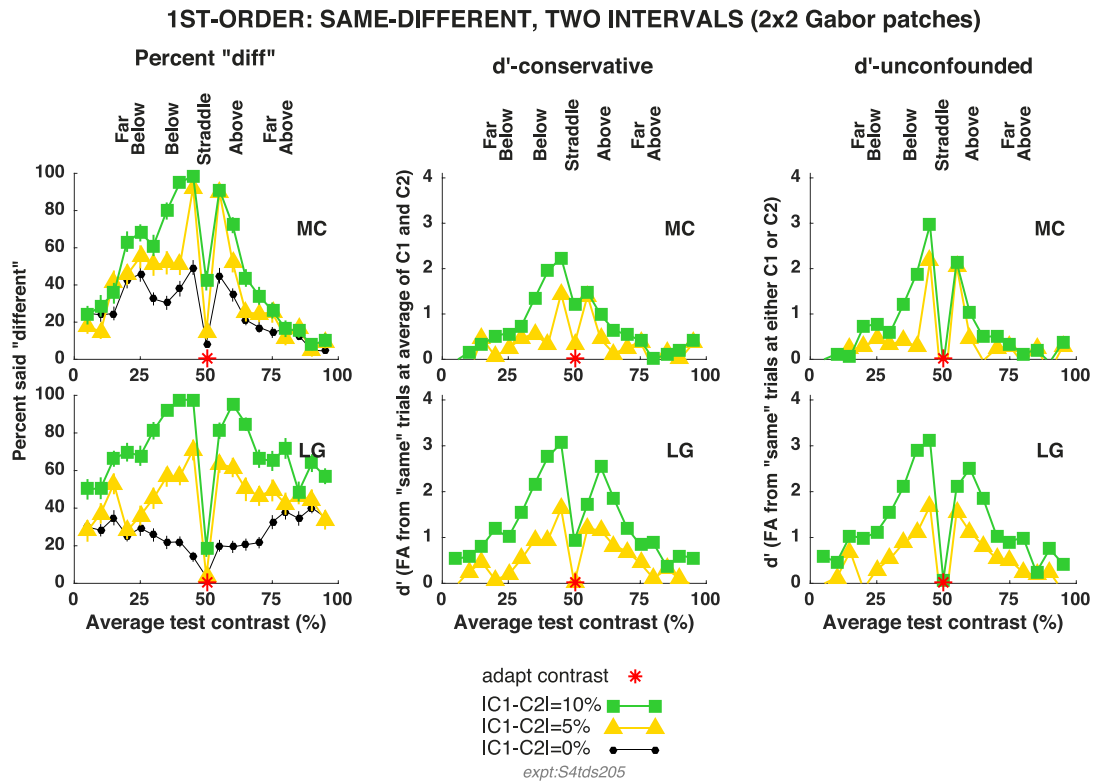
Figure 10. Results from another version of the first-order same–different two-interval experiment (using a 2 × 2 Gabor-patch grid) with observers MC and LG (two rows). This version used a wider range of average test contrasts and a 5% test-contrast difference (yellow triangles), as well as repeating 10% (green squares) and 0% (small black dots). Other aspects are the same as in Figure 9.

calculated *d′* values to be quite informative and prevent errors of interpretation.

We have chosen to display two *d′* measures: *d′-conservative* and *d′-unconfounded*. Appendix B describes these measure in detail; here we will describe them tersely. The middle column of Figure 9 shows the measure we call *d′-conservative* because it is conservative with respect to our major conclusions here. In particular, *d′-conservative* underestimates the depth of the notch corresponding to the straddle effect. The right column shows a measure we think better, which we call *d′-unconfounded* because it corrects for a substantial confound. The difference between these two *d′* measures is in how the false-alarm rate is estimated (and indicated cryptically on the vertical axis labels). For these *d′* values—as for the ones in Figure 4—we had to truncate *d′* values to avoid the problems caused by performances of 0% or 100% correct. Cases where truncation occurred are shown by open symbols in the figure.

There are differences in results among observers in this first-order task, but they are beyond the scope of this article. We have previously noted individual differences in the second-order task (Graham & Wolfson 2007, 2013.)

Figure 10 shows further results for observers MC and LG. There is a wider range of average test contrasts

(farther out toward 0% and 100%) and a test contrast difference of 5% (yellow triangles) in addition to repeating 10% (green squares) and 0% (little black dots).

All the results shown from the first-order experiment (Figures 9 and 10) display the same general shape as second-order results we have published previously and replicated here (Figure 3). The next subsection looks more carefully at the comparison between these first- and second-order results.

Initial comparison of first- and second-order spatial-vision results

Figure 11 juxtaposes the first- and second-order results expressed as *d′* values. We used the traditional *d′* computation for the second-order orientation-identification results (Figure 4), the values shown here as gray stars. The first-order results are the *d′*-unconfounded values shown in Figures 9 and 10 (same symbols as in the earlier figures). The results for test-contrast differences |C1 − C2| equal to 5%, 10%, and 20% are shown in the three columns.
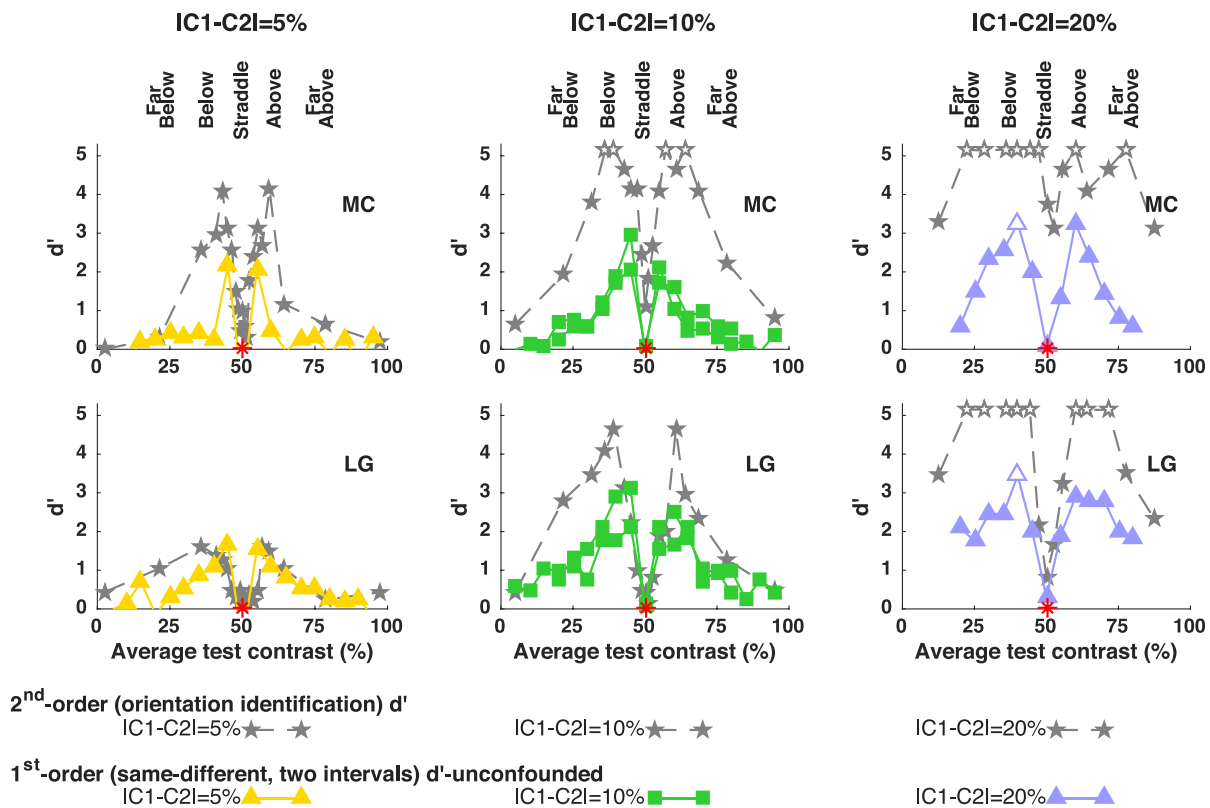
Figure 11. Second-order orientation-identification results (from Figure 4, in gray symbols here) juxtaposed with first-order same–different two-interval results (from Figures 9 and 10, colored symbols) using d' values for all. Each experiment used 2×2 Gabor-patch grids. The results for the two observers (MC and LG) are shown in the two rows. The results for the three different values of test-contrast difference (5%, 10%, and 20%) are shown in the three columns. Note that the test-contrast difference of 10% was run twice (both Figures 9 and 10), which is why there are two curves with green squares in the middle column. The open symbols show cases where the truncation rule had to be applied. There are a substantial numbers of points showing truncation in the second-order experiment (gray symbols), but the only instances of truncation in the first-order experiment (colored symbols) are for $|C1 - C2| = 20\%$. Thus it is not truncation artifacts that explain the lower performance in the first-order same–different two-interval experiment relative to that in the second-order orientation-identification experiment.

Qualitative comparison

As Figure 11 shows, the first-order results are very similar in shape to the second-order results. Generally, the curves from both experiments have a two-peaked shape, with a notch in the center and decreases at both ends, as in Figure 5.

Quantitative comparison

While there is no systematic difference between the shapes of the first- and second-order results, there is a clear substantial difference in the heights of the curves (Figure 11). In general, the d' values from the second-order experiment (gray stars) are substantially higher than those from the first-order experiment (colored symbols).

Note that the different truncation ceilings for these two experiments (further described in Appendix B) do not affect this conclusion, because there is almost no truncation (no open symbols) in the results of the experiment producing lower performance (the first-order experiment).

It would be unwise to conclude from this quantitative difference in observers' performances, however, that the performance of the underlying second-order spatial-vision processes is actually superior to that of the underlying first-order processes. While we tried to make the first- and second-order experiments as similar as possible, they differ in many ways. These differences might affect how well an observer performs the task even if the underlying systems are equally sensitive. We will explore this topic further under the
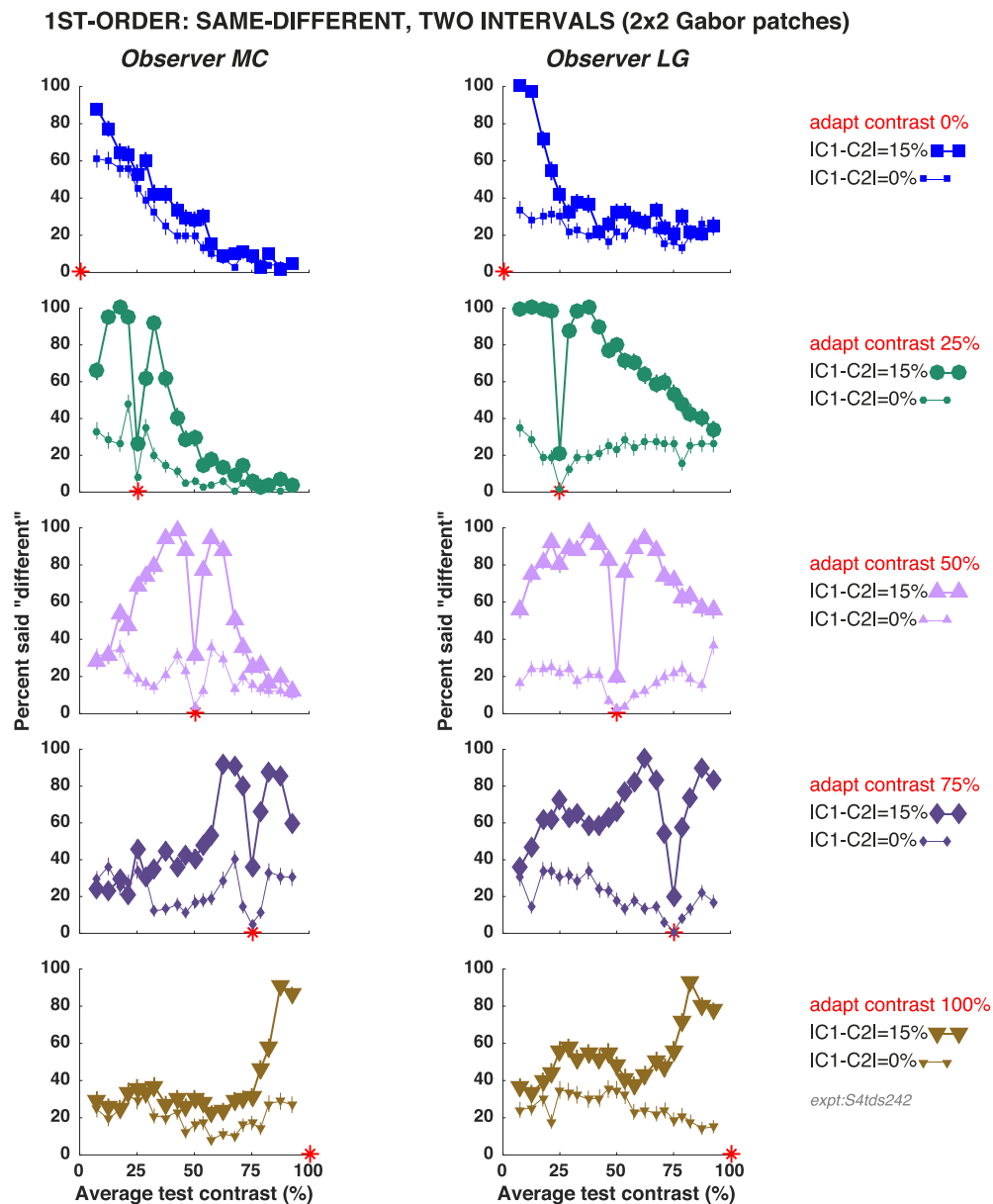
Figure 12. Results of varying adapt contrast in a first-order same–different two-interval experiment (using a 2×2 Gabor-patch grid) for observers MC and LG (two columns). The adapt contrast varies from 0% (top row) to 100% (bottom row). Percentage "diff" is plotted on the vertical axis, and average test contrast on the horizontal axis. Error bars show ± 1 standard error across blocks. These results are plotted on alternate vertical and horizontal axes in Figures C1–C3. Note that for adapt contrasts of 0% and 100% it is impossible to get results where the average test contrast actually equals the adapt contrast.

heading Further comparisons of first- and second-order spatial vision.

In summary

The straddle effect in contrast perception is *not* confined to second-order spatial pathways but also occurs in first-order pathways. This is the major novel result of the present article. The accompanying Weber-law behavior occurs in first-order as well as second-order pathways. This result is less surprising, as Weber-law behavior has been shown in many situations.

For both first and second-order pathways, the notch in the center of the results curves—the straddle effect—can be explained by a shifting, rectifying, contrast-comparison process; the decreases in performance at both tails, reflecting Weber-law behavior, can be explained by a contrast-normalization process.
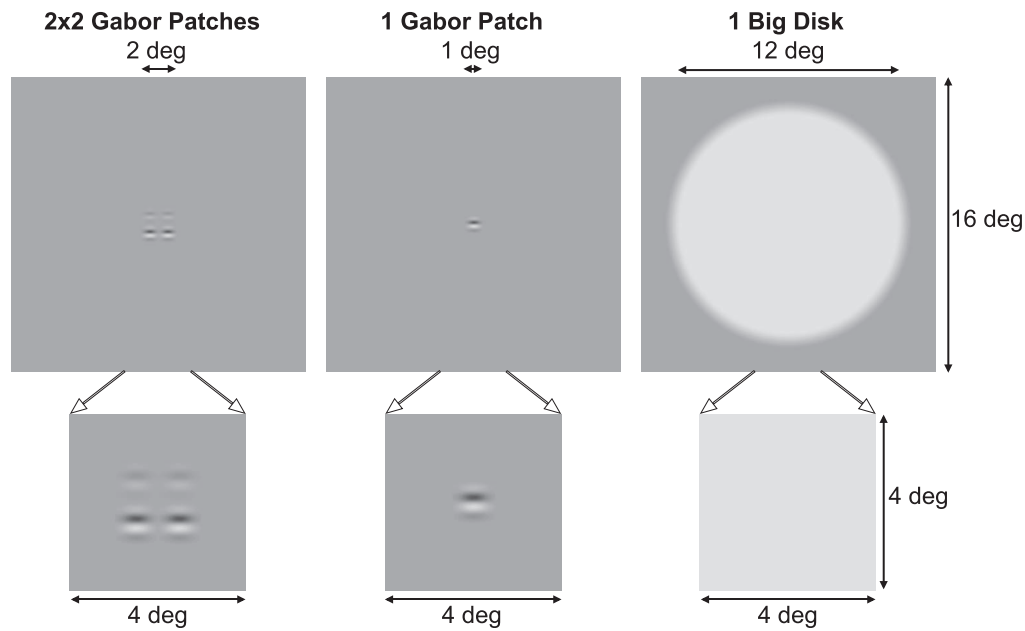
2x2 Gabor Patches
2 deg

1 Gabor Patch
1 deg

1 Big Disk
12 deg

16 deg

4 deg

4 deg   4 deg   4 deg

Figure 13. The three spatial patterns we used in first-order same–different two-interval experiments in this article. The 2 × 2 Gabor-patch grid is in the left panel. The one foveal Gabor patch is in the middle panel, and the big disk is in the right panel. The top row shows the full 16° × 16° area viewed by the observer, and the bottom row shows the central 4° × 4° region expanded to show the details of the patterns better. Further details of these patterns are given in Appendix A.

