Varying test-pattern duration to explore the dynamics of contrast-comparison and contrast-normalization processes

Norma V. Graham

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

S. Sabina Wolfson

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

In this paper, we examine the dynamics of contrast-comparison and contrast-normalization processes. Observers adapted (for 1 second) to a grid of Gabor patches at one contrast; then a test pattern (which varied in duration from 12 ms to 3012 ms) was shown; and then the adapt pattern was shown again (1 second). All the Gabor patches in all the adapt patterns had 50% contrast. The test pattern was the same as the adapt pattern except that the Gabor patches in the test pattern had two different contrasts; the test contrasts varied from row to row (horizontal test pattern) or column to column (vertical test pattern). The task was to identify the orientation of the contrast variation in the test pattern (in other words, the observer performed a second-order orientation identification task). The two contrasts in each test pattern were varied while keeping the difference between the two contrasts constant. We have previously found that the observer's performance is poor for test patterns containing contrasts both above and below the adapt patterns' contrast (what we have called the "straddle effect") when the test duration is approximately 100 ms. Here, we find the straddle effect persists at all test durations we used. Other features of the results varied dramatically with test duration. We find that a simple model containing contrast-comparison and contrast-normalization processes provides a good explanation for the psychophysical results. The results provide some insight into the dynamics of these processes.

Introduction and methods

For most of their waking hours, human eyes are looking at scenes that are temporally varying. The change occurs both because the world outside is changing and because the human eyes are changing their position in that world.

Lower stages of visual processing constrain what can possibly happen at higher stages of processing. Knowledge about the functional roles of the lower stages is useful in understanding the higher stages.

In this paper, we concentrate on two intermediate-level processes: contrast comparison and contrast normalization. We have been studying their static properties (Wolfson & Graham, 2009; Graham & Wolfson, 2018). This paper is a first step in investigating their dynamic properties, which we have not previously investigated systematically. Here, we present results of varying one particular temporal characteristic of the visual pattern – namely, the duration of a test pattern – on the performance of human observers in discriminating simple patterns. In addition, we develop a simple model to account for many aspects of these results.

In this section, we first introduce the psychophysical paradigm (as briefly as possible) with the help of Figures 1, 2, and 3. Then, we summarize older experimental results, and we preview new results.

Figure 1 shows the spatial and temporal characteristics at each stage in a trial. The pattern is a 2 × 2 grid of Gabor patches. During any single trial, all the Gabor patches have the same spatial frequency and orientation and occupy the same spatial positions. The orientation of the Gabor patches (either vertical or horizontal) varies randomly from trial to trial. During the 1-second adapt pattern, all of the Gabor patches have the same contrast, 50%. The test pattern follows the adapt pattern. The Gabor patches in the test pattern are spatially identical to those in the adapt pattern, but have contrasts (T1 and T2) that vary from row to row (horizontal test pattern) or column to column (vertical test pattern). In our previous experiments, the duration of the test pattern has been approximately 100 ms (in the new experiments reported here, the duration of the test pattern varies). Following the test pattern, a pattern that is identical to the adapt pattern is shown for a second; we call this the post-test pattern.

The possible arrangements of the Gabor patches in the test pattern are shown in the box on the righthand

Citation: Graham, N. V., & Wolfson, S. S. (2023). Varying test-pattern duration to explore the dynamics of contrast-comparison and contrast-normalization processes. *Journal of Vision*, *23*(1):15, 1–22, https://doi.org/10.1167/jov.23.1.15.

https://doi.org/10.1167/jov.23.1.15 Received July 22, 2022; published January 23, 2023 ISSN 1534-7362

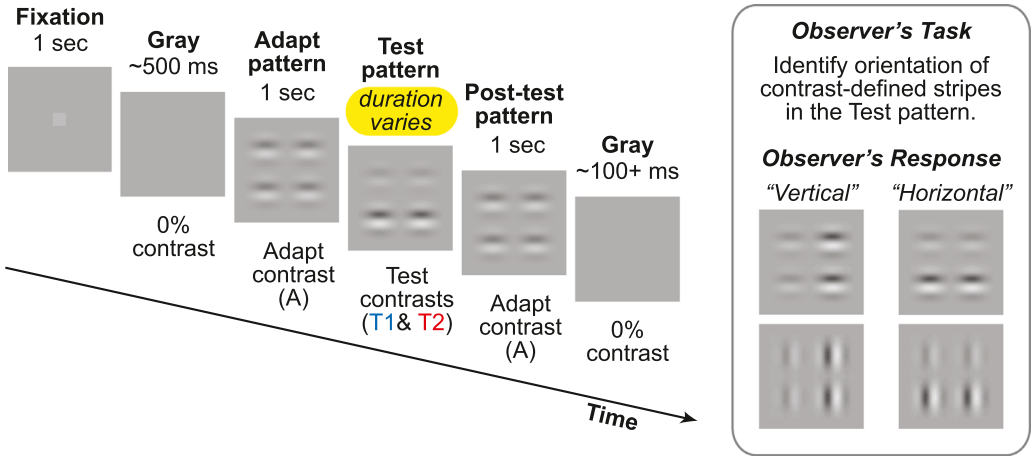

Figure 1. Diagram of a typical trial. The sinusoidal grating in our Gabor patches had a period of 0.5 degrees which corresponded to a spatial frequency of 2 c/deg. The Gaussian window had a full-width-at-half-height of 0.5 degrees. The contrast of a Gabor patch was computed by taking the difference between the luminance at the peak of the Gaussian and the mean luminance of the pattern, and then dividing that difference by the mean luminance. The task of the observer was to identify the orientation of the contrast-defined stripes in the test pattern. The words "vertical" and "horizontal" are the correct responses of the observer to the patterns shown below them.