The input to the contrast-normalization process is assumed to be the output from the contrast-comparison process.

About the remaining sections

The next section presents more first-order experiments, continuing to use the same–different two-interval task (shown in Figures 7 and 8) but with different adapt contrasts and spatial patterns. The section following that presents a new second-order task and further comparisons between the results of first- and second-order experiments. The final section presents conclusions.

Further first-order spatial-vision experiments

This section asks two further questions: Do first-order results show the same changes with adapt contrast that second-order results have previously shown (e.g., Wolfson & Graham, 2009)? And are the first-order results for two other kinds of spatial patterns the same as for the 2 × 2 Gabor-patch patterns used in the previous section?

Varying adapt contrast in a same–different two-interval first-order experiment

An important property of the results from previously published second-order orientation-identification experiments is that the whole curve of performance versus average test contrast shifts with adapt contrast: No matter what the adapt contrast, there is a notch in the curve centered at the point where the average test contrast equals the adapt contrast (e.g., Wolfson & Graham, 2009, figure 12). In other words, the straddle effect shifts so that it is always centered at the adapt contrast.

A priori, this need not have been the result. Perhaps all adapt contrasts (not just the adapt contrast of 50% that we have used so far in this article) would produce notches at 50%, and it was just a coincidence that we chose that adapt-contrast value in our initial experiments. Or perhaps the notch would move continuously with adapt contrast, but not as far as the adapt contrast moves. Or perhaps all low adapt contrasts would produce a notch at 0% (a one-sided curve), all middling ones at 50%, and all high ones at 100%.

This result of varying adapt contrast that we found in previously published second-order experiments must occur if a contrast-comparison process like that we have hypothesized is occurring. It seems wise, therefore, to check that the same result of varying adapt contrast does occur in first-order experiments; so we ran a first-order experiment with five different adapt contrasts using our observers LG and MC.
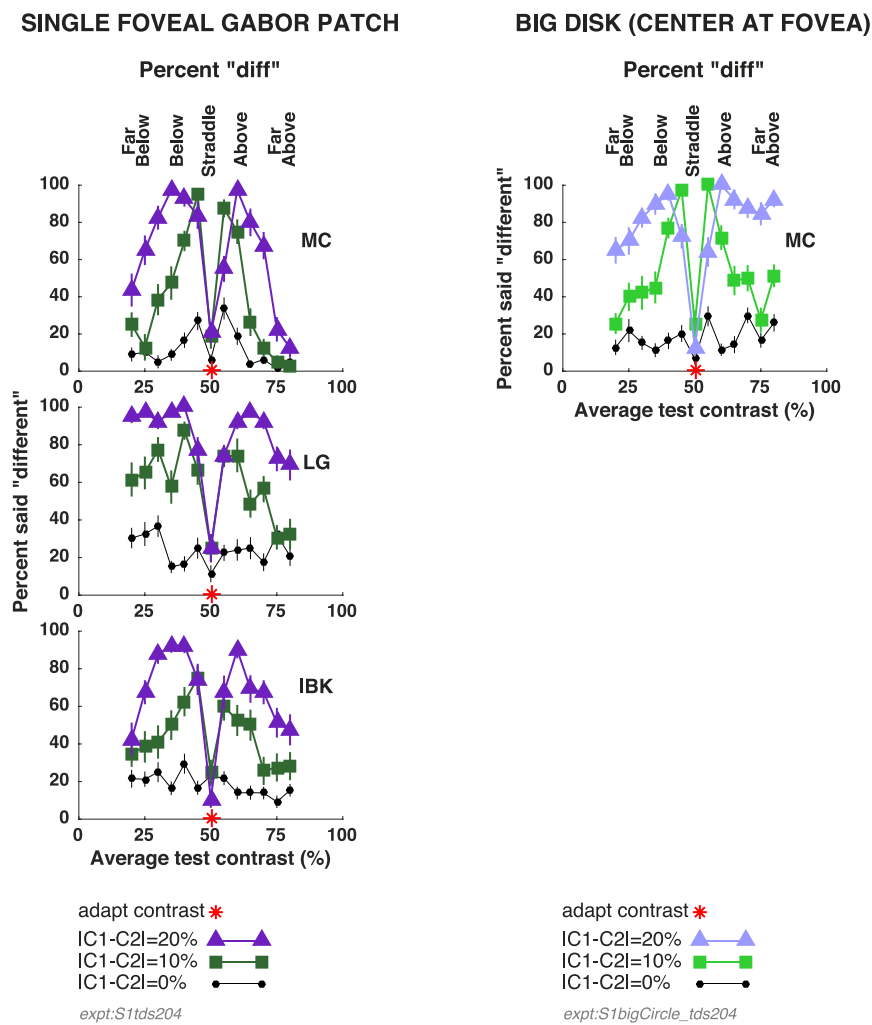
Figure 14. Results from first-order same–different two-interval experiments using a single Gabor patch (left column) or a big disk (right column). For the results plotted here, percentage "diff" is on the vertical axis and average test contrast is on the horizontal axis. Error bars show ± 1 standard error across blocks. For figures showing results plotted as d'-conservative and d'-unconfounded, see Figures C6, C10, and C11.

The procedure and patterns are very similar to those of the first-order same–different two-interval experiment of the previous section (Figures 7 and 8 and accompanying text). The differences are that in this new experiment the adapt contrast changed randomly from trial to trial and could assume any of the following values: 0% (gray), 25%, 50%, 75%, or 100%; and that the difference |C1 − C2| between the two test contrasts was always either 0% (for which the correct response was "same") or 15% (for which the correct response was "different"). See Appendix A for more details of the methods and procedures.

Results from this experiment, plotted as percentage "diff," are shown in Figure 12. (The other performance measures—d'-conservative and d'-unconfounded—are shown, along with percentage "diff," in Figures C1–

C3.) The horizontal axes show average test contrast. The value of the adapt contrast is indicated by a red asterisk on each horizontal axis and is labeled at the right end of the row. The position of the straddle-effect notch moves as the adapt contrast moves, following the adapt contrast so that the notch is always centered where the average test contrast equals the adapt contrast. The first-order results share this general and important property with second-order results.

Using a single Gabor patch or a big disk in first-order experiments

The first-order same–different two-interval paradigm allows us to use some spatial patterns that were

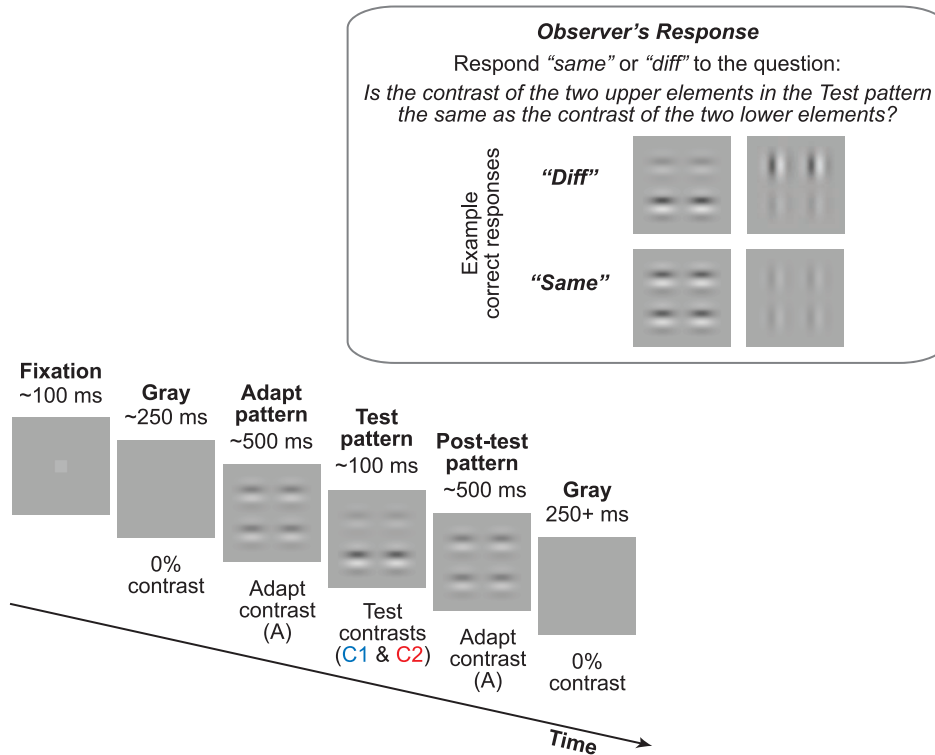## 2ND-ORDER: SAME-DIFFERENT, TWO POSITIONS (2x2 Gabor patches)

Figure 15. The stages of a typical trial in the second-order same–different two-position experiment. The test pattern is a 2 × 2 Gabor-patch grid. On half the trials, randomly chosen, all four Gabor patches in the test pattern have the same contrast. On the other half of the trials, the upper two Gabor patches in the test pattern have a different contrast from that of the lower two Gabor patches. (Thus, the second-order stripes are always horizontal in this experiment.) The adapt and posttest patterns are, as in the previous experiments, composed of four Gabor patches all at the same contrast. In these experiments, the durations of the adapt and posttest patterns are 500 ms, as in the first-order same–different experiments. The spatial frequencies, spatial positions, and Gabor-patch orientations are the same as in the previous experiments. The box in the upper right of the figure illustrates the response the observer is instructed to make by showing four different test patterns. Feedback as to the correctness of the response was provided.

logically impossible to use in our second-order orientation-identification task. So we tried experiments using two such patterns: a single Gabor patch (see Figure 13, middle) and a big uniform disk with soft edges (Figure 13, right).

The single Gabor patch seems of interest for two reasons. Since the 2 × 2 grid of Gabor patches used previously was fixated at the center, the individual patches were somewhat parafoveal. Using a single fixated Gabor patch allows us to investigate first-order vision at the fovea. A single Gabor patch is also a prototypical stimulus used to investigate first-order spatial vision.

The other new pattern, the big disk, was uniform throughout its center and had soft edges. We wondered whether quasiperiodic stimuli like Gabor patches are necessary to get a straddle effect, or can a straddle effect be obtained with a stimulus like a big disk?

Further details of the methods and procedures for these two experiments are given in Appendix A. Figures showing trial diagrams and contrast-versus-time profiles for these two experiments are shown in Appendix C (Figures C4 and C5 for the single Gabor patch, C8 and C9 for the big disk).

Figure 14 shows results using the single Gabor-patch pattern (left column) and the big-disk pattern (right column). The results in this figure are plotted as percentage "diff" versus average test contrast. The results plotted using both d'-conservative and d'-unconfounded are in Appendix C (Figures C6, C10, and C11).

Since only one observer completed the big-disk experiment before the lab was closed, these results with the big disk should be treated more cautiously than others. This observer was, however, very experienced and one of the two observers that performed all the other experiments presented in this article. There is no reason to suspect that there is anything peculiar about his results here.

Results from the experiment using a single Gabor patch (Figure 14, left column) and from the experiment using a big disk (Figure 14, right column) show the same general shape as those using a 2 × 2 grid of Gabor

2ND-ORDER: SAME-DIFFERENT, TWO POSITIONS (2x2 Gabor patches)

*Is the contrast of the two upper elements in the Test pattern
the same as or different from the contrast of the two lower elements?*
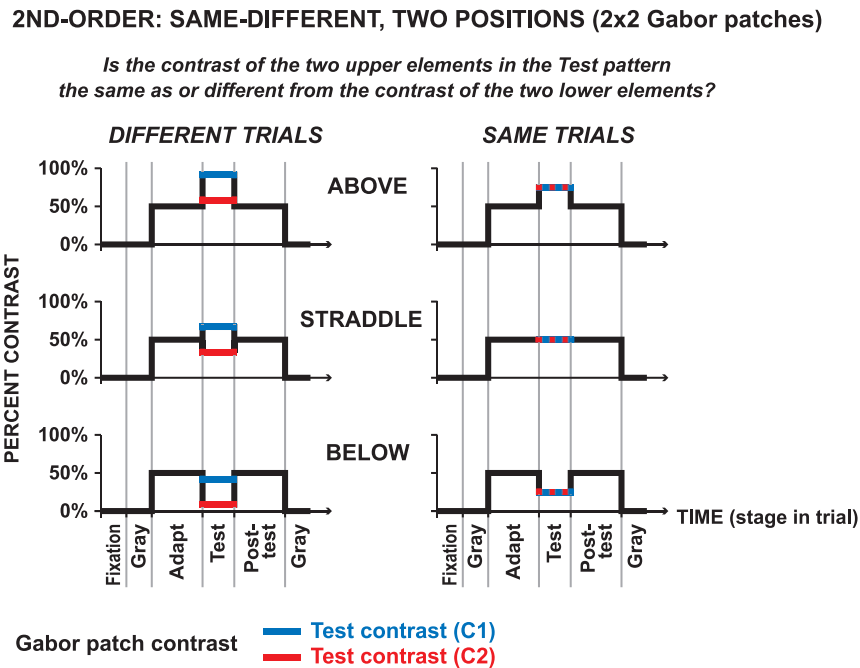

Figure 16. The contrast-versus-time profiles for six types of trial in the second-order same–different two-position experiment. The blue lines indicate test contrast C1 and the red lines indicate test contrast C2. (The upper two patches are at one of the two contrasts C1 or C2, and the lower two patches are at the other.) Thick black lines indicate contrasts of 0% (when the screen is homogeneous gray) and 50% (during the adapt and posttest patterns). The correct response is "different" for the trials in the left column and "same" for the trials in the right column. The top (bottom) row shows profiles for the cases where the test contrasts are both above (both below) the adapt contrast. The middle row shows cases where the average of the two test contrasts equals the adapt contrast. (Thus in the middle row of the right-hand column, the two test contrasts are both equal to the adapt contrast.)

patches (Figures 9 and 10). The curves show two peaks surrounding a center notch, with declining performance as one goes farther away from the peaks in either direction. All these results can be explained qualitatively by contrast-comparison and contrast-normalization processes. (A superimposed plot of results for a single Gabor patch and 2×2 Gabor patches—which are very similar—is shown in Figure C7.)

If the same contrast-comparison and contrast-normalization processes underlie all these results, one can use these results to suggest some likely spatial characteristics of these processes. In particular, neither second-order spatial channels nor any other kind of integration of contrast over a wide area is necessary for contrast-comparison and contrast-normalization processes to operate.

Further comparisons of first- and second-order spatial vision

As shown earlier (Figure 11) better performance occurs in the second-order orientation-identification experiment (gray stars) than in the first-order same–different two-interval experiment (colored symbols);

however, both show the straddle effect and Weber-law behavior consistent with contrast-comparison and contrast-normalization processes. Is the higher performance an intrinsic difference between first- and second-order visual pathways, or are there other differences due to decision and performance processes outside the visual pathways?

We try here to remove one difference between the two experiments that particularly worried us. The first-order experiment used a same–different task. Same–different tasks tend to lead to asymmetric response criteria and may produce substantial criterion variability, which itself lowers performance. (See Appendix B and Figure B1 for more about symmetric and asymmetric response criteria.) The second-order experiment used a two-alternative forced-choice task, which typically produces symmetric response criteria. Perhaps, therefore, the lower performance in our first-order experiment relative to our second-order experiment (Figure 11) results from the typical differences between same–different tasks and two-alternative tasks. To remove this difference between first- and second-order experiments, we studied the second-order pathway again, but this time using a same–different task.
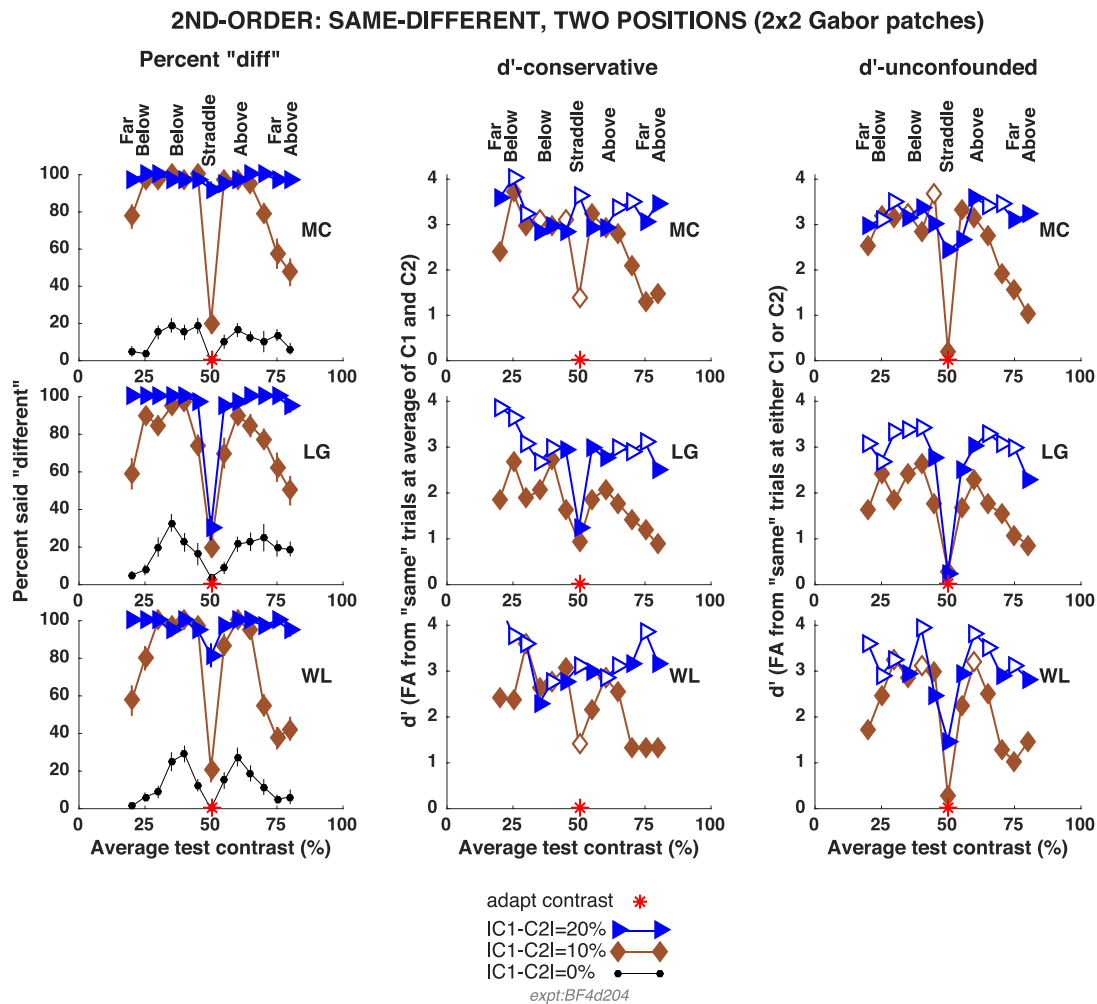
Figure 17. Results from the second-order same–different two-position experiment (using a 2 × 2 Gabor-patch grid) with observers MC, LG, and WL (three rows). The columns show three different measures of observer performance: percentage "diff", d'-conservative, and d'-unconfounded. Each curve gives the results for a constant-difference series (20% in dark-blue right-pointing triangles, 10% in brown diamonds, and 0% in small black dots). The error bars in the percentage "diff" plots show ±1 standard error across blocks. The open symbols in the d' plots show cases where the truncation rule had to be applied. See text and Appendix B for more details.