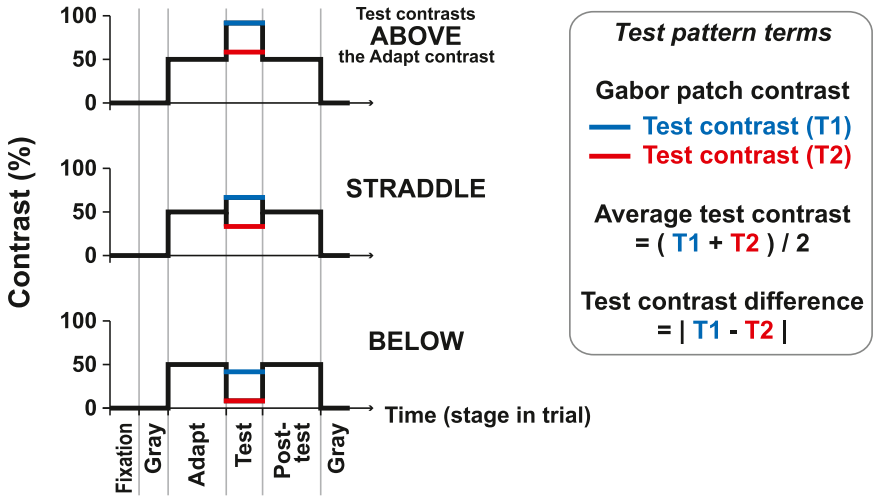

Figure 2. An illustration of contrast plotted as a function of time for three types of trial (above, straddle, and below). The light entering an observer's eyes in our experiments varies as a function of two spatial dimensions and one temporal. (It is unchanging in wavelength composition.) For our purposes, here, however, it is sufficient to use a one-dimensional function of time for each of the two test contrasts. In each panel, the function for one test contrast (T1) is plotted in black and blue, while the function for the other test contrast (T2) is plotted in black and red. The thick black lines show the single contrast when all four Gabor patches have the same contrast (50% during the adapt, 50% during the posttest, and 0% during the gray stages).

side of Figure 1. On each trial, the observer responds by indicating whether the contrast-defined orientation in the test pattern is vertical or horizontal. The labels "Vertical" and "Horizontal" above the columns give the correct responses. After the post-test pattern, the screen returns to gray (for at least 100 ms). While the screen remains gray, the observer responds. Auditory feedback as to the correctness of the response is given immediately after the response.

We have previously used the methods and procedures just briefly described. Further details of these methods and procedures are given in the description of "2nd-order orientation identification" in Appendix A of Graham and Wolfson (2018).

If one wishes to know about the dynamics of the processes underlying the psychophysical performance in these experiments, test-pattern duration is an obvious parameter to vary (as are several others which we will

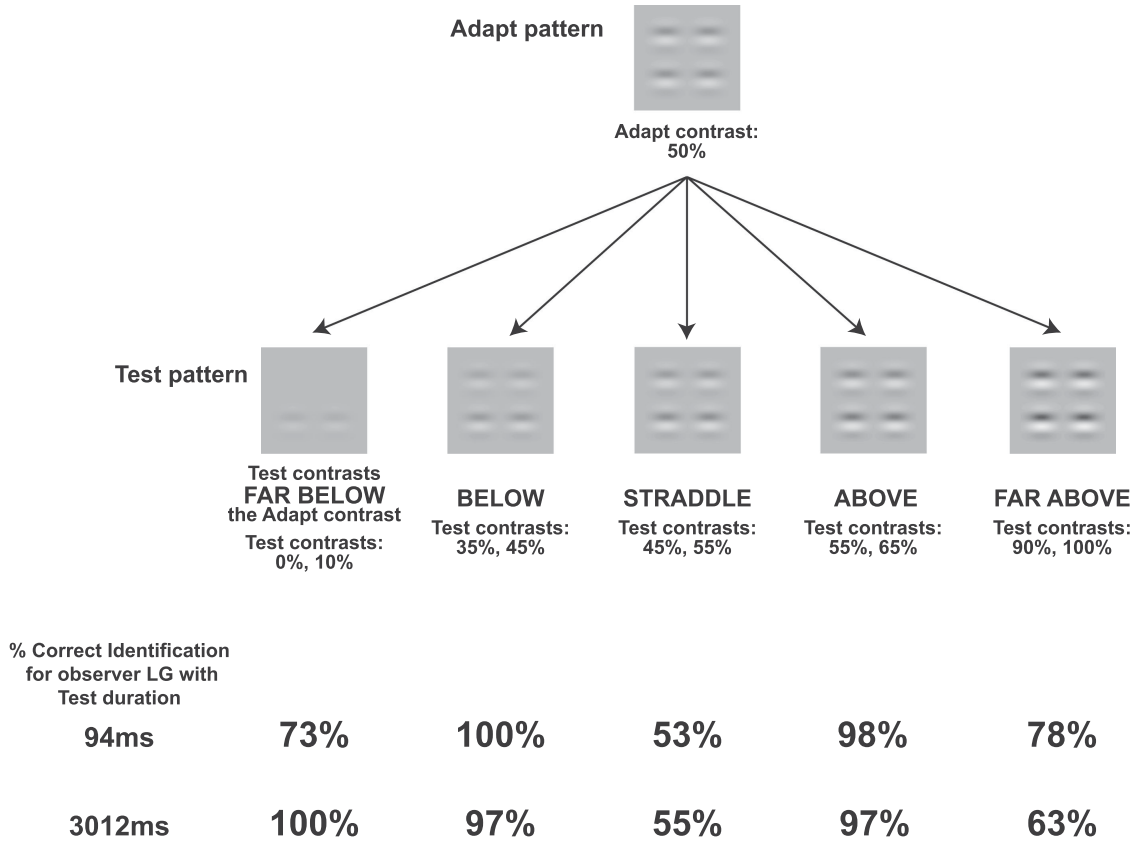

Figure 3. Illustration of the adapt and test patterns on five trials from a constant-difference series of patterns (where the test contrast difference is 10%). Results are shown for one observer with two different test durations: 94 ms (like the duration we have used in most of our prior work) and 3012 ms (a much longer duration than we have used before). Performance with the two different test durations is similar except on the far below pattern. Contrast differences in the gray-level images are exaggerated to increase their salience (the data are from observer LG in Figure 5).

visit in later papers). In the new experiments reported here, the duration of the test pattern varied from 12 ms to 3012 ms. The exact test-pattern durations are indicated in the data figures.

We generally refer to our procedure as short-term adaptation to visual contrast. It could be called masking instead (e.g. Foley, 2011). Further discussion of this terminology can be found in the Introduction section of Wolfson and Graham (2009).

Figure 2 diagrams three types of trials which we call above, straddle, and below. The names reflect the values of the test contrasts (T1 and T2) relative to the adapt contrast. The blue lines show the contrast T1, and the red lines show the contrast T2. In an above trial (top row), both T1 and T2 are higher than the adapt contrast. In a straddle trial (middle row), T1 is higher than the adapt contrast and T2 is lower than the adapt contrast. In a below trial (bottom row), both T1 and T2 are lower than the adapt contrast. The box on the righthand side defines two quantities that we often use: average test contrast and test contrast difference.

Figure 3 shows a single adapt pattern and five possible test patterns. The value of the adapt contrast is always 50%. The test-contrast difference | T1-T2 | is 10% for each test pattern, and the values of the two test contrasts vary from very low (lefthand pattern) to very high (righthand pattern). We refer to a set of patterns like this – in which the average test contrast varies but the test-contrast difference is held constant – as a constant-difference series.

The bottom of Figure 3 gives the percent correct on each of the five test patterns for two test durations for a single observer. The results for the 94 ms test duration replicate results of our previous experiments (e.g. Wolfson & Graham, 2009). The results for the 3012 ms test duration are a small subset of the new results presented later in this paper.

Performance on the middle test pattern – the Straddle pattern – is very poor at both test durations. We have called this effect the straddle effect. (Foley, 2011, has also shown the straddle effect with a test duration of 100 ms.)

For both test durations in Figure 3, performance peaks on the two patterns adjacent to the middle – the below and above patterns.

At a test duration of 94 ms (like that used in our previous work), performance at both ends – on far-below and far-above patterns – drops down from the peak performance.

At a long test duration of 3012 ms, performance on the righthand (far-above) pattern was again substantially lower than peak performance. However, performance on the lefthand (far-below) pattern is as high as at the peak! This was originally very surprising to us.

The curve in Figure 4 is a sketch of a typical observer's performance on a constant-difference series of patterns at a test duration like 94 ms.

To understand the results presented in this paper, for a variety of test durations we developed a dynamic, functional model starting from the static modeling we had already done for our previous results. This modeling is based on two processes: a contrast-comparison process and a contrast-normalization process. The labels in Figure 4 indicate where on this curve each of these two processes is dominant according to the static model we have previously developed. A description of this static model is given in Wolfson and Graham (2009).

A brief verbal description of the two processes follows.

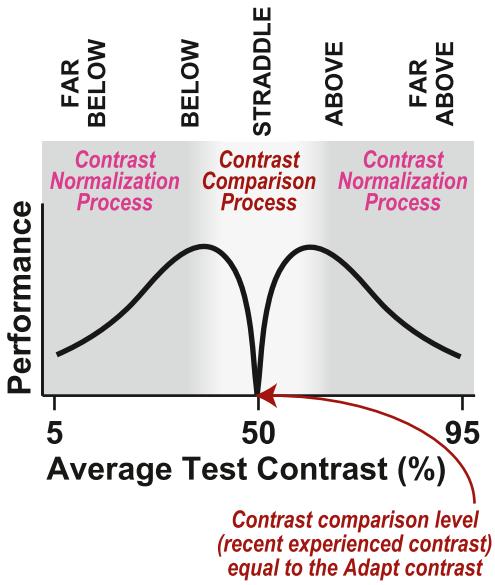

Figure 4. Sketch of a typical observer's performance on a constant-difference series of patterns with a test duration of about 100 ms. The labels "Contrast Normalization Process" and "Contrast Comparison Process" indicate where on this plotted curve each of two processes in our model is dominant. See text for description of the model and these two processes.

The contrast-comparison process. We have explained the straddle effect previously – and will continue to do so here (both verbally and then formally in a model) – in terms of a contrast-comparison process (Wolfson & Graham, 2007). In more detail, the test-pattern contrast at each point in space is compared to a contrast-comparison level, which is the weighted average of recent contrasts at that point. The comparison is a computation of the magnitude – neglecting the sign – of the difference between the current contrast and recently-weighted average contrast. When the result of that comparison is large, the observer is easily able to see the pattern and thus to perform well. Performance is poor when the result of that comparison is small, as can happen when both current contrasts are far from the recently-weight average contrast but on opposite sides of it. (In Foley's 2011 model, there is a V-response that acts much like the contrast-comparison process we use.)

The contrast-normalization process. For test durations of about 100 ms: the observer's performance reaches a peak on both sides of the local minimum at the straddle pattern; and then it declines to very low values for far-below and far-above test patterns. We have attributed these declines to a contrast-normalization process which is a form of divisive inhibition that has been suggested in many other situations: in physiology by, for example, Heeger (1992), Carandini and Heeger (2012), Solomon and Kohn (2014), and Sawada and Petrov (2017); and in psychophysics by, for example, Legge and Foley (1980), Teo and Heeger (1994), Watson and Solomon (1997), Foley (1994), Graham and Sutter (2000), and Shooner and Mullen (2020).

In our model presented in this paper, the dynamics of these processes are made explicit. For the contrast-comparison process, we postulate transient pulses of excitation occurring at both increases and decreases of contrast. For the contrast-normalization process, we found ourselves forced to postulate two kinds of divisive inhibition, one that is transient and one that is sustained.

Experimental results

Figures 5 and 6 show results from our experiments varying the test duration for four human observers (M.C., L.G., W.L., and B.S.G.). Different observers' results appear in different columns. The exact values of the test durations are given at the right of the figures. Each panel shows the results for two constant-difference series, one for a 10% difference (green open squares) and the other for a 5% difference (purple solid discs).

The curves for the 10% and the 5% constant-difference series are very similar except that the observer's performance on the 10% series is generally