A new second-order spatial-vision experiment: Same–different, two positions

The same–different task we used for this new second-order experiment is illustrated in Figures 15 (trial diagram) and 16 (plots of contrast versus time). As shown in these figures, in each trial there is only a single interval, and the test pattern in that interval has four patches. On half the trials (randomly selected), all four patches in the test pattern have the same contrast. On the other half of the trials, the contrast of the upper two patches is different from the contrast of the lower two patches. The observer responds to indicate whether the contrast of the upper two patches is the same as or different from that of the lower two patches. An adapt pattern of 50% contrast was

presented immediately before the test pattern, and a posttest pattern of 50% contrast followed immediately after.

The results of this new second-order same–different two-position experiment are shown in Figures 17 and 18. These two versions of the experiment were the same except that the version in Figure 18 uses a wider range of average test contrasts and a lower test-contrast difference than the version in Figure 17. Qualitatively, all these results are very similar to the results of the second-order orientation-identification experiment (shown in Figure 4) and to the results of the first-order same–different two-interval experiment (shown in Figures 9 and 10). In particular, they show the straddle-effect notch in the middle and decreasing performance in both tails.
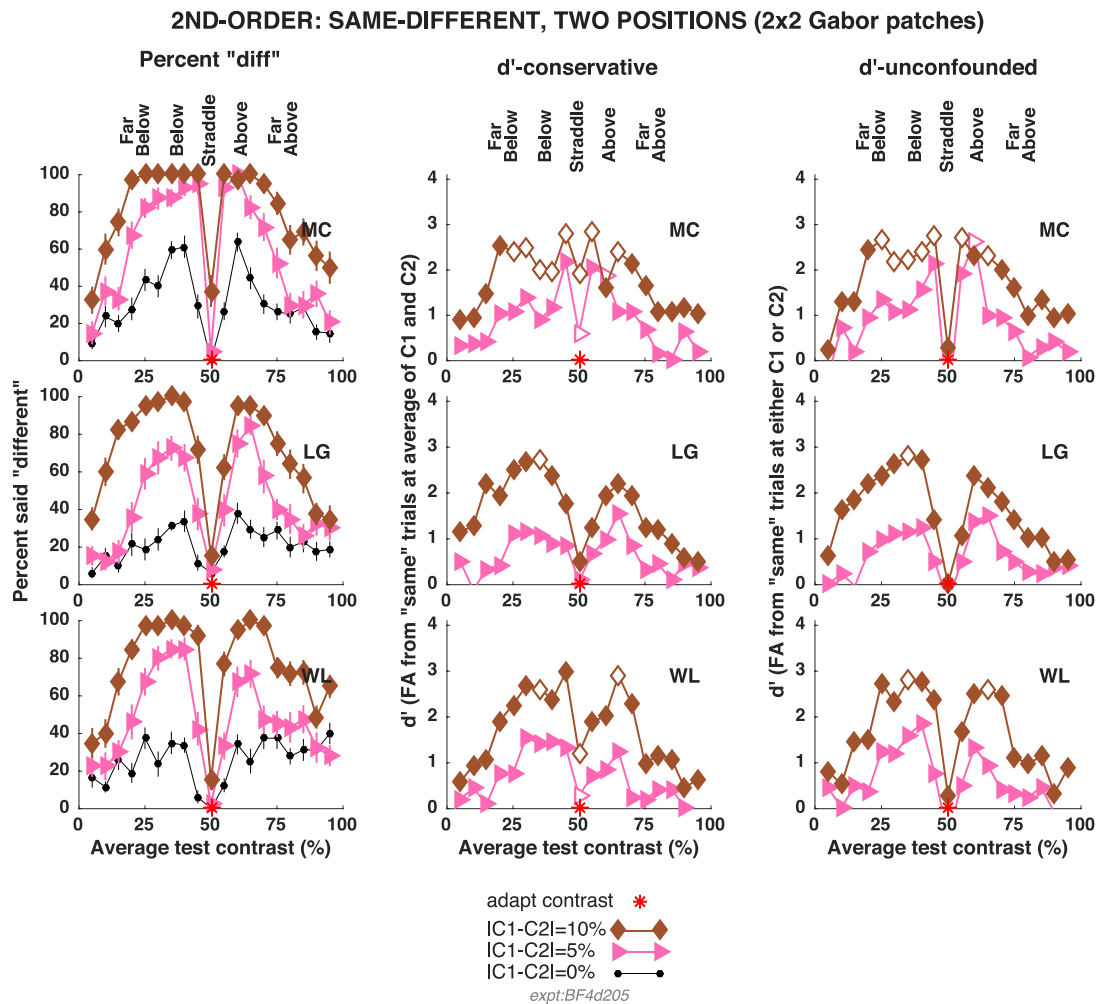
2ND-ORDER: SAME-DIFFERENT, TWO POSITIONS (2x2 Gabor patches)

Figure 18. Results from another version of the second-order same–different two-position experiment (using a 2 × 2 Gabor-patch grid) with the same observers MC, LG, and WL (three rows). This version used a wider range of average test contrasts and a lower test-contrast difference (5% in pink right-pointing triangles) than did the experiment in Figure 17. Other conventions are the same as in Figure 17.

Figure 19 compares the new second-order same–different two-position experiment (Figures 17 and 18) and the second-order orientation-identification experiment (Figure 4). Performance in the same–different task is somewhat lower than in the orientation-identification task, as we thought it might be. (There is some truncation in both experiments, but even paying attention only to cases where the lower point is not truncated shows the difference.)

Figure 20 compares the new second-order same–different two-position experiment (Figures 17 and 18) and the first-order same–different two-interval experiment (Figures 9 and 10). This figure shows that even when both second- and first-order experiments use a same–different procedure, performance in the second-order experiment (right-pointing triangles and diamonds) is somewhat higher than in the first-order experiment (upright triangles and squares).

Why is performance in the first-order experiment somewhat poorer than in the second-order experiment (Figure 20)? This may result from less effective processes operating in the first-order pathway; however, we are not convinced of this. The poorer performance might be due to differences in the details of the experiments. Even though they are both same–different tasks, in the second-order experiment the observer makes a judgment about Gabor patches that are presented in the same temporal interval, whereas in the first-order experiment the Gabor patches are presented in different temporal intervals. It is quite possible that in the first-order experiment there is some memory loss between the first and second temporal intervals. Thus, a memory component—rather than something different about underlying first- versus second-order processes—may be causing the slightly poorer performance.
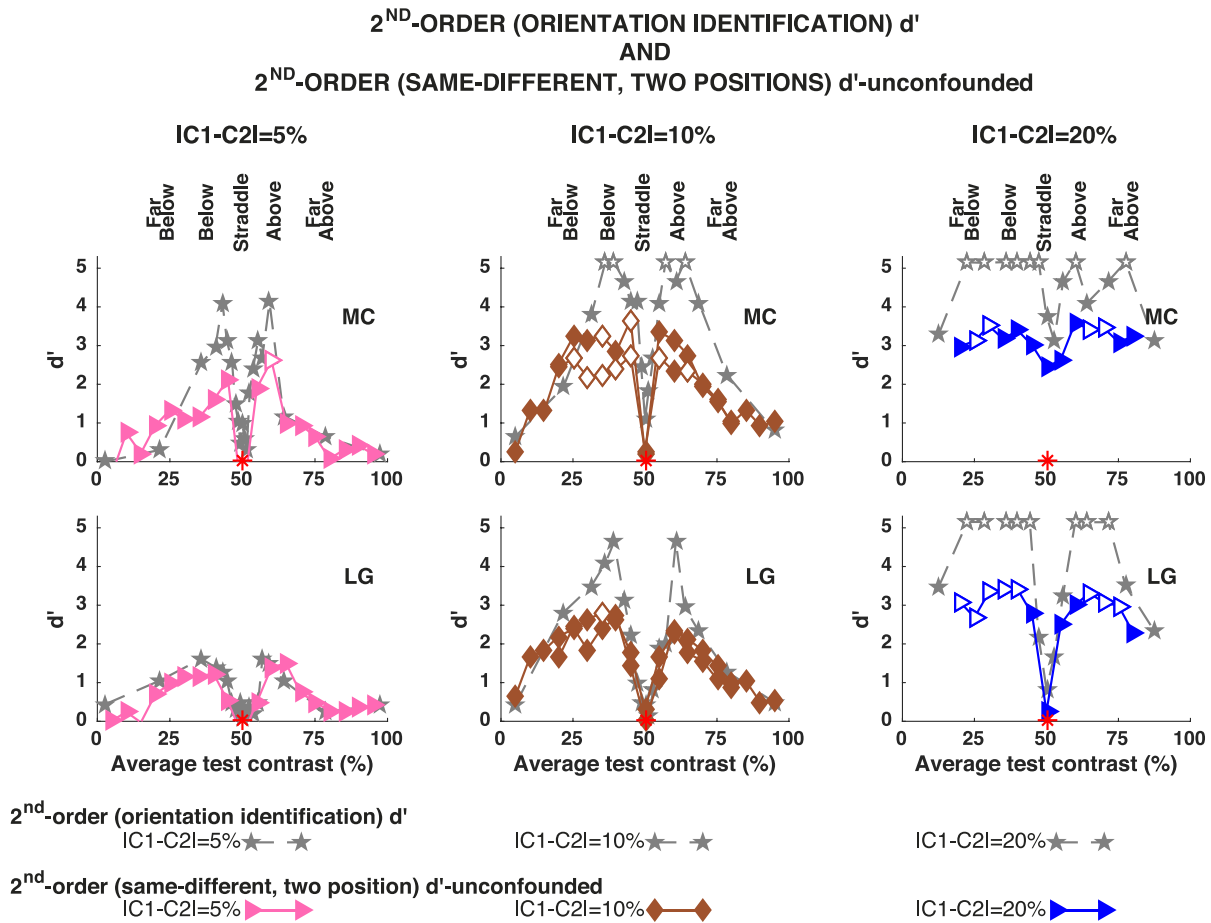
Figure 19. Direct comparison of results from the second-order same–different two-position experiment and results from the second-order orientation-identification experiment. Each experiment used a 2 × 2 Gabor-patch grid. The gray stars show d' values in the second-order orientation-identification experiment (originally shown in Figure 4 as colored symbols and repeated as gray stars in Figure 11). The pink, brown, and blue colored symbols (that are either right-pointing triangles or diamonds) show d'-unconfounded in the second-order same–different two-position experiment. (These d'-unconfounded values were previously shown in the right-hand columns of Figures 17 and 18.) The open symbols represent cases where truncation was used. There is some truncation in both experiments in this figure; but if you consider only the average test contrasts at which the colored symbols are solid (i.e., not truncated), you will see that the performance is generally lower for the colored symbols (second-order same–different two-position experiment) than for the gray stars (second-order orientation-identification experiment). This difference is particularly clear at the peaks of the curves.

Conclusions

We initially discovered the straddle effect using large (15 × 15) grids of Gabor patches with a second-order experimental task (second-order orientation identification). We presumed that the effect was second order, but since then we began to question this notion. In this article we have shown that the straddle effect is also found in first-order experiments.

Figure 5 shows idealized data and our current thoughts on the underlying processes. The straddle effect—the very poor performance when the two test contrasts straddle the adapt contrast—is consistent with a contrast-comparison process. Performance reaches maxima at average test contrasts somewhat below and

somewhat above the adapt contrast. For average test contrasts that are even smaller or even larger, performance decreases again. These decreases at the tails are consistent with a contrast-normalization process.

Although in our experiments there is no qualitative difference between the behaviors in first- and second-order pathways, there may be a quantitative difference. We have shown a quantitative difference between first- and second-order experiments even when both use same–different tasks (Figure 20). We are not sure, however, that the better performance in the second-order experiment results from an intrinsic difference between first- and second-order pathways; it may result instead from a greater memory demand in the first-order experiment.
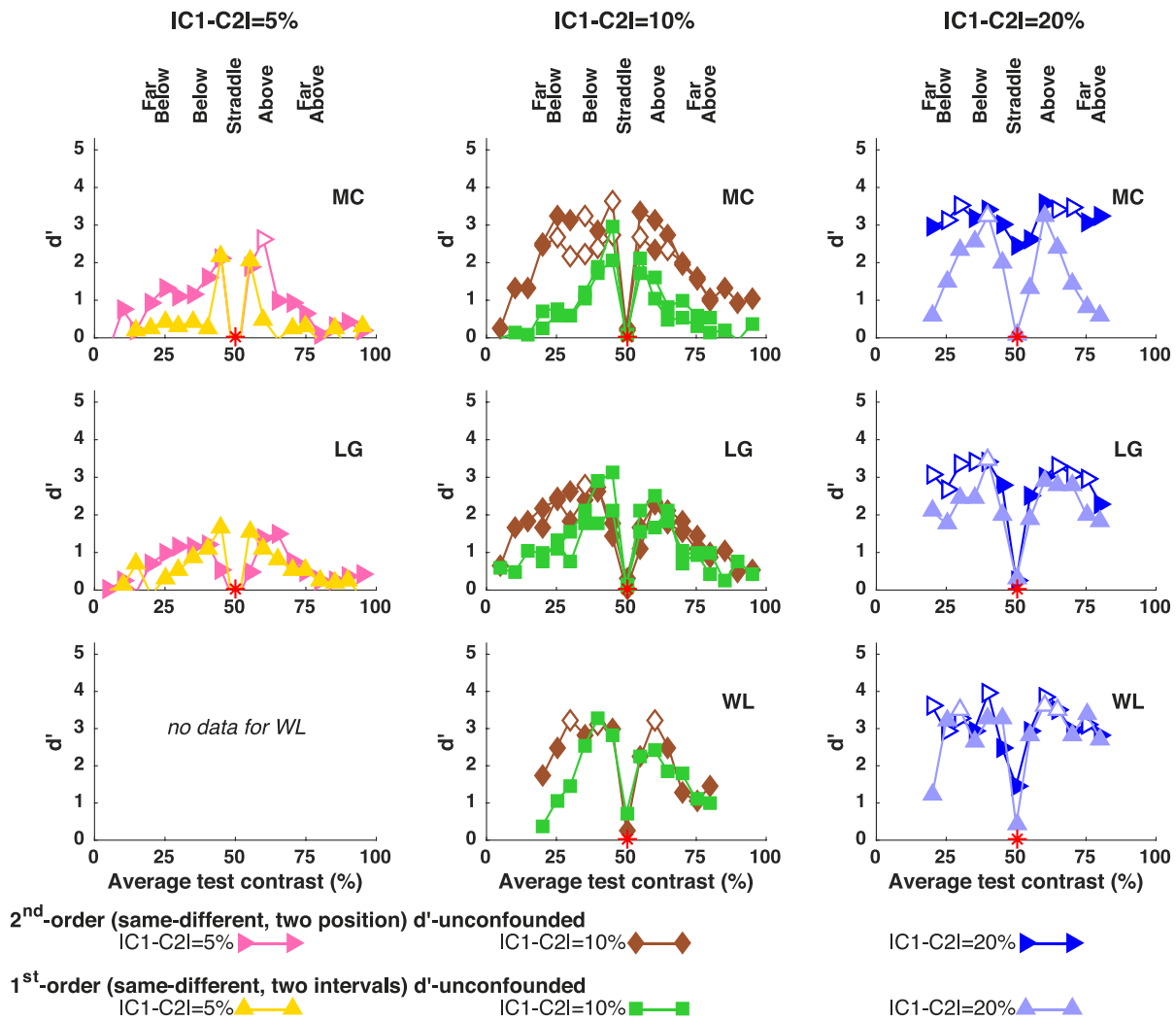
Figure 20. Direct comparison of results from the second-order same–different two-position experiment and the first-order same–different two-interval experiment, each using a 2 × 2 Gabor-patch grid. The vertical axis shows d'-unconfounded, and the horizontal axis shows average test contrast. The pink, brown, and dark-blue symbols (which are either right-pointing triangles or diamonds) show results from the second-order same–different two-position experiment and are identical to the colored symbols in Figure 19. The yellow, green, and light-blue symbols (which are either upright triangles or squares) show results from the first-order same–different two-interval experiment and are identical to the colored symbols in Figure 11. The open symbols represent cases where truncation was used. Performance is generally—but only slightly—higher in the second-order same–different two-position experiment than in the first-order same–different two-interval experiment. We suspect that this difference is due to the greater memory demands of the two-interval experiment than of the two-position experiment.

Overall, adapting to a particular contrast diminishes performance for some test patterns but improves it for many others. For example, after adaptation to 0% contrast, performance on a pattern with an average test contrast of about 60% is poor (Figure 12, top row). But after adaptation to 50% contrast, performance on the pattern with 60% average test contrast is very high (Figure 12, middle row). We suspect that the processes that produce the straddle effect are generally helpful, not detrimental, in the course of everyday vision.

It might be that for *all* abrupt changes in contrast, there is a rectification process whose result is that increases and decreases in contrast act similarly and are confusable. We suspect now that there is a shifting, rectifying contrast-comparison process acting on all visual patterns, not just gratings and textures.

Acknowledgments

Commercial relationships: none.
Corresponding author: Norma V. Graham.
Email: nvg1@columbia.edu.
Address: Department of Psychology, Columbia University, New York, NY, USA.

References

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436.

Carandini, M., & Heeger, D. J. (2011). Normalization as a canonical neural computation. *Nature Reviews Neuroscience, 13*(1), 51–62.

Foley, J. M. (2011). Forward–backward masking of contrast patterns: The role of transients. *Journal of Vision, 11*(9):15, 1–24, https://doi.org/10.1167/11.9.15. [PubMed] [Article]

Graham, N. (2011). Beyond multiple pattern analyzers modeled as linear filters (as classical V1 simple cells): Useful additions of the last 25 years. *Vision Research, 51,* 1397–1430.

Graham, N., & Sutter, A. (2000). Normalization: Contrast-gain control in simple (Fourier) and complex (non-Fourier) pathways of pattern vision. *Vision Research, 40,* 2737–2761.

Graham, N., & Wolfson, S. S. (2007). Exploring contrast-controlled adaptation processes in human vision (with help from Buffy the Vampire Slayer). In M. Jenkins & L. Harris (Eds.), *Computational vision in neural and machine systems* (pp. 9–47). Cambridge, UK: Cambridge University Press.

Graham, N., & Wolfson, S. S. (2013). Two visual contrast processes: One new, one old. In C. Chubb, B. A. Dosher, Z.-L. Lu, & R. M. Shiffrin (Eds.), *Human information processing: Vision, memory, and attention* (pp. 13–27). Washington, DC: American Psychological Association.

Graham, N., Wolfson, S. S., Kwok, I., & Grinshpun, B. (2010). "Buffy contrast adaptation" with a single Gabor patch. *Journal of Vision, 10*(7): 1386, https://doi.org/10.1167/10.7.1386. [Abstract]

Hochberg, J. E. (1978). *Perception* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Kachinsky, E. S., Smith, V. C., & Pokorny, J. (2003). Discrimination and identification of luminance contrast stimuli. *Journal of Vision, 3*(10):2, 599–609, https://doi.org/10.1167/3.10.2. [PubMed] [Article]

Landy, M. S. (2013). Texture analysis and perception. In J. S. Werner & L. M. Chalupa (Eds.), *The new visual neurosciences* (pp. 639–652). Cambridge, MA: MIT Press.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442.

Pokorny, J. (2011). Review: Steady and pulsed pedestals, the how and why of post-receptoral pathway separation. *Journal of Vision, 11*(5):7, 1–23, https://doi.org/10.1167/11.5.7. [PubMed] [Article]

Solomon, S. G., & Kohn, A. (2014). Moving sensory adaptation beyond suppressive effects in single neurons. *Current Biology, 24*(20), R1012–R1022.

Wolfson, S. S., & Graham, N. (2007). An unusual kind of contrast adaptation: Shifting a contrast-comparison level. *Journal of Vision, 7*(8):12, 1–7, https://doi.org/10.1167/7.8.12. [PubMed] [Article]

Wolfson, S. S., & Graham, N. (2009). Two contrast adaptation processes: Contrast normalization and shifting, rectifying, contrast comparison. *Journal of Vision, 9*(4):30, 1–23, https://doi.org/10.1167/9.4.30. [PubMed] [Article]

Appendix A: Detailed description of methods and procedures

The first phrase in each item in the following list gives the observer's task. We used three different tasks in the experiments in this article: second-order orientation identification, same–different two positions, and same–different two intervals.

The middle phrase in each item in the list gives the spatial pattern used in each experiment. Three different patterns were used in these experiments: a 2×2 grid of Gabor patches, a single Gabor patch, and a single big disk. Only one of these patterns (the 2×2 grid of Gabor patches) was used with all three tasks. The other two patterns were used to investigate secondary points.

The last phrase in each item in the list identifies the figures illustrating that experiment. Table A1 summarizes the details of all the experiments.

List of five experiments

Second order

- Orientation identification—2×2 Gabor patches—Figures 1 and 2
- Same–different two positions—2×2 Gabor patches—Figures 15 and 16

First order

- Same–different two intervals—2×2 Gabor patches—Figures 7 and 8
- Same–different two intervals—single Gabor patch—Figures C4 and C5
- Same–different two intervals—single big disk—Figures C8 and C9

Patterns

In the experiments described in this article, three spatial patterns were used. Figure 13 shows each of these spatial patterns, all at the same scale:

- A 2×2 grid of horizontal or vertical Gabor patches (all of the same orientation; horizontal is shown in Figure 13), with fixation at the center of the four patches.
- A single horizontal or vertical Gabor patch (horizontal is shown in Figure 13), with fixation at the center of the patch.
- A single big disk, with fixation at the center of the disk.

Details of the Gabor patches

A Gabor patch is a sinusoidal grating windowed by a two-dimensional Gaussian function. The sinusoidal grating in our Gabor patches had a period of 0.5° (32 pixels), which corresponded to a spatial frequency of 2 c/°. The Gaussian window had a full width at half height of 0.5° (32 pixels). Each Gabor patch was

truncated at $1° \times 1°$ (64×64 pixels) at the viewing distance of 90 cm. (Viewing distance and therefore spatial dimensions are approximate because observer head movements were not constrained.)

The orientation of a single-Gabor-patch pattern and of all four patches in a 2×2 grid pattern was either vertical or horizontal throughout an individual trial but could vary between trials. The positive zero crossing of the sinusoid was always at the center of each patch; the "dark bar" of the sinusoid was to the left of center for vertical patches and on top for horizontal patches.

The contrast of a Gabor patch was computed by taking the difference between the luminance at the peak of the Gaussian and the mean luminance of the pattern, and then dividing that difference by the mean luminance. All Gabor patches had a mean luminance of 55 cd/m². The single Gabor patch or the 2×2 grid of Gabor patches was centered within a $16° \times 16°$ ($1{,}024 \times 1{,}024$ pixel) gray square at 55 cd/m². Therefore, the mean luminance of the whole pattern was also 55 cd/m². (Further details of the Gabor-patch patterns appear in the subsection More about the adapt, test, and posttest patterns, and their relation to one another.)

Details of the big-disk pattern

The uniform center of the disk was 12° in diameter and there was a soft edge of a half-cosine shape forming a 1° wide annulus around the center. Thus the total disk subtended 14°.

The disk itself was centered within a $16° \times 16°$ ($1{,}024 \times 1{,}024$ pixel) gray square at 55 cd/m². This 55 cd/m² will be referred to as the background luminance for the big-disk pattern, as the luminance of the disk was in general greater than 55 cd/m² and thus the mean luminance of the full big-disk pattern was in general greater than 55 cd/m².

The contrast of the big disk was computed by taking the difference between the luminance in the central area of the disk and the background luminance, and then dividing that difference by the background luminance.

When the big disk was an adapt or posttest pattern, it was always an increment on the background. A test pattern, however, might be either an increment or a decrement from the adapt pattern.

More about the adapt, test, and posttest patterns, and their relation to one another

The adapt, test, and posttest patterns on a given trial all had identical spatial characteristics. The contrasts in the adapt and posttest patterns were also identical, but the contrast (or contrasts in some cases) in the test pattern was generally different.

Row	Fig. #	Test contrast differences	Average test contrasts	Adapt contrasts (also Post-Test contrast)	Pattern	Pattern details	In lab expt name
2ND-ORDER: ORIENTATION IDENTIFICATION (2x2 Gabor patches)							
1	3, 4	5% 10% 15% 20% and more	many & depends on test contrast difference	50%	2x2 Gabor patches (1 deg center-to-center distance)	Gabor patch spatial frequency is 2 cycles-per-degree; full-width-at-half-	BF4i52
1ST-ORDER: SAME-DIFFERENT, TWO INTERVALS (2x2 Gabor patches)							
2	9	0% 10% 20%	20% 25% 30% 35% … 80%	50%	2x2 Gabor patches	2 cpd; 0.5 deg	S4tds204
3	10	0% 5% 10%	5% 10% 15% 20% … 95%	50%			S4tds205
COMPARES 2ND-ORDER (ORIENTATION IDENTIFICATION) TO 1ST-ORDER (SAME-DIFFERENT, TWO INTERVALS)							
4	11	*compares row 1 to rows 2 & 3*					
1ST-ORDER: SAME-DIFFERENT, TWO INTERVALS (2x2 Gabor patches), varying adapt contrast							
5	12, C1, C2, C3	0% 15%	21 different ones btwn 7.5% and 92.5%	0% 25% 50% 75% 100%	2x2 Gabor patches	2 cpd; 0.5 deg	S4tds242
1ST-ORDER: SAME-DIFFERENT, TWO INTERVALS (1 Gabor patch)							
6	14, C6	0% 10% 20%	20% 25% 30% 35% … 80%	50%	1 foveal Gabor patch	2 cpd; 0.5 deg	S1tds204
COMPARES 1 GABOR PATCH TO 2x2 GABOR PATCHES (BOTH ARE 1ST-ORDER, SAME-DIFFERENT, TWO INTERVALS)							
7	C7	*compares row 6 to row 2*					
1ST-ORDER: SAME-DIFFERENT, TWO INTERVALS (1 big disk)							
8	14, C10	0% 10% 20% [rel lum 0 0.1 0.2]	20% 25% 30% 35% … 80% [rel lum 1.2 1.25 1.3 1.35 … 1.8]	50% [rel lum 1.5]	1 big soft edged disk	about 12 deg	S1bigCircle tds204
9	C11	0% 5% 10% [rel lum 0 0.05 0.1]	5% 10% 15% 20% … 95% [rel lum 1.05, 1.1, 1.15, 1.2, … 1.95]	50% [rel lum 1.5]			S1bigCircle tds205
2ND-ORDER: SAME-DIFFERENT, TWO POSITIONS (2x2 Gabor patches)							
10	17	0% 10% 20%	20% 25% 30% 35% … 80%	50%	2x2 Gabor patches	2 cpd; 0.5 deg	BF4d204
11	18	0% 5% 10%	5% 10% 15% 20% … 95%	50%			BF4d205
COMPARES ORIENTATION IDENTIFICATION TO SAME-DIFFERENT, TWO POSITIONS (BOTH ARE 2ND-ORDER)							
12	19	*compares row 1 to rows 10 & 11*					
COMPARES 2ND-ORDER TO 1ST-ORDER (BOTH ARE SAME-DIFFERENT, TWO INTERVALS)							
13	20	*compares rows 10 & 11 to rows 2 & 3*					

→

Row	Fixation duration	Gray duration (after Fixation)	Adapt duration (also Post-Test duration)	Test duration	Gray duration (after Post-Test)	For each observer, the number of trials per point in figures LG	MC	BSG	IBK	SYP	WL	Trials per block	Number of blocks run LG	MC	BSG	IBK	SYP	WL
1	1000ms	506ms	1000ms	94 ms	106ms+	100	102					145	100	102				
						Number of "same" trials/point (Number of "different" trials/point)												
2	106ms	247ms	506ms	94 ms	247ms+	40 (80)	40 (80)	32 (64)		20 (40)	40 (80)	104	20	20	16		10	20
3						86 (172)	80 (160)					152	43	40				
5	106ms	247ms	506ms	94 ms	247ms+	100 (100)	100 (100)					210	100	100				
6	106ms	247ms	506ms	94 ms	247ms+	40 (80)	40 (80)			40 (80)		104	20	20			20	
8	106ms	247ms	506ms	94 ms	247ms+		40 (80)					104		20				
9							52 (104)					152		26				
10	106ms	247ms	506ms	94 ms	247ms+	40 (80)	40 (80)				40 (80)	104	20	20				20
11						40 (80)	40 (80)				40 (80)	152	20	20				20

Table A1. Details of the experiments. All the figures listed in the column labeled "Fig. #" are figures showing experimental results. The adapt pattern's Gabor-patch orientation was always 0° or 90°. The test pattern's Gabor-patch orientation was the same as the adapt pattern's. The position of the test pattern's Gabor patches was the same as that of the adapt pattern's. The posttest pattern was identical to the adapt pattern. Feedback as to the correctness of the response was provided. Note that for the big-disk experiment (rows 8 and 9), relative luminance values are given in addition to contrast values; see Figure C9 for the relationship between relative luminance and contrast. Note that for the first-order same–different two-interval experiment varying adapt contrast (row 5), all adapt contrasts were intermixed in every block; the 21 different average test contrasts used in this experiment were 7.5%, 12.5%, 17.5%, 21.25%, 25%, 28.75%, 32.5%, 37.5%, 42.5%, 46.25%, 50%, 53.75%, 57.5%, 62.5%, 67.5%, 71.25%, 75%, 78.75%, 82.5%, 87.5%, and 92.5%. Further details of the methods and procedures are described in the text of Appendix A.

Only two Gabor-patch orientations were used in these experiments: vertical and horizontal. When the pattern was a 2×2 grid of Gabor patches, all individual Gabor patches in the grid were identical in orientation. Although fixed within any individual trial, the Gabor-patch orientation varied randomly from trial to trial.

In every trial of these experiments there were two test contrasts, C1 and C2, either in the one test pattern of a single-interval trial or in two different intervals of a two-interval trial.

In the second-order orientation-identification experiment, the test pattern could have either horizontally or vertically defined stripes. However, in the second-order same–different two-position experiment, the test pattern always had horizontally defined contrast stripes. (The response in that experiment was whether the upper and lower rows were of the same contrast or not.)

Alternate description of our patterns in terms of backgrounds and probes

In this article we describe our experiments in terms of adapt and test patterns, where the adapt pattern is *not* present during the test pattern. Alternately, the test pattern here could be described as the adapt (background) pattern *plus* an added (or subtracted) probe pattern. The relationship of these experiments here to earlier experiments in the literature using the background–probe terminology is discussed in Wolfson and Graham (2009), at the end of the Procedure subsection under Methods.

Durations

The duration of the test patterns was approximately 100 ms. Within any one experiment, the duration of the adapt, test, and posttest patterns stayed constant. The duration of the adapt (and posttest) patterns was always approximately either 1 s (for the second-order orientation-identification experiment) or 0.5 s (for the same–different experiments); see Table A1 for exact values. Previous work (Graham & Wolfson, 2013) has shown that for the conditions here and the conclusions of this study, the difference between 0.5- and 1-s adapt-pattern durations does not matter. We used the shorter durations in most experiments to diminish the burden on the observers by speeding up the collection of data.

Some terminology: Average test contrast, test-contrast difference, and constant-difference series

The relationship on any individual trial between the two test contrasts C1 and C2 and the adapt contrast A is critical. Thus we use short terms to describe the relationship succinctly.

The *average test contrast* is the mean of C1 and C2—that is, $(C1 + C2)/2$.

When one of the two test contrasts is greater than the adapt contrast and the other is less than the adapt contrast, we call it a Straddle test pattern, since the test contrasts straddle the adapt contrast. When the average test contrast is somewhat greater than the adapt contrast, we call it an Above test pattern. When the average test contrast is much greater than the adapt contrast, we call it a Far Above test pattern. Similarly, when the average test contrast is less than the adapt contrast we call it a Below or a Far Below test pattern.

The absolute value $|C1 - C2|$ of the difference between the two contrast values is called the *test-contrast difference*.

To show the results of these experiments, we frequently plot performance on several curves. Each curve shows the results for a series of test patterns in which the test-contrast difference is constant. We call each such series a *constant-difference series*.