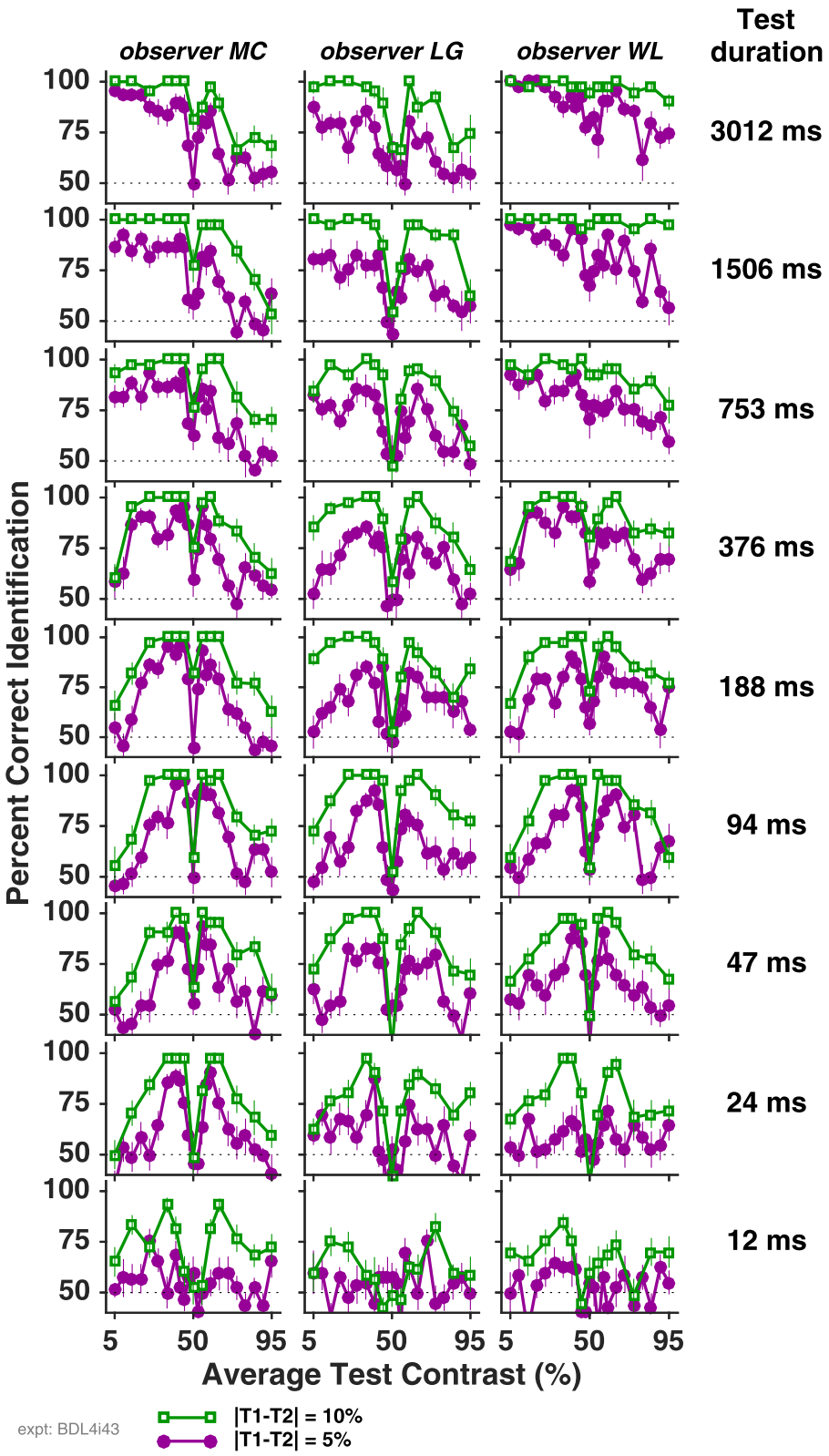

Figure 5. Results from observers M.C., L.G., and W.L. (in different columns) at test durations ranging from 12 ms (bottom row) to 3012 ms (top row). Error bars show +/- 1 standard error across blocks. The green open squares show results from the 10% constant-difference series. The purple solid discs show results from the 5% constant-difference series.

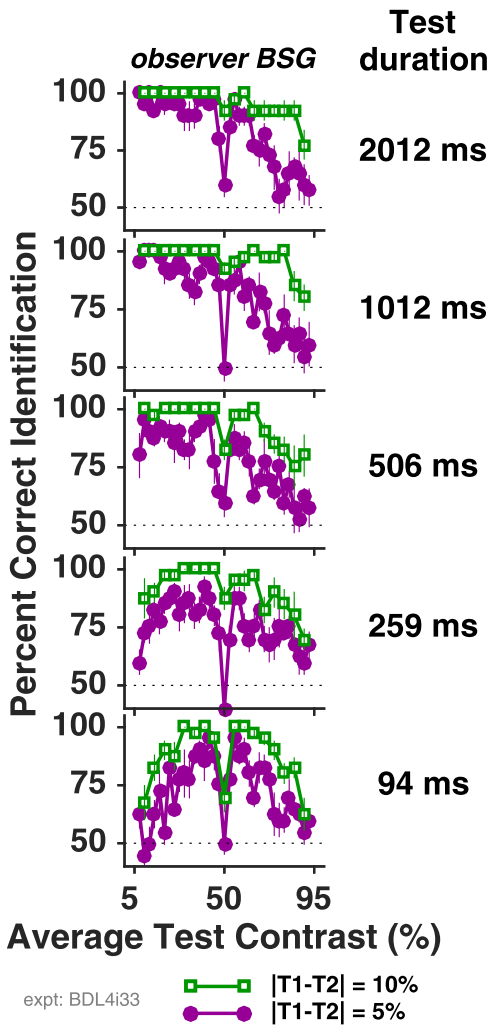

Figure 6. Results from observer BSG at test durations ranging from 94 ms (bottom row) to 2012 ms (top row). Symbols have the same meaning as in Figure 5.

higher than that on the 5% series. This is not surprising because a 10% difference in test contrasts should make processing the information substantially easier than a 5% difference would.

In each block of trials in the experiment, the test duration was the same on all trials. But the order of blocks of different durations was random. (Blocks in Figure 5 were 144 trials long, and observers ran 10 or 11 blocks at each test duration. Blocks in Figure 6 were 416 trials long, and the observer ran five blocks at each test duration.)

Effect of varying test duration

What happens as the test duration is varied? An example of the results for a constant-difference series (at just 5 points) at two durations was already presented in Figure 3. In Figure 7, the complete shapes of the constant-difference-series curves at two test durations are idealized. We will now go through the full results in some detail.

Results for test duration of 94 ms

The results at a test duration of 94 ms in Figures 5 and 6 show a shape like that idealized in the left-hand panel of Figure 7. Namely, performance generally dips to a very low value for the Straddle test pattern (where the average test contrast equals the adapt contrast 50%). Performance rises to peaks for average test contrasts somewhat below and somewhat above the adapt contrast. Performance declines again for far-below test patterns (left end) and far-above test patterns (right end).

As an aside, in this paper, the adapt contrast is always 50%. Previously, we have used adapt contrasts varying from very low to very high (e.g. Wolfson & Graham, 2009) with a test duration of about 100 ms. Regardless of the adapt contrast, the observers' performance always reached a very low value when the average test contrast equaled the adapt contrast. In other words, the dip at the straddle test pattern was not always at 50%, but depended on the adapt contrast.

Results for test durations increasing from 12 ms to 94 ms

Consider what happens as test duration increases from very short (12 ms) to a duration of 94 ms in Figure 5. Performance on the straddle test pattern remains around chance (50% correct); performance on the test patterns with average test contrast very far below (or above) also remains around chance; but performance on the other patterns improves greatly. Looked at from a slightly different perspective, the whole constant-difference-series curve tends to rise, but the rise is greatest at the peaks and least at the center (the straddle pattern) and two ends (the far-below patterns, and far-above patterns). This is clearer for the results with a 5% constant-difference series than with a 10% series, because the results in the 10% series are more often at ceiling.

The kind of general increase in performance from very short to longer test durations is what one might well expect from almost any dynamic process that integrates incoming stimulation for some time period (i.e. accumulates information over some time period). When the test duration is too short, there is not very much information for the brain to use in doing the task. As the test duration gets longer, more and more information is available and performance gets better.

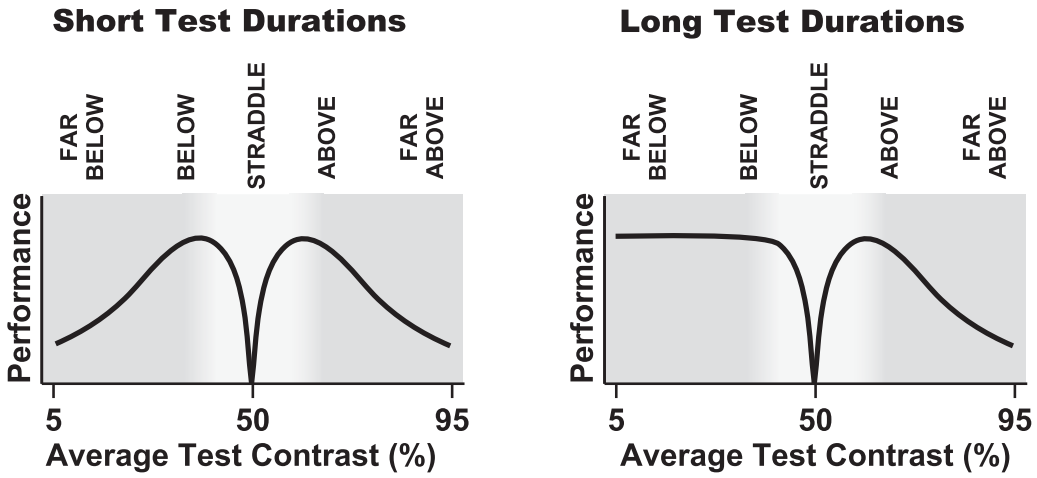

Figure 7. An idealized version of the experimental results shown in Figures 5 and 6. At short test durations, the results generally have the shape shown in the left panel. At the longest test durations, the results generally have the shape shown in the right panel.

Results for test durations increasing from 94 ms to 3012 ms

There is a dramatic change in the shape of the results curves as the test duration continues to increase from 94 ms to 2012 ms (Figure 6) or to 3012 ms (Figure 5). The performance continues to get better and better, but only for the lower average-test-contrast end of the curve (the left end), not for the higher average-test-contrast end of the curve (the right end). This is idealized in Figure 7.

In general, there is little improvement in the performance on the straddle test pattern even at the longer test durations. This initially surprised us. At an intuitive level, we expected observers to overcome their difficulty seeing the straddle test pattern if the test pattern was on for a second or more. There are individual differences; in particular, observer W.L. shows more improvement at longer test durations.

Informal explanation of the long-test-duration results

Let us get back to the question of why the left end of the constant-difference series curve is elevated at long test durations but not at short test durations. At an informal, verbal level, the answer is as follows.

If the test duration is long and the test pattern is composed of low test contrasts (e.g. the image of the test pattern on the left of Figure 3), the observer is seeing low contrasts for a relatively long time. Therefore, it is very much as if the system has been staring at a gray blank field for quite a while before the test offset. When the test offset occurs, the system might be in a heightened state of sensitivity relative to its state at the test offset of a high average-test-contrast pattern (like that in the image of the test pattern on the right end of Figure 3). Thus, the system might be able to respond to the low average contrast test pattern because it is able to "see" the test offset even if it could not "see" the test onset.

Thus, when test duration is long, performance on the left end of the constant-difference-series curve might be substantially better than performance on the right end because it is as if the observer has adapted to something like a gray blank field during the long test pattern. That is, a constant-difference-series curve for a long test duration will have higher performance on the left end than on the right end, as shown in Figure 7 in the panel labeled "Long Test Durations."

In the next section, we introduce a simple model that will account for many of these results in terms of the dynamics of three mechanisms: a transient excitatory mechanism (produced by contrast comparison) and two inhibitory mechanisms – one sustained and one transient – which carry out contrast normalization.

Modeling

The simple model developed here is not intended to describe in any detail the brain's processing of visual patterns. It is an attempt to characterize and give insight into functional dynamics of intermediate-level human vision. We are not using fine-grained neural modeling of the sort used by, for example, Heeger and Zemilianova (2020) or Chariker, Shapley, and Young (2020).

How the model responds to contrast as a function of time (the input side of the model)

This model breaks naturally into two parts. The first part is shown in Figure 8. It is the input side of the model, going from Gabor-patch contrast to the model response $R(t)$. The second part is the output side of the model and will be discussed in "How the model predicts the observer's performance."

Figure 8 shows the three key mechanisms: one excitatory and two inhibitory. The two key inhibitory mechanisms are divisive. All three mechanisms are described further below. We give equations for all three mechanisms in Figure 9 and show time courses of their action in Figure 10.

Recall that the patterns used in our experiments were collections of Gabor patches, where the Gabor patches had one of two possible contrasts (called T1 and T2) during the test interval. The contrast of one Gabor patch as a function of time will be given in symbols as $contrast(t)$. An example of this function for one Gabor patch is shown in the top row of Figure 10 with a test duration of 1506 ms. (The same example is shown twice, once in the left column for the numerator computation and once in the right column for the denominator computation.) $Contrast(t)$ is shown toward the left of Figure 8.

Key excitatory mechanism

The key excitatory mechanism is shown in Figure 8 as an oval labeled "transient excitation." Critically for the explanation of our psychophysical data, this excitatory mechanism shows positive transient output pulses to the test onset and to the test offset. An example of this is shown in the left column of Figure 10, fourth row from the top. For a test duration of 1506 ms as shown here, the key excitatory mechanism is fast enough that its output pulses at the onset and offset of the test pattern are separate and of identical height. The onset and offset pulses are separate for all test durations of 94 ms or longer. (Illustrations for a test duration of 94 ms are shown in the Appendix). For shorter test durations (12, 24, and 47 ms), output pulses to the onset and offset of the test pattern interact, and the first pulse is substantially higher than the second pulse. The exact shape and timing of the key excitatory mechanism's output does not matter very much for the conclusions of this paper, as verified by subsidiary calculations we have done.

In symbols, we will refer to the output of the key excitatory mechanism as:

$$E_{tran}(t)$$

The subscript *tran* is short for transient. Similarly, the subscript *sus* will be short for sustained.

Details about the key excitatory mechanism for the interested reader

The equations that produce $E_{tran}(t)$ are given in Figure 9. $E_{tran}(t)$ is computed as the rectified difference between two sustained excitatory mechanisms, one of which is slightly slower than the other. The left column, second row of Figure 10 shows the outputs of the two sustained excitatory mechanisms: The faster one is shown by the darker brown line $O(t)$ labeled "measure of current local contrast"; the slower one is shown by the orange line $Z(t)$ labeled "current contrast comparison level." Both are linear with contrast in the pattern.