Fixation patterns

Two fixation patterns were used:

- A fixation "point," namely a small square filled with a lighter gray than the mean luminance as shown to scale in, for example, Figure 1. This was used for all the 2×2 Gabor-patch grid experiments and the big-disk experiment.
- An outline square, shown to scale in Figure C4. This was used for the single-Gabor-patch experiment.

The duration of the fixation pattern for each experiment is given in Table A1. In the initial instructions, each observer was told to fixate in the middle of the screen throughout a trial as indicated by the fixation pattern.

Gray screen

At times during these experiments the observer was looking at an entirely gray screen. For the gray screen, the whole $16° \times 16°$ area was at 55 cd/m^2—that is, equal to the mean luminance in the case of Gabor-patch experiments and to the background luminance in the case of the big-disk experiment. Note that a 0% contrast adapt or test pattern is exactly the same as a gray screen.

Three observer tasks

Second-order orientation identification

The inset at the right of Figure 1 shows example test patterns in the second-order orientation-identification task, with correct responses indicated. The observer's
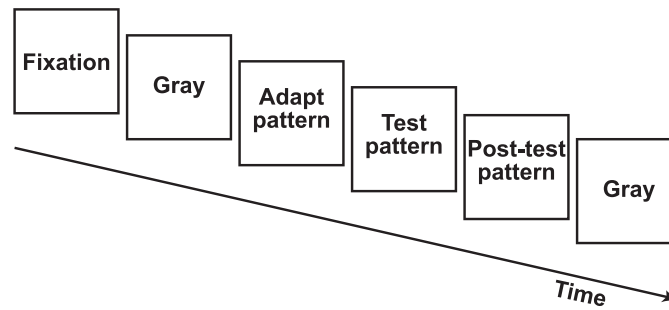
Figure A1. Generic trial diagram showing each stage in a single interval in any of the experiments here. Diagrams for specific experiments are shown in Figures 1, 7, 15, C4, and C8.

task is a conventional identification task with two alternatives: On each trial, the observer is required to identify which of two possible categories the test pattern falls into. More specifically, the observer is required to respond with whether the two different test contrasts in the 2×2 Gabor-patch test pattern define horizontal or vertical stripes. The orientation of the contrast-defined stripes is technically a spatially second-order orientation.

In this article, we use this task only with 2×2 grids of Gabor patches.

Second-order same–different two positions

The inset at upper right of Figure 15 shows example test patterns in the second-order same–different two-position task, with correct responses indicated. This is a conventional same–different task, using single-interval trials. What must be judged as same or different are the Gabor patches occupying two different positions in the test pattern. More specifically, the test pattern is a 2×2 grid of Gabor patches, and the observer is required to respond indicating whether the contrast of the top two Gabor patches is the same as or different from the contrast of the bottom two.

We use this task only with 2×2 grids of Gabor patches. Note also that we only use horizontal contrast-defined arrangements of the 2×2 Gabor grids, as shown in Figure 15.

First-order same–different two intervals

The first-order same–different two-interval task is also a conventional same–different task, except it uses two intervals in each trial, and the aspect to be judged is whether the test patterns in the two intervals are identical or different. (If they are different, they are different only in contrast.)

All three types of patterns are used with this task: the 2×2 grid of Gabor patches (Figure 7), the single Gabor patch (Figure C4), and the big disk (Figure C8). All three test patterns are show at the same scale in Figure 13.

Structure of each trial

Each trial in all the experiments started when the observer pressed the 0 key or space bar.

Each trial in a particular experiment contained either a single interval or two intervals. Every interval in every experiment in this article had the same general timing and pattern sequence, which is described later.

Each trial ended with the observer giving a response and receiving feedback as to the correctness of the response.

Between trials, the screen was a homogeneous gray until the observer started the next trial by pressing the 0 key or space bar.

Sequence and timing of stages within an interval

Figure A1 shows an abstract diagram of each stage in any single interval of any of the experiments here. The diagrams for each experiment are shown in more detail in Figures 1, 7, 15, C4, and C8. The precise timing of each stage for each experiment is listed in Table A1.

The interval begins with a fixation pattern followed by a short gray period. This is immediately followed by a brief adapt pattern.

There follows immediately, for approximately 100 ms, a test pattern. In all these experiments, the test pattern is identical to the adapt pattern in spatial characteristics but generally contains a different contrast or contrasts.

The test pattern is followed immediately by a brief posttest pattern, which in these experiments is always identical to the adapt pattern (in duration as well as spatial characteristics and contrast).

If an interval is the first in a two-interval trial, there is a short fixed-duration gray period at the end of the first interval. Then the second interval begins with a short fixation pattern and proceeds as the first interval did.

If an interval is the second in a two-interval trial, or if it is the only interval in a single-interval trial, there is a final gray period that is terminated by the observer

pressing a response key. (The observer is not allowed to respond until a short period of time has passed.)

Immediately following the key-press response, there is a tone giving feedback as to the correctness of the response.

Structure of blocks of trials: Intermixing different patterns and contrasts within a block

The contrast of the adapt pattern was almost always 50%. The one exception is the experiment varying adapt contrast in the first-order same–different two-interval task (Figures 12, C1, C2, and C3). In this experiment, five different adapt contrasts were used. The adapt contrasts are listed in Table A1. The set of test contrasts used was generally different from experiment to experiment. In Table A1, the set of test contrasts is given by specifying what values were used for the average test contrasts and the test-contrast differences.

Some experiments had two versions. The only difference between two versions is the particular test contrasts used. For example, we ran two versions of the first-order same–different two-interval experiment: The results from one version are plotted in Figure 9, and the results from the other version are plotted in Figure 10. In any block of a version of an experiment (e.g., Figure 9), all test contrasts used in that version were intermixed. To put it another way, all test-contrast differences and all average test contrasts used in a version of an experiment were presented in a random order within a block. All contrast combinations were shown an equal number of times within a block. Similarly, in the only experiment where more than one adapt contrast was used, those adapt contrasts were randomly intermixed in a block.

All the substantive characteristics (e.g., test and adapt contrasts) were randomly sampled *without* replacement across trials to make the numbers of trials of each the same. But the nonsubstantive characteristics were sampled *with* replacement. The nonsubstantive characteristics include whether Gabor-patch orientation was horizontal or vertical; whether the contrast-defined stripes were horizontal or vertical; which interval of a two-interval trial which test pattern came in; and which position of the two-position same–different experiment had which test contrast.

Table A1 lists the trials per point in the results figures. Two numbers are given for the same–different experiments: number of trials per point when the correct answer was "same" and number when the correct answer was "different." In a block, we wanted always to have the number of Same trials (trials on which the correct response was "same") be equal to the number of Diff trials (trials on which the correct

response was "different"). Thus, in experiments that had two nonzero constant-difference series, there are twice as many trials per point for the 0% constant-difference series (the Same trials) as for the nonzero (e.g., 10% and 20%, as in Figure 9) constant-difference series (the Diff trials). In the second-order orientation-identification experiment, only one number is give, which is the total number of trials per point.

Observers

This research adhered to the tenets of the Declaration of Helsinki. All observers were Columbia University students who were long-term observers in our laboratory. They gave their informed consent and were paid for their work.

Four observers (SYP, BSG, IBK, and WL) had extensive knowledge of the experiments, but the other two (LG, MC) did not. The two relatively inexperienced observers (LG and MC) are the two who participated in almost every experiment here. All reported having normal or corrected-to-normal visual acuity, and wore their prescribed glasses if any had been prescribed.

Equipment

The experiments were run on Macintosh G4s with NEC MultiSync FE992 CRT monitors and ATI Radeon 9000 video cards. The resolution was 1,280 × 1,024 pixels at 85 Hz. The mean luminance of the Gabor-patch experiments, the background luminance in the big-disk experiment, and the luminance of the blank gray screen when no pattern was present were all the same and approximately 55 cd/m^2. The room was dark except for the CRT screen and a dim lamp. Each CRT's lookup table was linearized. Stimuli were generated and presented using MathWorks' MATLAB with the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). The actual duration of each stage (adapt, test, etc.) of each trial was checked for accuracy. In the event of inaccuracy—typically an extra refresh of the CRT (1/85 of a second)—the trial was excluded from subsequent analysis. Such inaccuracies occurred on less than 1% of the trials.

Appendix B: About computing *d′* values

We use conventional signal-detection theory to summarize and compare performances in the different experiments. Descriptions of signal-detection theory

can now be found in many places, as it has become widely used to disentangle response biases from sensitivity estimates. The source we refer to here is Macmillan and Creelman (2005).

The ways we calculate d' in this article are not standard because the experimental and theoretical frameworks here are somewhat different from most standard cases. The reader already familiar with signal-detection theory can skip ahead to Three ways of calculating d' values for all our same–different experiments.

A simple form of signal-detection theory can be used for our purposes. It starts with two probability distributions as shown in Figure B1 (for two different criteria). These two distributions are often referred to as the signal and noise distributions, and they represent the distribution of values along some internal dimension resulting from, respectively, signal and noise stimuli. It is a hypothetical dimension on which the observer is making a two-alternative decision. This decision is assumed to be based on a criterion—that is, the observer gives one response (e.g., "Yes, there is a signal") if the internal response on this trial is above the criterion, and the other response (e.g., "No, there is not a signal, there is only noise") if the internal response on this trial is below the criterion.

The internal-response dimension can be labeled in various ways depending on the application of the model—for example, internal sensory magnitude, perceptual strength, memory strength, or decision variable. The two classes of stimuli—frequently labeled signal and noise in the abstract model—are more generally two mutually exclusive and exhaustive subsets of all stimuli being considered.

We will make the conventional assumption of equal-variance normal (Gaussian) distributions. The sensitivity parameter d' will be defined as the difference between the mean of the signal and noise distributions divided by the standard deviation of either distribution. For this equal-variance Gaussian model, the value of d' is then easily computed (e.g., Macmillan & Creelman, 2005, pg. 8, equation 1.5) as

$$d' = z(H) - z(F),$$

where the symbols in the equation are defined as follows:

- z(x) is the z score characterizing the standard normal distribution that is associated with the cumulative probability x.
- H (for "hit") is the conditional probability that an observer responds "yes" given that a signal stimulus has occurred on a trial. It is called the *hit rate*.
- F (for "false alarm") is the conditional probability that an observer responds "yes" given that a noise stimulus has occurred on a trial. It is called the *false-alarm rate*.

A third quantity (for "miss") is often useful. It is the conditional probability that an observer responds "no" given that a signal stimulus has occurred on a trial—that is, $M = 1 - H$.

(An aside for those who are interested: Rather than assuming equal-variance Gaussian signal and noise distributions, another approach would be to measure full receiver operating characteristic curves—see, e.g., Macmillan & Creelman, 2005, pp. 57–63—and from these curves to calculate estimates of the individual shapes of both distributions. But measuring full curves would have greatly expanded the amount of data that needed to be collected, and from what we know about very similar tasks, it is highly unlikely to affect the conclusions here.)

This general signal-detection-theory framework can be used for many tasks with two alternative responses, including all those we used here: the orientation-identification task, the same–different two-position task, and the same–different two-interval task. How we used this framework to calculate d' values in the three tasks follows. First, however, a general point that applies to all our graphs of d' values, no matter how they were computed.

Truncation of d' values to deal with proportions of 1 and 0

When the measured hit rate or false-alarm rate was 100% or 0%, we made an adjustment to keep the d' values from becoming infinite. More specifically, we used the following simple rule to truncate the values of d':

If all trials were correct, we adjusted the number correct to be equal to the total number of trials minus 0.5 (e.g., 40 correct out of 40 was adjusted to be 39.5 correct out of 40). If no trials were correct, we adjusted the number correct to equal 0.5 rather than 0 (e.g., 0 correct out of 40 was adjusted to be 0.5 correct out of 40). Points adjusted in this manner are indicated by open symbols in the figures.

Thus, truncation ceilings will be the same for two different experiments if the same number of trials was used in both. This is true in Figure 20 but is generally not true. The number of trials in each experiment is listed in Table A1.

Calculating d' for the second-order orientation-identification experiment

The second-order orientation-identification experiment can be seen as an example of the signal-detection-theory case already described and shown in Figure B1. There are two mutually exclusive and exhaustive subsets of the trials (those with horizontal contrast-defined stripes vs. those with vertical contrast-defined

SYMMETRIC CRITERION: P(False Alarms) = P(Misses)
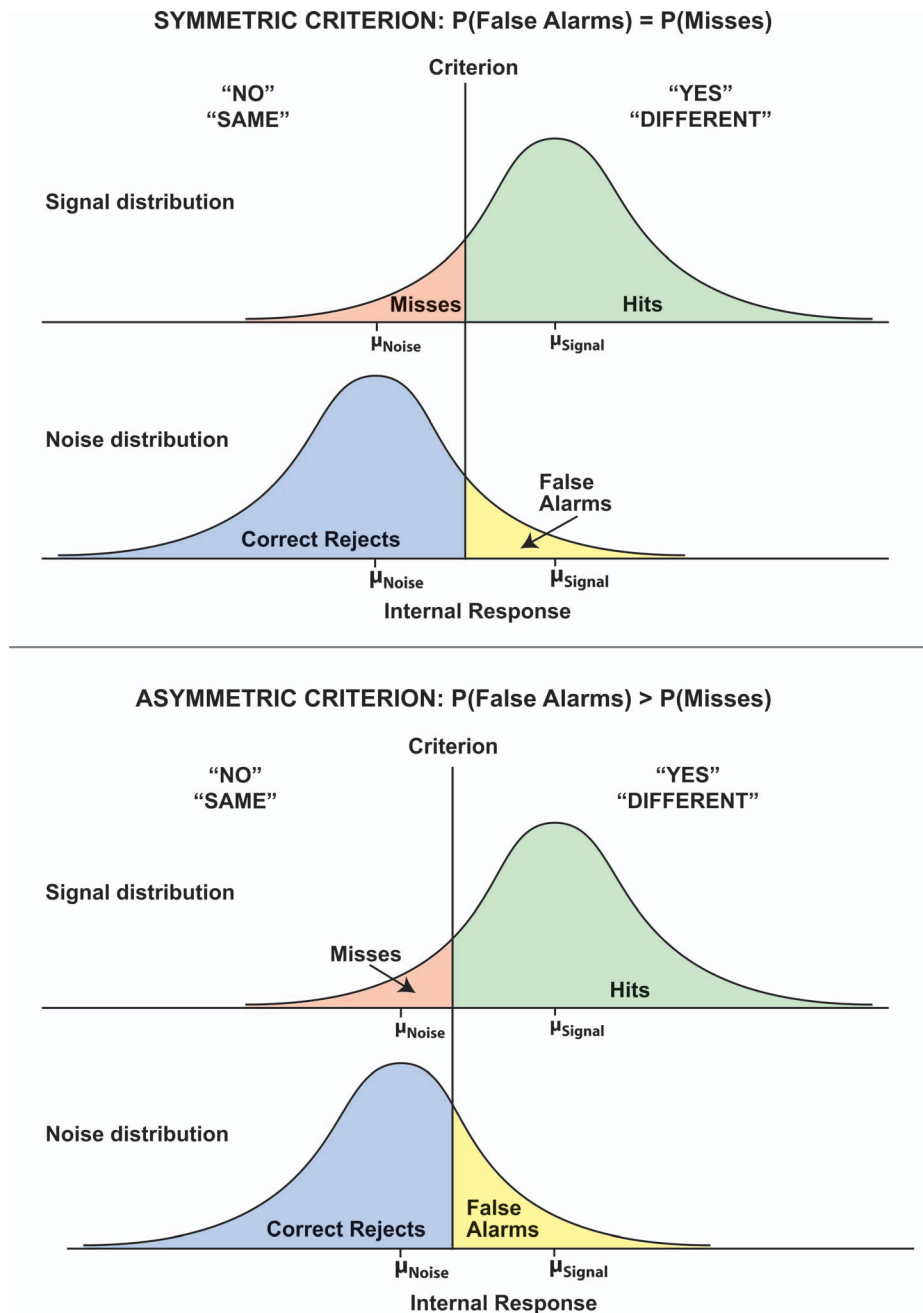
ASYMMETRIC CRITERION: P(False Alarms) > P(Misses)

Figure B1. Illustration of a simple form of signal-detection theory. Each panel contains two equal-variance Gaussians and a criterion line. The top panel shows the line for a symmetric criterion and the bottom for an asymmetric criterion. The horizontal axis is labeled "Internal Response"; other terms like decisions axis or internal sensory magnitude dimension are also used for this concept.

stripes) and two corresponding responses ("horizontal" or "vertical"). To view this as a signal-versus-noise case, you arbitrarily choose one of the two subsets (e.g., vertical contrast-defined stripes) to be signal and the other (e.g., horizontal contrast-defined stripes) to be noise.

Our second-order orientation-identification experiment can be treated even more simply than the signal-detection-theory case described. This simplicity arises because the observer can reasonably be assumed to be

unbiased—that is, to favor neither response over the other (as shown in Figure B1, top panel). This is a reasonable assumption, since the presentation probabilities are equal (the horizontal and vertical contrast-defined stripes occurred equally often) and neither kind of error seems worse than the other (to respond "horizontal" when it should be "vertical," or vice versa). This lack of bias is represented by placing the criterion on the internal axis so that false alarms and misses are equally likely. In symbols, the criterion is set

to make F = M = 1 − H. This is frequently called the *symmetric criterion*, because the line representing the criterion falls exactly in the middle of the two distributions, making the whole drawing symmetric around that criterion line.