Formally, both $O(t)$ and $Z(t)$ are the result of six stages of exponential filtering where the time constant for the faster mechanism $O(t)$ was 5 ms and that for the slower mechanism $Z(t)$ was 10 ms. Our reason for using exponential filtering here was its convenience

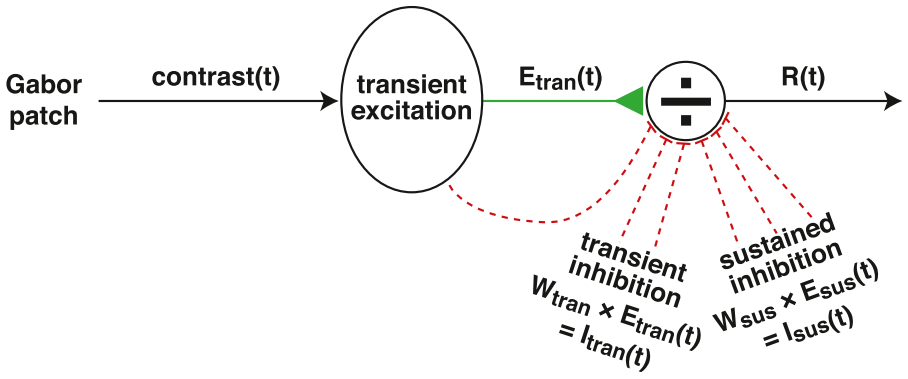

Figure 8. Flow diagram of the input side of the model. The dashed red lines are inhibitory connections, and the solid green line is an excitatory connection. There is one key excitatory mechanism. This mechanism responds positively and briefly to step increases and step decreases in contrast. This positive-going transient output results from the contrast-comparison process. There are two key inhibitory mechanisms – one sustained and one transient. Together they perform contrast normalization. The weights W_{tran} and W_{sus} can represent the number of connections or the strength of connections. The output of this flow diagram is the input to Figure 11.

$R(t)$ is the model's response as a function of time to a Gabor patch.

$$R(t) = \frac{E_{tran}(t)}{\sigma + I_{tran}(t) + I_{sus}(t)}$$

The contrast of the Gabor patch as a function of time is denoted by the symbol $contrast(t)$.

$Numerator(t) = E_{tran}(t) = |O(t) - Z(t)|$

where $O(t)$ is the output to $contrast(t)$ of a low-pass filter,
and $Z(t)$ is the output to $contrast(t)$ of a slightly slower low-pass filter.

$Denominator(t) = \sigma + I_{tran}(t) + I_{sus}(t)$

$I_{tran}(t) = W_{tran} \times E_{tran}(t)$ where $W_{tran}=1$ and $\sigma=5$

$I_{sus}(t) = W_{sus} \times E_{sus}(t)$
where $E_{sus}(t)$ is the output to $contrast(t)$ of a very slow low-pass filter,
and where $W_{sus}= 0$ **or** 1

$R1(t)$ and $R2(t)$ are the model's responses as a function of time to two Gabor patches; one of these Gabor patches has contrast $T1$ during the test interval, and the other has contrast $T2$ during the test interval. The pooled visual response, $V(T1,T2)$, is computed by the following equation:

$$V(T1,T2) = \left[\sum_t |(R1(t) - R2(t))|^k\right]^{1/k} \quad \text{where } k=2$$

Figure 9. Equations for the model developed in this paper. The top box refers to the input side (shown in Figure 8 and discussed in "How the model responds to contrast as a function of time"). The bottom box refers to the output side (shown in Figure 11 and discussed in "How the model predicts the observer's performance"). The period of time over which the summation *V(T1,T2)* occurs depends on whether the observer is using the WHOLE, ONSET, or OFFSET decision rule.

in producing sustained (low-pass) responses that we could vary the time course of. For more details of the exponential filtering done here, see Graham and Hood (1992) in the appendix section titled "Linear lowpass module - Cascaded exponential filters" on page 1392. The time step we used in these calculations was 0.005 ms.

The difference, $O(t) - Z(t)$, is shown in the left column, third row of Figure 10. It shows a positive-going pulse whenever there is a contrast increase, and a negative-going pulse whenever there is a contrast decrease.

That difference is then full-wave rectified – in other words the absolute value of the difference is computed – to become $|O(t) - Z(t)|$. It shows positive-going pulses to both contrast increases and contrast decreases, as in the left column, fourth row, of Figure 10.

(The V-response of Foley's 2011 model embodies a full-wave rectification that performs a similar function to that performed by the full-wave rectification here.)

For the results and predictions considered in this paper, full-wave rectification is good enough. But a quantitative explanation of some other aspects of the results (at least for some observers) requires partial rather than full rectification. For example, with only full-wave rectification, the observer's performance on a straddle test pattern, no matter how large the difference between the two test contrasts, would always be at chance. For further discussion about the need for partial rectification see Motoyoshi and Kingdom (2007), and (Graham & Wolfson, 2007; Graham & Wolfson, 2013). Although we could allow for partial rectifications for our observers here, it would require two further parameters (for each observer) without

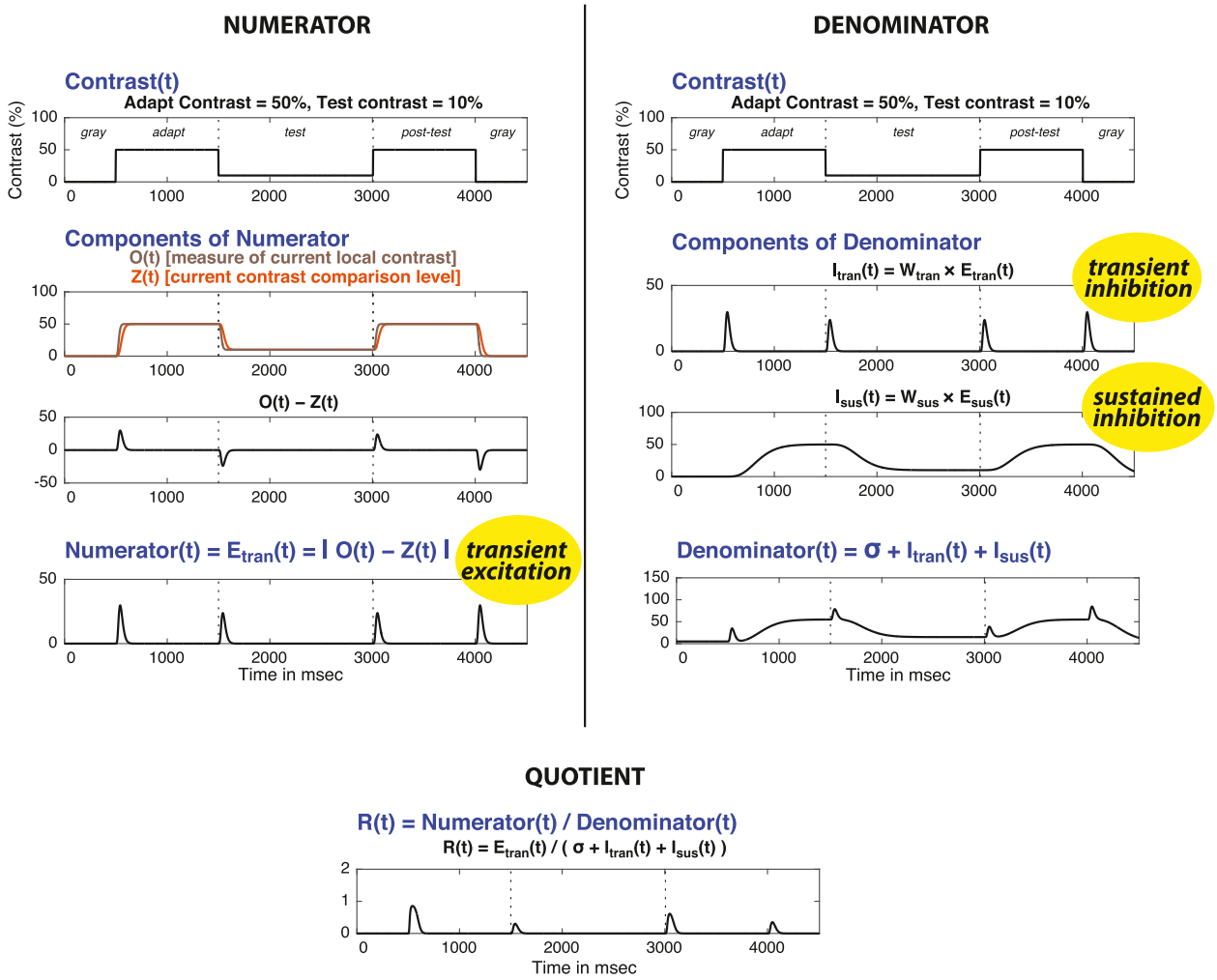

Figure 10. Responses of different stages in our model's input side as a function of time. The model is responding to a single Gabor patch having a time-course specified by *contrast(t)*. This particular Gabor patch is of very low test contrast (10%). The adapt contrast is 50%, which is the adapt contrast used in all the experiments reported in this paper. The test duration is 1506 ms. The left column shows the functions of time that produce the numerator. The right column shows the functions that produce the denominator. The centered row at the bottom is the response R(t), which equals the moment-by-moment division of the numerator by the denominator. More illustrations for a test duration of 1506 ms are given in the Appendix. In addition, illustrations for a test duration of 94 ms are given in the Appendix.

gaining anything about the aspects of the results we are interested in.

Two key inhibitory mechanisms

In our model, there are two key inhibitory mechanisms. They are labeled "transient inhibition" and "sustained inhibition" in Figure 8.

The key transient inhibitory mechanism is shown in the right column of Figure 10, second row. The exact time-course does not matter much for our conclusions here. We assume it is a weighted version of the key excitatory mechanism's time-course (see Figure 10, left column, fourth row). For this transient key inhibitory mechanism, we use the symbol:

$$I_{tran}(t)$$

The second key inhibitory mechanism is sustained and acts over a substantially longer period of time. The exact time-course for the sustained key inhibitory mechanism does not matter much. However, the approximate time course is important. For a sketch of the exact time course, see Figure 10, right column, third row. For this sustained key inhibitory mechanism, we use the symbol:

$$I_{sus}(t)$$

The transient inhibitory mechanism's response has a time course which is a weighted version of the key

excitatory response $E_{tran}(t)$ with the weight W_{tran}. That is:

$$I_{tran}(t) = W_{tran} \times E_{tran}(t)$$

The sustained inhibitory mechanism's response is:

$$I_{sus}(t) = W_{sus} \times E_{sus}(t)$$

where $E_{sus}(t)$ is the result of six stages of exponential filtering with a time constant of 60 ms.

In addition to the two key inhibitory mechanisms, there is a constant baseline inhibition represented by the parameter σ.

The existence of sustained and transient mechanisms has been suggested by many investigators to explain both psychophysical and physiological results. Early suggestions include Kulikowski and Tolhurst (1973) in psychophysics and Cleland, Dubin, and Levick (1971) in physiology. Relying on the physiological correlates, more recent suggestions about psychophysics (e.g. Pokorny & Smith, 1997; Shooner & Mullen, 2020) refer to the sustained and transient mechanisms as the parvocellular (PC) and magnocellular (MC) mechanisms.

The two key inhibitory mechanisms are assumed to act together as a divisive inhibition. The divisive inhibition will be very simply modeled here as a moment-by-moment division. That is, the key excitatory mechanism's output is divided by the total amount of inhibition (the sum of the constant inhibition σ plus the two key inhibitory mechanisms' outputs). The exact value of the parameter σ (in concert with the weights W_{sus} and W_{tran}) sets the overall scale of the inhibition. In addition, any value of σ greater than zero guarantees that the denominator – the total amount of inhibition – can never reach zero. The exact value of σ matters rather little for the predictions (although the necessary order-of-magnitude for σ depends heavily on the weights used for the excitatory and inhibitory mechanisms).

Response R(t) of the model's input side

We use $R(t)$ as the symbol for the response of the model's input side as a function of time to a Gabor patch:

$$R(t) = \frac{E_{tran}(t)}{\sigma + I_{tran}(t) + I_{sus}(t)}$$

$R(t)$ is used for the general case. $R1(t)$ and $R2(t)$ are used to refer explicitly to the two contrasts in a test pattern. $R1(t)$ is the response to a Gabor patch having contrast T1 during the test interval. $R2(t)$ is the response to a Gabor patch having contrast T2 during the test interval.

Crucially, as will be shown in the Discussion section, the model's ability to account for our experimental results depends on the value of the weight W_{sus}. We found a clear distinction between two types of behavior: (1) behavior that could be predicted by very small values of W_{sus} (relative to W_{tran}) and (2) behavior that could be predicted by substantially larger values of W_{sus} (approximately equal to W_{tran}).