The calculation of d' from an equal-variance Gaussian model with a symmetric criterion (symmetric bias) reduces to (Macmillan & Creelman, 2005, p. 9, equation 1.7):

$$d' = 2 \times z(\text{probability correct})$$

These are the d' values plotted in Figure 4, right column.

Three ways of calculating *d′* values for all our same–different experiments

Our same–different experiments can also be seen as an example of the single-detection-theory case already described and shown in Figure B1. The stimuli for which the correct response is "diff" are called the signal stimuli (and denoted by Diff). The stimuli for which the correct response is "same" are called the noise stimuli (and denoted by Same). Then the equation for d' becomes:

$$d' = z(H) - z(F),$$

where H = probability that observer responds "diff" on Diff trials and F = probability that observer responds "diff" on Same trials.

In cases like our same–different experiments, observers are likely to show an asymmetric bias such that false alarms and misses are not equally likely. In symbols: F ≠ M and therefore F ≠ 1 − H. In fact, it is easy to find cases in our results where that happened. Thus we cannot use here the symmetric criterion (Figure B1, top panel); we will have to allow the criterion to be asymmetric (Figure B1, bottom panel).

For the same–different experiments we considered three different ways to estimate d' values. The three different estimates of the true d' value came from three different ways of calculating the false-alarm rate. They are described in the following, after two general points:

First, no matter which of these three ways one uses to calculate d', the important general results from these experiments remain true, in particular: A straddle effect exists—that is there is a notch in the graph of performance versus average test contrast centered at the point where the average test contrast equals the adapt contrast; and there is maximal performance for patterns containing both test contrasts somewhat below (or both somewhat above) the adapt contrast. Once both the test contrasts become far from the adapt contrast, performance is below maximum.

Second, only the second and third ways of calculating d' values are used for our figures here (e.g., Figure 9, middle and right columns). Why the first way is not shown is described in the next subsection.

First way of calculating *d′* values (a way that is biased in this context)

This first way is the simplest calculation of false-alarm rate and therefore of d'. It assumes that the false-alarm rate is the same for all the trials in a given experiment no matter how the test-contrast values differ from trial to trial. This single false-alarm rate is estimated as the proportion of all Same trials on which the observer responded "diff." In our figures here, we do not show the d' values calculated this first way for two reasons, one practical and one theoretical.

The practical reason: As mentioned, this first way of calculating d' uses the same estimated false-alarm rate for every test stimulus. Thus, the calculated d' values for a set of stimuli will be monotonic with the hit rates for those stimuli. These hit rates are plotted in all our relevant figures in the left column (the large colored symbols in the left column of, e.g., Figure 9). In particular, the d' values calculated this first way will necessarily have all the peaks and valleys that the hit rates in the left column show.

The theoretical reason: The assumption underlying this first way of calculating d'—that the false-alarm rate on every trial is identical—is easily shown to disagree with the observers' performances. First note that all "diff" answers on Same trials are, by definition, false alarms. In our relevant figures these "diff" answers on Same trials are plotted as the bottom curve in the left column (small black dots connected by a black line in, for example, Figure 9). If the assumption underlying this first way of calculating d' were true, then these bottom curves would be straight horizontal lines except for statistical variability. In our data, however, these bottom curves are systematically bent away from a straight line, showing a dip when the average test contrast equals the adapt contrast (marked with a red asterisk)—that is, for the Same-Straddle stimulus.

Second way of calculating *d′* values: *d′*-conservative

The second way of calculating d' produces values shown in the middle column of results figures here, for example, Figure 9. We call values calculated this way *d′-conservative* for reasons described later. This second way of calculating d' assumes that the false-alarm rate varies from trial to trial because it is determined by the average test contrast on each trial.
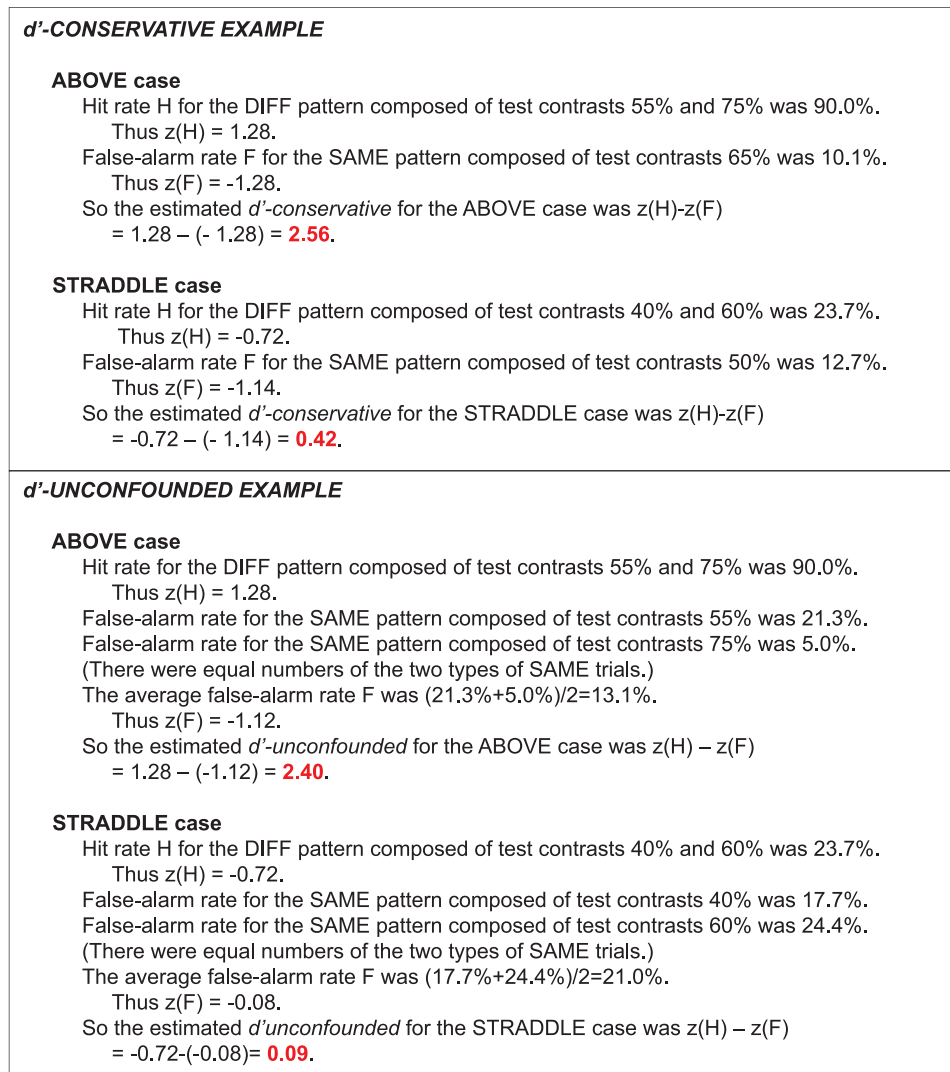
---

d'-CONSERVATIVE EXAMPLE

ABOVE case
    Hit rate H for the DIFF pattern composed of test contrasts 55% and 75% was 90.0%.
        Thus z(H) = 1.28.
    False-alarm rate F for the SAME pattern composed of test contrasts 65% was 10.1%.
        Thus z(F) = -1.28.
    So the estimated *d'-conservative* for the ABOVE case was z(H)-z(F)
        = 1.28 – (- 1.28) = **2.56**.

STRADDLE case
    Hit rate H for the DIFF pattern composed of test contrasts 40% and 60% was 23.7%.
        Thus z(H) = -0.72.
    False-alarm rate F for the SAME pattern composed of test contrasts 50% was 12.7%.
        Thus z(F) = -1.14.
    So the estimated *d'-conservative* for the STRADDLE case was z(H)-z(F)
        = -0.72 – (- 1.14) = **0.42**.

---

d'-UNCONFOUNDED EXAMPLE

ABOVE case
    Hit rate for the DIFF pattern composed of test contrasts 55% and 75% was 90.0%.
        Thus z(H) = 1.28.
    False-alarm rate for the SAME pattern composed of test contrasts 55% was 21.3%.
    False-alarm rate for the SAME pattern composed of test contrasts 75% was 5.0%.
    (There were equal numbers of the two types of SAME trials.)
    The average false-alarm rate F was (21.3%+5.0%)/2=13.1%.
        Thus z(F) = -1.12.
    So the estimated *d'-unconfounded* for the ABOVE case was z(H) – z(F)
        = 1.28 – (-1.12) = **2.40**.

STRADDLE case
    Hit rate H for the DIFF pattern composed of test contrasts 40% and 60% was 23.7%.
        Thus z(H) = -0.72.
    False-alarm rate for the SAME pattern composed of test contrasts 40% was 17.7%.
    False-alarm rate for the SAME pattern composed of test contrasts 60% was 24.4%.
    (There were equal numbers of the two types of SAME trials.)
    The average false-alarm rate F was (17.7%+24.4%)/2=21.0%.
        Thus z(F) = -0.08.
    So the estimated *d'unconfounded* for the STRADDLE case was z(H) – z(F)
        = -0.72-(-0.08)= **0.09**.

Figure B2. Step-by-step examples of calculating d'-conservative and d'-unconfounded.

Numerical example of calculating d'-conservative

We consider numerical examples for two kinds of trials that are particularly important in interpreting our experimental results. These example calculations will be gone through here in the text but can be found in a more succinct format in Figure B2. The top half of the figure illustrates the second way of calculating d' (d'-conservative). The bottom half illustrates the third way of calculating d' (d'-unconfounded), which will be described later. The numbers used in these examples are based on observer MC's performance, shown in the top row of Figure 9. The results of these numerical examples are summarized in Figure B3.

First we consider a Diff-Above pattern that produces near-peak performance: the Diff-Above pattern with a test-contrast difference of 20% (purple triangles) and an average test contrast of 65%. This pattern contains test contrasts of 55% and 75%. The observer's measured hit rate H for this pattern (that is, the percentage of trials

of this pattern for which observer MC responded "diff") was 90%, thus leading to a z(H) = 1.28. In order to calculate d'-conservative, the false-alarm rate on all trials of this Diff-Above pattern is determined by the average test contrast of 65%. So we estimate the false-alarm rate F by using the measured performance for the Same test pattern having both its test contrasts equal to 65% (small black dots). The observer's

Results from calculations shown in Fig. Y2

	ABOVE (55%,75%)	STRADDLE (40%,60%)
d'-conservative	**2.56**	**0.42**
d'-unconfounded	**2.40**	**0.09**

Figure B3. Summary of results from numerical examples in Figure B2.

measured false-alarm rate on that pattern was 10.1%, thus producing $z(F) = -1.28$. Thus the estimated d'-conservative for this Diff-Above pattern $= z(H) - z(F) = 1.28 - (-1.28) = 2.56$.

Now we consider the Diff-Straddle pattern, again with a 20% test-contrast difference—that is, the pattern having the two test contrasts 40% and 60% (and thus an average test contrast equal to the adapt contrast of 50%). The measured hit rate H for this Diff-Straddle pattern was 23.7%, which produces $z(H) = -0.72$. The false-alarm rate F for calculating d'-conservative is the probability of an observer's responding "diff" to the Same-Straddle pattern (i.e., the pattern having both test contrasts, and average test contrast, equal to the adapt contrast of 50%). The measured performance F in this example was 12.7%, leading to a $z(F) = -1.14$. Thus the estimated d'-conservative for this Diff-Straddle pattern $= z(H) - z(F) = -0.72 - (-1.14) = 0.42$.

Third way of calculating d' values: d'-unconfounded

The difference between d'-unconfounded and d'-conservative is in the way the false-alarm rate is calculated. The hit rate is the same for both calculations. We use the term d'-unconfounded for reasons that will be described later.

Numerical example of calculating d'-unconfounded

We use the same Diff-Above pattern used in the prior example: the Diff-Above pattern with a 20% test-contrast difference (purple triangles) and an average test contrast of 65% (test contrasts 55% and 75%). The hit rate H for calculating d'-unconfounded is identical to that for calculating d'-conservative (90%), and thus $z(H) = 1.28$. The false-alarm rate F for d'-unconfounded is calculated from the pool of all Same trials in which the test contrasts both equaled 55% (called Same-55%) or both equaled 75% (called Same-75%). The proportion of times the observer responded "diff" (gave a false alarm) on Same-55% trials was 21.3%. For Same-75% trials the false-alarm rate was 5%. Since the numbers of trials of Same-55% and Same-75% were identical, we can average the two false-alarm rates (21.3% and 5.0%) to get the false-alarm rate F on the whole pool (13.1%). Thus, $z(F) = -1.12$. So the estimated d'-unconfounded for the Diff-Above case is $z(H) - z(F) = 1.28 - (-1.12) = 2.40$. This d'-unconfounded estimate is quite similar to the d'-conservative estimate for this pattern.

Consider now the Diff-Straddle pattern containing test contrasts of 40% and 60%. The hit rate H for calculating d'-unconfounded is identical to that for calculating d'-conservative (23.7%), and accordingly $z(H) = -0.72$. The false-alarm rate F for d'-unconfounded is calculated from the pool of all Same trials in which the test contrasts both equaled 40% or both equaled 60%. The proportion of times the observer said "diff" (gave a false alarm) on Same-40% trials was 17.7%. For Same-60% trials the false-alarm rate was 24.4%. Since the numbers of trials of Same-40% and Same-60% were identical, we can average the two false-alarm rates (17.7% and 24.4%) to get the false-alarm rate F on the whole pool (21.0%). Thus, $z(F) = -0.08$. So the estimated d'-unconfounded for the Diff-Straddle case is $z(H) - z(F) = -0.72 - (-0.08) = 0.09$. This d'-unconfounded estimate is very different from the d'-conservative estimate for this pattern.

Some exceptions when calculating d'-unconfounded

We did not anticipate calculating d'-unconfounded before we ran the experiments reported here. Thus, for some experiments we did not use the Same patterns that would be necessary to compute d'-unconfounded exactly as above. Where we had not used a necessary Same pattern, we estimated the performance on that missing one from the performance on the Same pattern that was nearest (in test contrast) to the missing one. These exceptions occurred in two places in our experiments:

- All test patterns in constant-difference series characterized by a contrast difference of 5%.
- The very leftmost and very rightmost points of all constant-difference series.

The special nature of the Same-Straddle pattern and the observer's response

The empirical probability of an observer's responding "diff" to Same trials is plotted in the left column of, for example, Figure 9 with small black dots. Notice that probability is *not* constant across average test contrast. In particular, it shows a dip precisely where the dip occurs in the curves for the Diff patterns—that is, where the average test contrast equals the adapt contrast. At first we were puzzled by this dip in the curve for Same trials. But then we realized there was a unique feature that the observer could use to almost always correctly identify the Same-Straddle trials. Notice that in five of the six kinds of trials illustrated in Figure 8 (e.g., the two kinds in the top row of the figure), there are four changes in contrast during each interval: at the start of the adapt pattern, at the start of the test pattern, at the end of the test pattern, and at the end of the posttest pattern. Thus, during each two-interval trial, there are $2 \times 4 = 8$ transitions.

However for the Same-Straddle kind of trial (middle row, right column), the two test contrasts are both equal to the adapt contrast, and thus there is no change in contrast at either the start or the end of the test pattern. Thus there are only two transitions per interval, for a total of four transitions per trial.

According to our observers (and in agreement with our own observations), many fewer transitions are immediately and effortlessly perceived on Same-Straddle trials than on other kinds of trials. And thus, since the observers can relatively immediately and effortlessly see that there are only four transitions, they quickly learn from feedback that the correct answer on this trial is "same." So it is little surprise after all that the observers respond "diff" on very few of these trials, thereby producing a dip in the middle of the curves for Same trials in the left column of, for example, Figure 9.

There is a minor exception to this. That exception occurs in cases when the test contrasts are very, very close to the adapt contrast and thus not seen as different from it. We have rarely used test contrasts that close to the adapt contrast, and thus to make the remainder of this discussion simpler, we ignore these cases.

When one applies signal-detection theory in the way we have so far done in this appendix, one is assuming that there is something like a single perceptual dimension on which the observer is setting a criterion and making a judgment. In same–different tasks that dimension is the similarity between the two things being judged. However, this model cannot account for the behavior described on the Same-Straddle trials.

A more appropriate model for our task is to suppose two relevant perceptual dimensions: the usual one based on perceived similarity between two intervals of a trial, and then a special one based on something like the number of perceived transitions in a trial (or perceived flicker in a trial).

The usual dimension—perceptual similarity—is the more useful dimension for an observer who wants to be correct on five of the six kinds of trials we have been discussing. On the sixth kind, however—the Same-Straddle case—the number of transitions is the more useful dimension. It is extremely useful, since it leads the observer to respond almost perfectly (i.e., almost always respond "same").

We never told the observers that there was such a feature as number of transitions. However, feedback was always used, and feedback appears to have quickly taught observers to respond "same" to those special trials.

Why do we use the term d'-conservative?

Calculating d'-conservative systematically uses smaller false-alarm rate estimates $z(F)$ for the Straddle pattern than for the other patterns because it uses the false-alarm rate from the Same-Straddle pattern. This elevates the d'-conservative estimates for the Straddle patterns—by subtracting a smaller $z(F)$—relative to the other two ways of calculating d'. And thus d'-conservative plots almost always show a shallower (more conservative) straddle-effect notch than d'-unconfounded plots.

One might argue that d'-conservative is not actually conservative with respect to reality. This argument would make sense if the assumption used in calculating d'-conservative were true—that is, if the false-alarm rate operating on a given trial were determined by the average test contrast on that trial. However, as described in the previous subsection, we are reasonably certain that Same-Straddle trials are in a class by themselves and do not reflect the false-alarm rate on any other trial.

Why do we use the term d'-unconfounded?

To look at the same facts we have just mentioned from a different perspective, we can say that the d'-unconfounded calculations show deeper straddle-effect notches for the following reason: They do *not* allow the confounding effect of the special nature of the Same-Straddle case (the unusual number of transitions) to completely dominate the calculation of d' at the center of the notch.

Therefore, we think that d'-unconfounded is the most valid of our three ways of calculating d' and that it most closely reflects the similarity dimension. It is this similarity dimension that can show the effects of the contrast-comparison and contrast-normalization processes, and it is these spatial processes that we are studying here.

Note that we do not mean to imply that this way of calculating d' produces totally unconfounded values, but it seemed to be the best term we could think of.

More detail for interested readers: A person might suggest that it would be even better (less confounded) to prevent the Same-Straddle case from ever entering into any d' estimates. Calculating d'-unconfounded does allow it to enter into the d' calculation for Diff patterns that have one test contrast equal to the adapt contrast. But even there its effect is diluted by the patterns that have the other test contrast (not equal to the adapt contrast). We did some calculations excluding the Same-Straddle pattern entirely and it made little difference except to make the straddle effect notch even deeper occasionally.

Appendix C: More plots from the first-order same–different experiments

This appendix shows more plots from the first-order same–different experiments described in the main text. These plots should be looked at in conjunction with the main text.
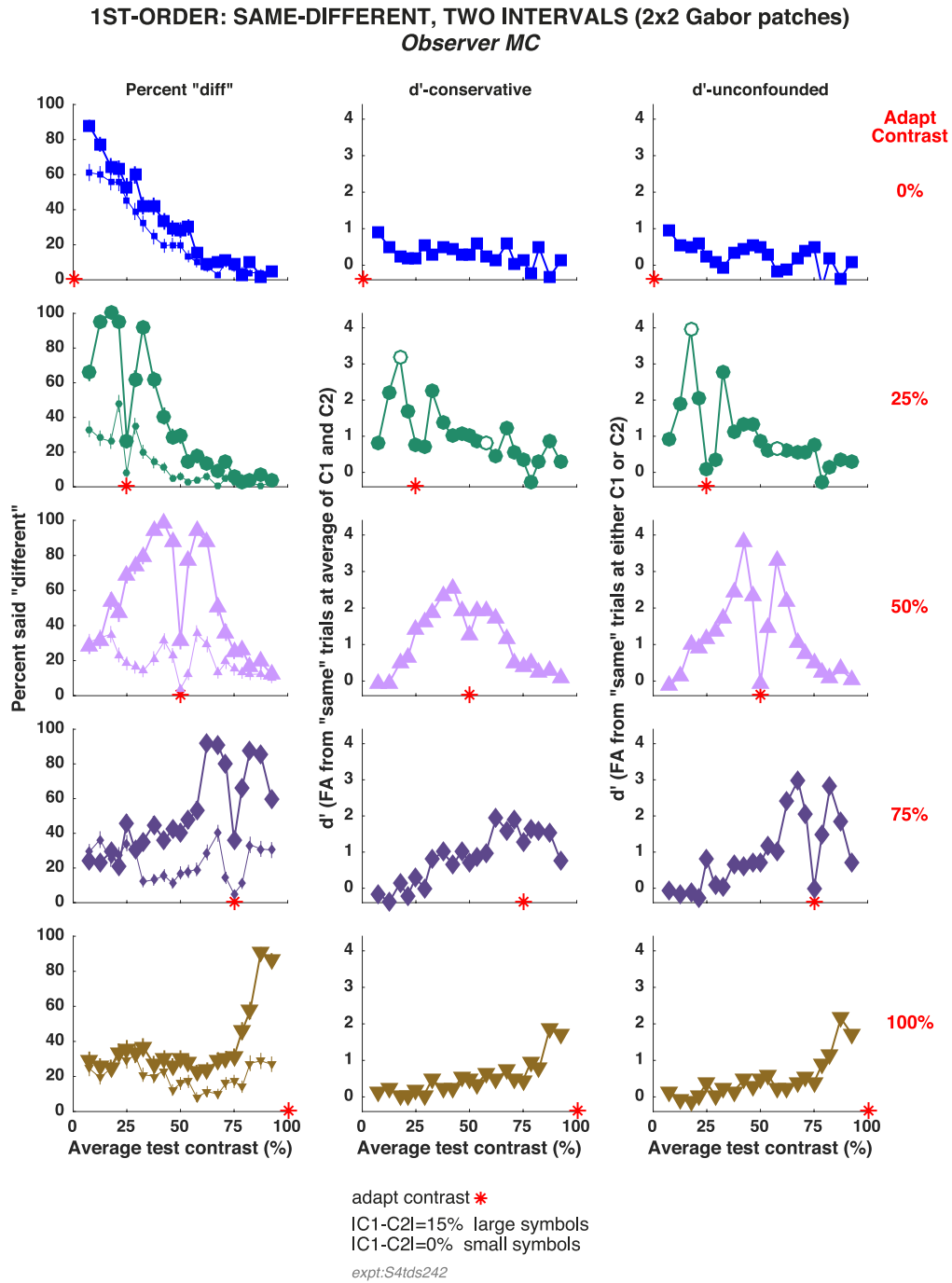
Experiment varying adapt contrast

Figure C1. Results for observer MC. Results of varying adapt contrast in a first-order same–different two-interval experiment (using a 2 × 2 Gabor-patch grid) from 0% (top row) to 100% (bottom row). They are plotted as three different measures: percentage "diff," d'-conservative, and d'-unconfounded, in the left, middle, and right columns, respectively. The horizontal axis gives average test contrast, and the value of the adapt contrast is indicated by a red asterisk and labeled on the right edge of each row. The left column of this figure was shown in Figure 12 (left column).
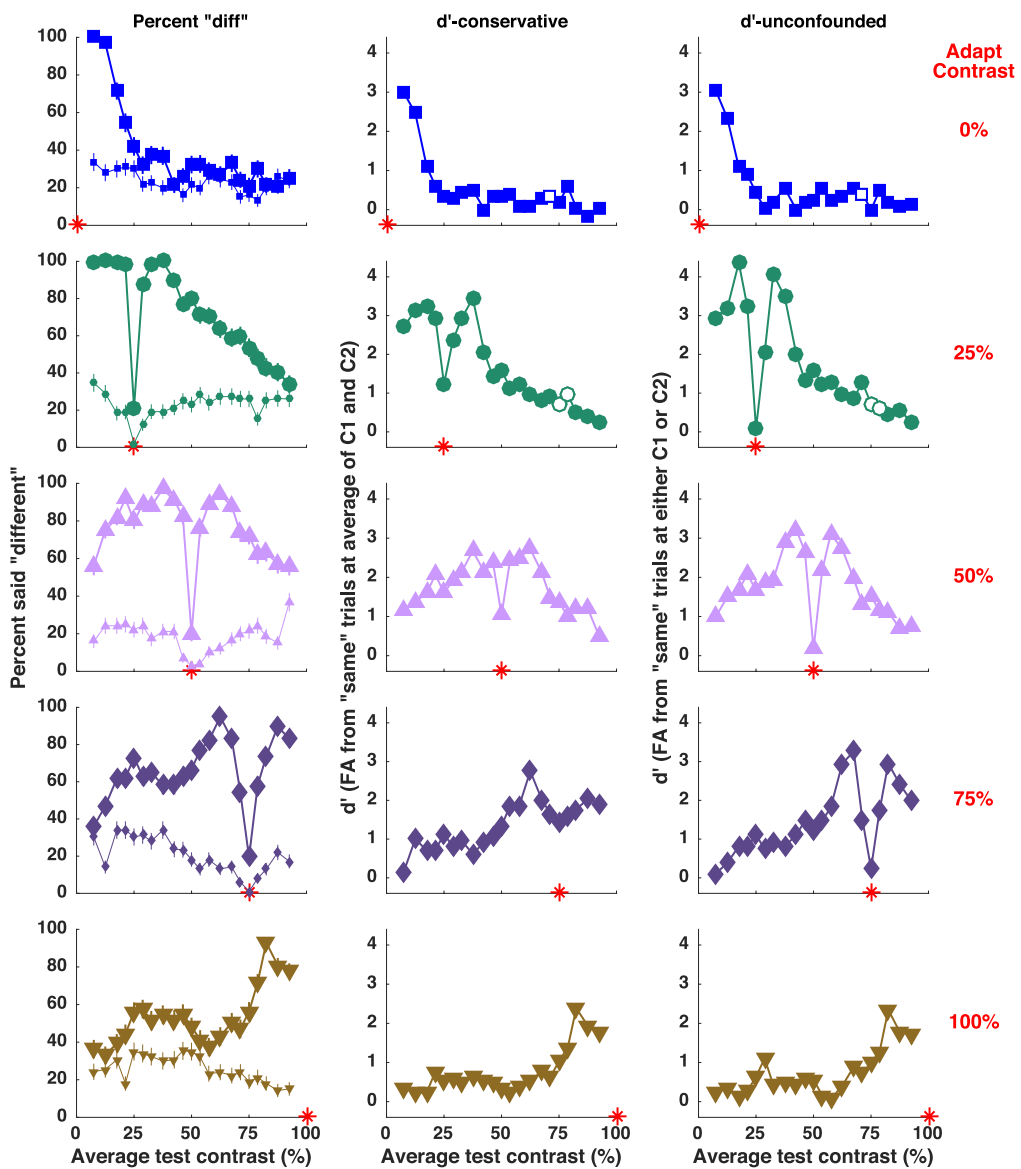
Figure C2. Results for observer LG plotted identically to those for observer MC in Figure C1. The left column of this figure was shown in Figure 12 (right column).

1ST-ORDER: SAME-DIFFERENT, TWO INTERVALS (2x2 Gabor patches)
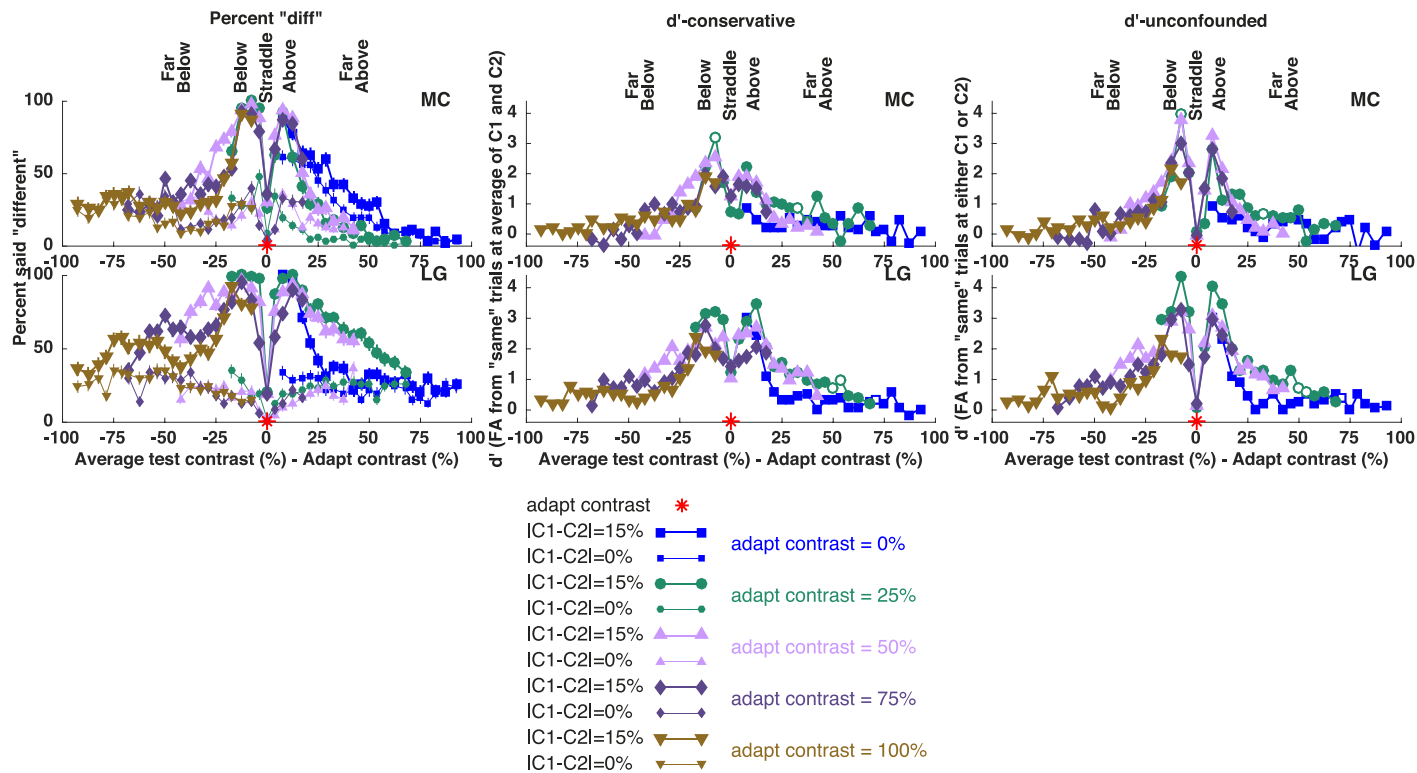
Figure C3. Curves for all five adapt contrasts plotted in Figures C1 and C2 are juxtaposed here. The three columns of the figure are for the three different measures of performance: percentage "diff," d'-conservative, and d'-unconfounded. The top row shows MC's results, and the bottom LG's. The horizontal axis gives the difference in contrast (preserving the sign) between test contrast and adapt contrast; that is, each curve is horizontally shifted so that the adapt contrast is always in the middle of the horizontal axis. The curves for different adapt contrasts on this axis superimpose, forming an empirical "butterfly curve" like that idealized in Figure 5.

Experiments using a single Gabor patch and a big disk

1ST-ORDER: SAME-DIFFERENT, TWO INTERVALS (1 Gabor patch)
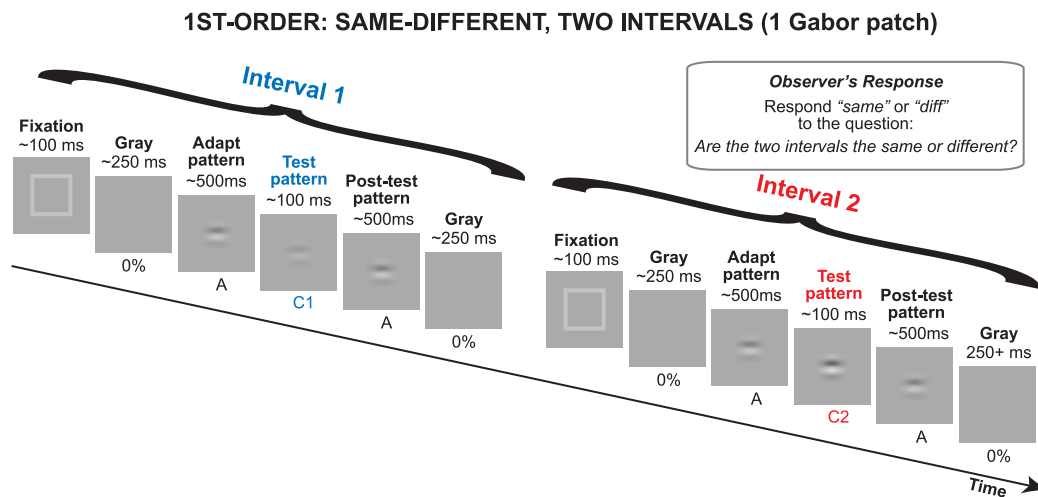
Figure C4. A diagram of a typical trial from the experiment using a single foveal Gabor patch. This trial diagram is identical to that for the 2 × 2 Gabor-patch pattern (Figure 7, top half), except that a single foveal Gabor patch replaces the 2 × 2 Gabor-patch pattern, and the fixation pattern is also changed to avoid perceptual interference with the single Gabor patch. This example shows a horizontal Gabor patch; a vertical Gabor patch was used on half the trials. Feedback was provided. (The experiment with a single foveal Gabor patch is discussed in the main text and illustrated in Figures 13 and 14.)

1ST-ORDER: SAME-DIFFERENT, TWO INTERVALS (1 Gabor patch)
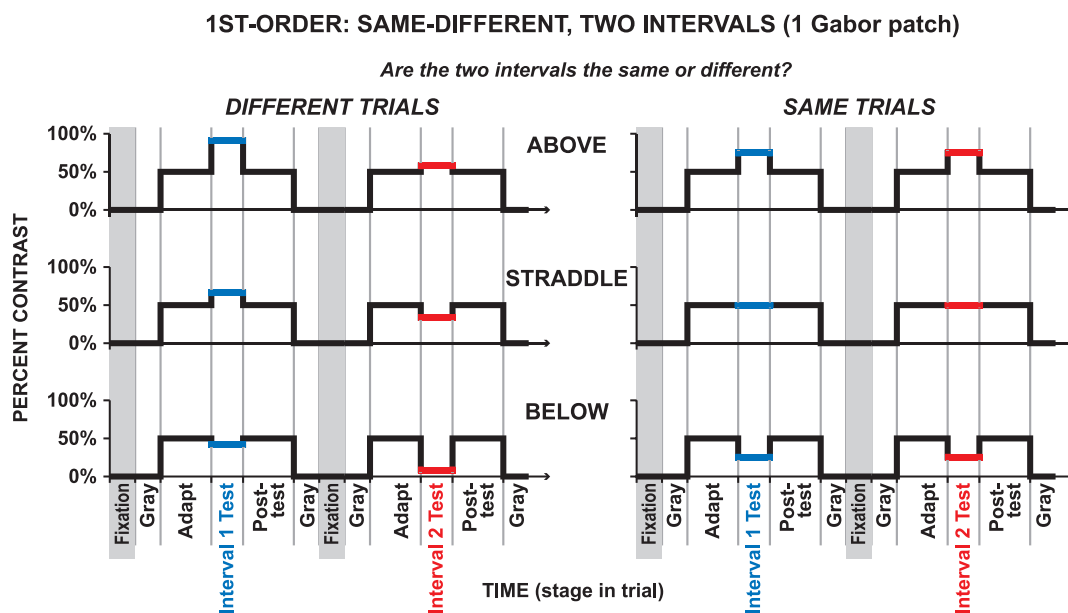
Figure C5. The contrast-versus-time profiles for the single-Gabor-patch experiment shown in Figure C4. This figure follows the conventions of Figure 8.
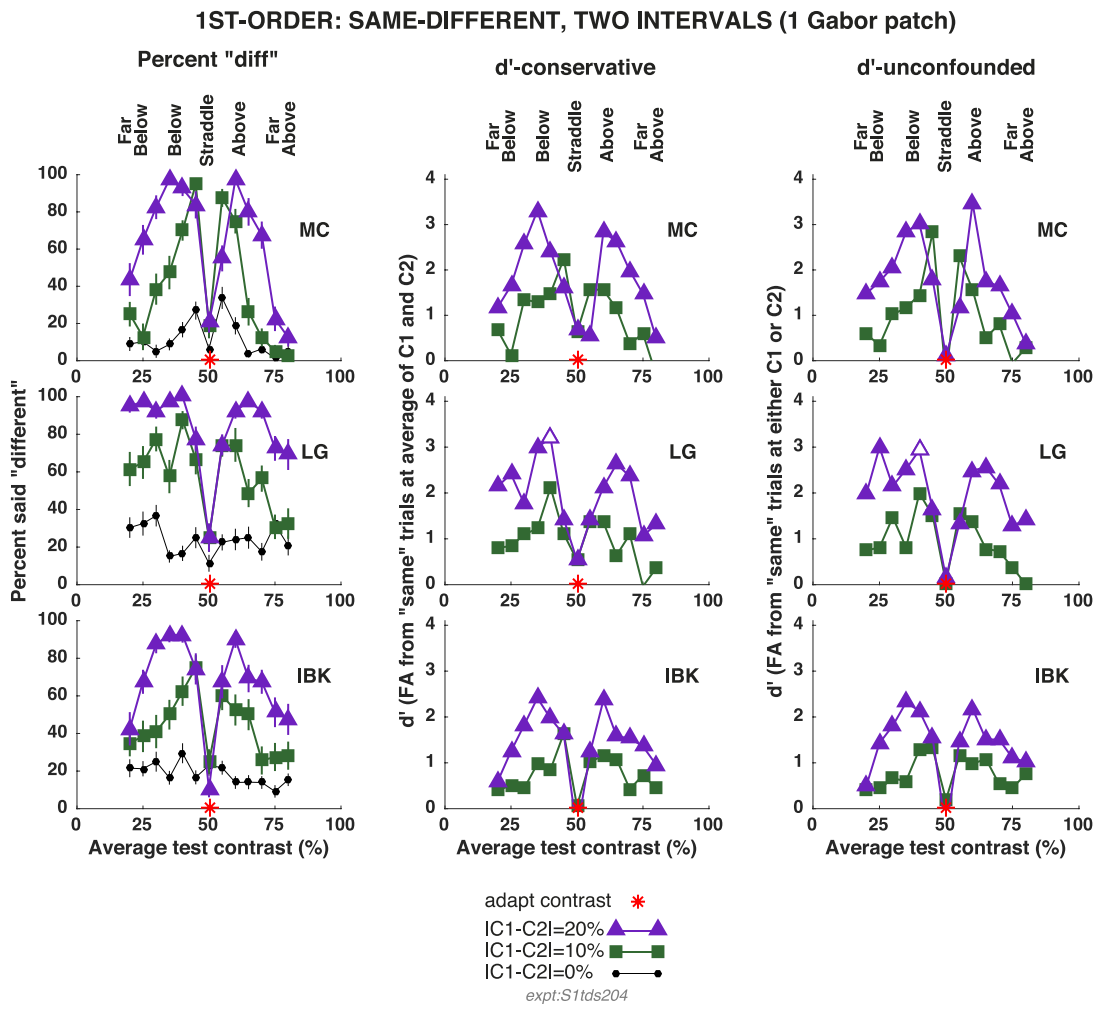
Figure C6. Results from the experiment using a single foveal Gabor patch. This figure follows the conventions of Figure 9. The left column of this figure is the same as the left column of Figure 14.

1 Gabor patch (1ST-ORDER, SAME-DIFFERENT, TWO INTERVALS) d'-unconfounded
AND
2x2 Gabor patches (1ST-ORDER, SAME-DIFFERENT, TWO INTERVALS) d'-unconfounded
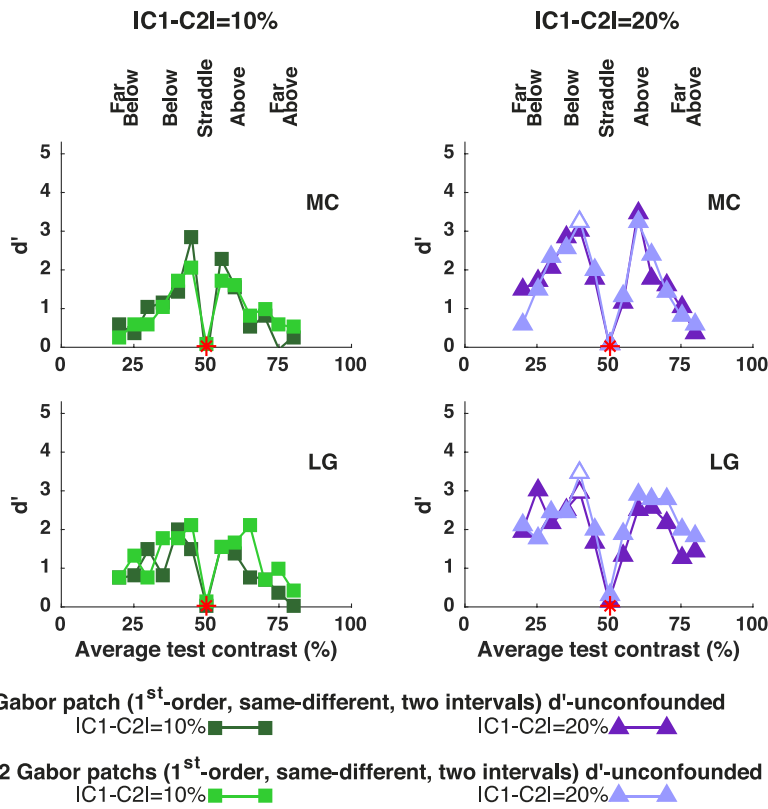
1 Gabor patch (1st-order, same-different, two intervals) d'-unconfounded
IC1-C2I=10%   IC1-C2I=20%

2x2 Gabor patchs (1st-order, same-different, two intervals) d'-unconfounded
IC1-C2I=10%   IC1-C2I=20%

Figure C7. Direct comparison of results from the same–different two-interval experiments using a single foveal Gabor patch and a 2 × 2 Gabor-patch grid. This figure has the same general format as Figures 11, 19, and 20. Like them, this figure shows only *d'*-unconfounded, but the other performance measures look similar. As shown here, the performances are very similar for the experiments with the single Gabor patch and the 2 × 2 Gabor patch. This similarity likely results from the balancing out of two sets of factors. One set favors the experiment with the single Gabor patch (e.g., greater foveal sensitivity); the other set favors the experiment with the 2 × 2 Gabor-patch (e.g., greater number of Gabor patches). The single-Gabor-patch data are also plotted in Figure C6, right column; the data for the 2 × 2 Gabor patch are also plotted in Figure 9, right column.
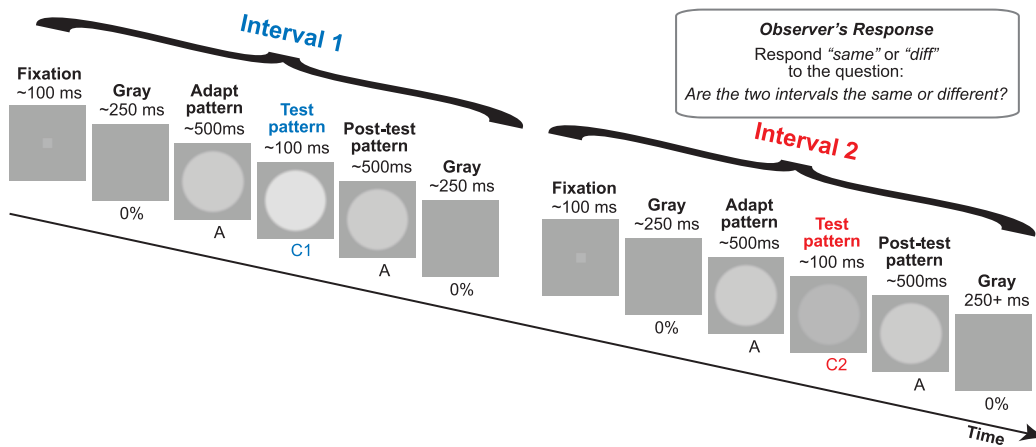

1ST-ORDER: SAME-DIFFERENT, TWO INTERVALS (1 big disk)

Figure C8. A diagram of a typical trial from the experiment using a big-disk pattern. This trial diagram is identical to that for the 2 × 2 Gabor-patch pattern (Figure 7, top half), except that a big disk replaces the 2 × 2 Gabor patch. Feedback was provided. (The big-disk experiment is discussed in the main text and illustrated in Figures 13 and 14.)

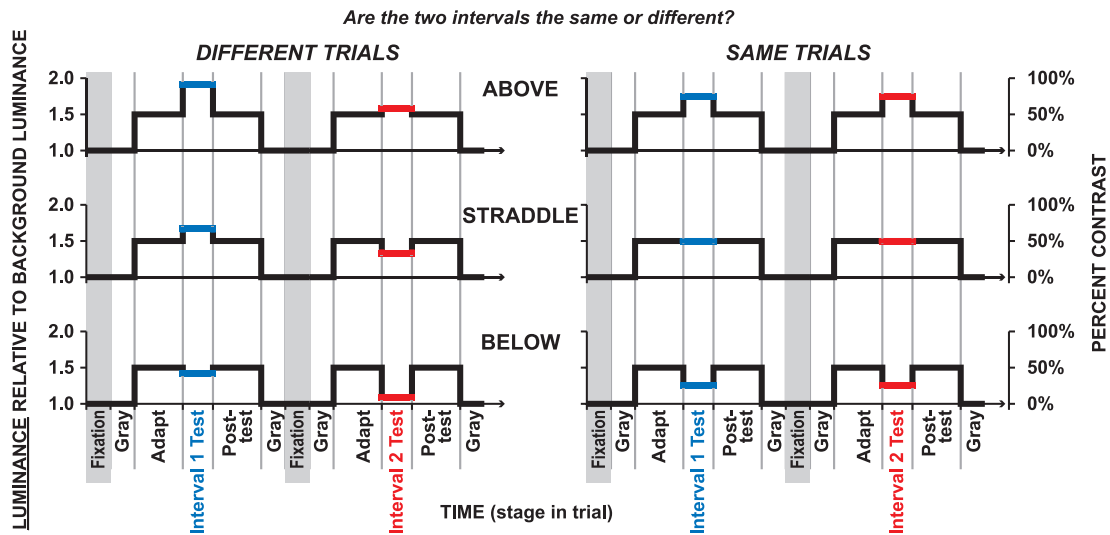## 1ST-ORDER: SAME-DIFFERENT, TWO INTERVALS (1 big disk)

Figure C9. Contrast/luminance-versus-time profiles for the big-disk experiment shown in Figure C8. Luminance in the wide central region of the disk relative to the background luminance (around the disk) is given on the left-hand vertical axis. The contrast of the big disk (shown on the right-hand vertical axis) is defined to equal the luminance in the central area of the disk minus the background luminance, divided by the background luminance. When the big disk is an adapt or posttest pattern, it is always an increment on the background. But the test pattern might be either an increment or a decrement from the adapt pattern. This figure follows the other conventions of Figure 8.
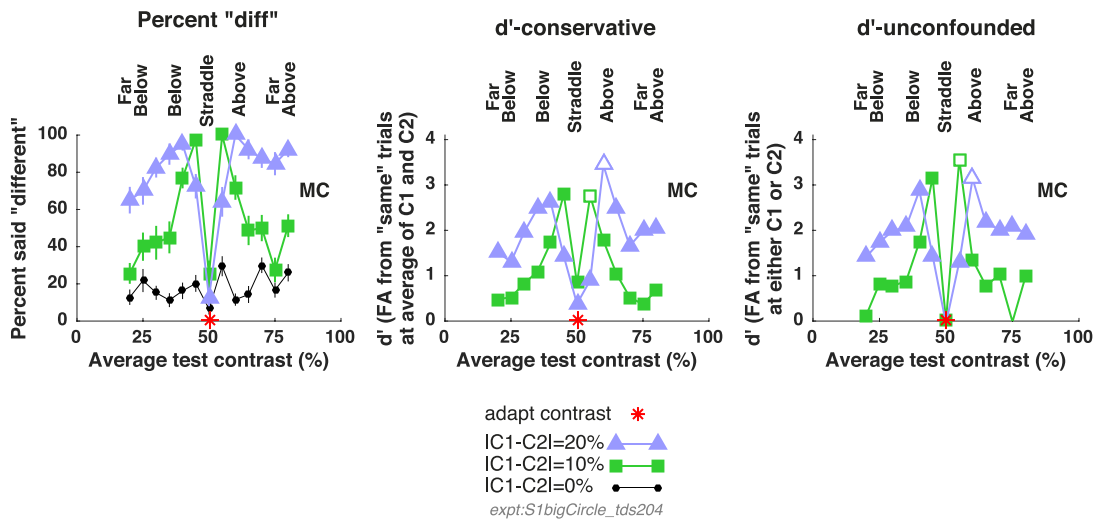

Figure C10. Results from the big-disk experiment. This figure follows the conventions of Figure 9. The left column of this figure is the same as the right column of Figure 14.
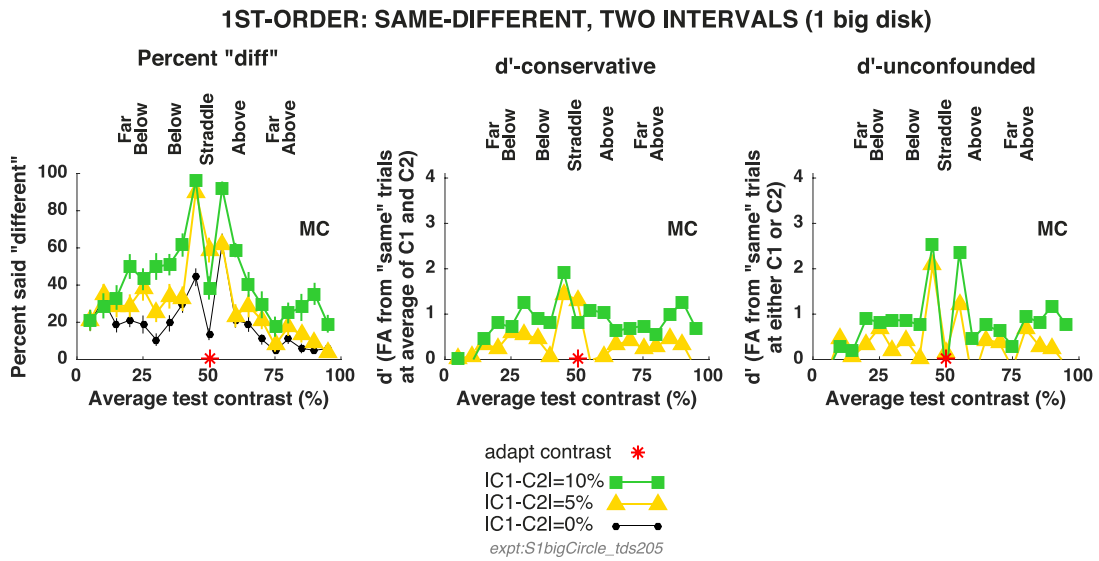
Figure C11. Results from another version of the big-disk experiment. This version used a wider range of average test contrasts and a lower test-contrast difference (5% in yellow triangles), as well as repeating 10% (green squares). This figure follows the conventions of Figure 10.