How the model predicts the observer's performance (the output side of the model)

The previous subsection described the three key mechanisms (transient excitation, transient inhibition, and sustained inhibition) that work together to produce the responses $R1(t)$ and $R2(t)$. In line with the general modeling philosophy of this paper, we use a very simplified two-step calculation to represent all the higher-level processing occurring between the

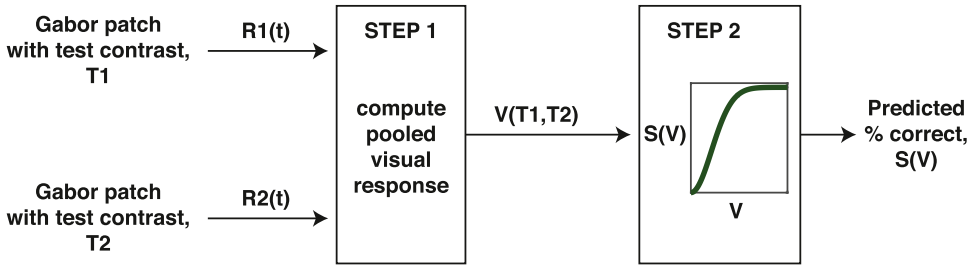

Figure 11. Flow diagram of the output side of the model. There are two inputs – *R1(t)* and *R2(t)* – shown in this flow diagram. *R1(t)* and *R2(t)* are responses to two different Gabor patches, each of which goes through the stages shown in the diagram of Figure 8. Those two Gabor patches are identical in spatial characteristics but have different contrasts (*T1* and *T2*). The value of *V(T1,T2)* will depend on the period of time over which the summation to compute *V(T1,T2)* occurs, i.e. whether the observer is using the WHOLE, ONSET, or OFFSET decision rule. The computed value of *V(T1,T2)* then goes through a monotonic S-function to produce a prediction of the observer's percent correct.

two responses, $R1(t)$ and $R2(t)$, and the observer's response. These two steps are shown in Figure 11.

STEP 1 - Compute pooled visual response

We define an intermediate quantity – the *pooled visual response* V – which is a single number for each pattern in the experiment. $V(T1,T2)$ is a measure of the overall difference between $R1(t)$ and $R2(t)$ pooled across time. In symbols,

$$V(T1, T2) = \left[\sum_t |\, R1(t) - R2(t)|^k\right]^{1/k}$$

where $T1$ and $T2$ are the values of the two test contrasts that led to the responses $R1(t)$ and $R2(t)$. The sum is taken over many discrete time points during the trial; the particular time points depend on the decision rule as explained below. In this paper, we will let the exponent $k = 2$. Although we have tried other values for k, the exact value makes little if any difference for the conclusions here.

The kind of computation in the equation for $V(T1,T2)$ has been called various names in different contexts, for example, the Minkowkski metric, power summation, and Quick pooling.

If all times t from the beginning to the end of the trial are summed over, we call this the *pooled visual response V for the whole trial* or, in symbols, $V_{WHOLE}(T1,T2)$. If an observer uses this quantity to make their decision on each trial in our experiment, we say the observer is using the WHOLE decision rule.

If all times from the beginning of the trial to 1 ms before the offset of the test pattern are summed over, we call it the *onset pooled visual response V* or, in symbols, $V_{ONSET}(T1,T2)$. This quantity only includes the response to the test-pattern onset; logically it does NOT include the response to the offset because the offset has not happened yet. If an observer uses this quantity to make their decision on each trial in our experiment, we say the observer is using the ONSET decision rule.

If all times from a moment after the offset of the test pattern to the end of the trial are summed over, we call it the *offset pooled visual response V* or, in symbols, $V_{OFFSET}(T1,T2)$. This quantity includes the response to the test-pattern offset, and generally it does NOT include any of the response to the onset. (For very short test durations, however, this quantity does in

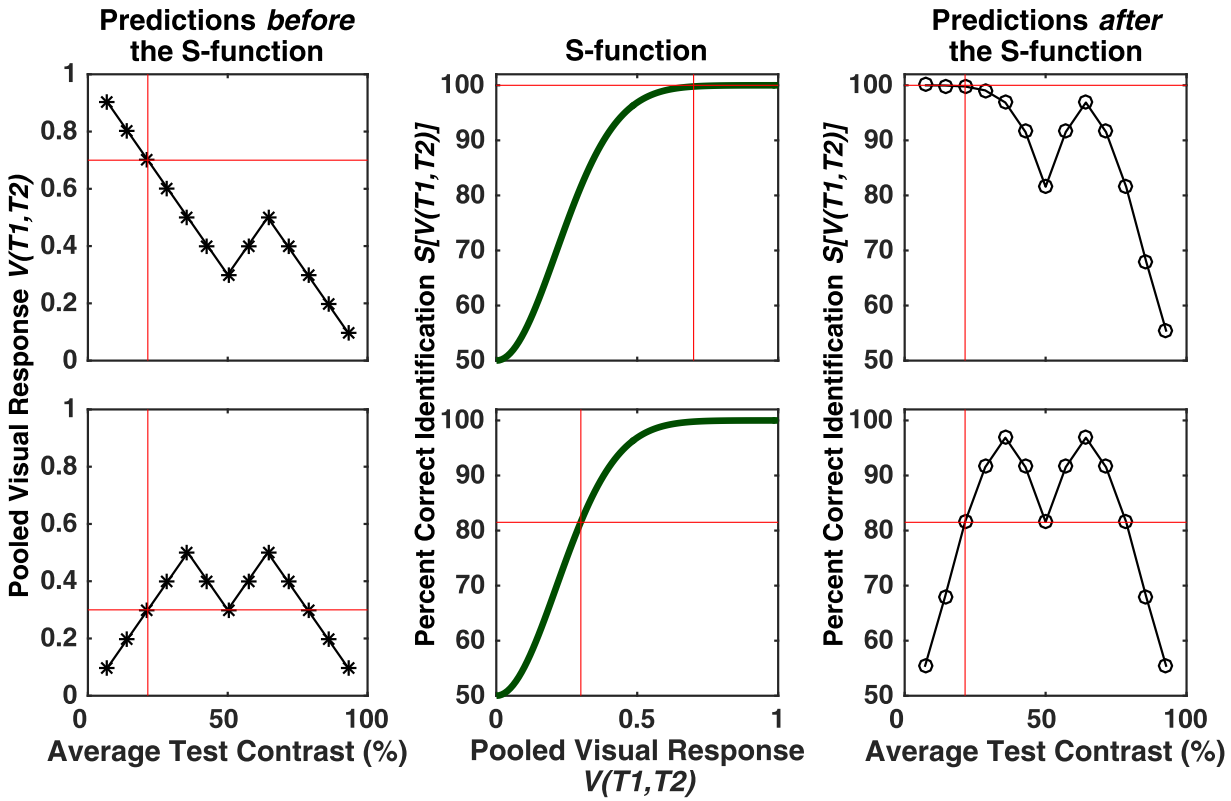

Figure 12. Illustration of the S-function's action. The top row illustrates the case of a predicted constant-difference series like the idealized one shown in Figure 7 for long test durations. The bottom row illustrates an example like the idealized curve in Figure 7 for short test durations. The left column shows the action at STEP 1 where the pooled visual response *V(T1,T2)* is computed. The middle column shows the action at STEP 2 where the S-function itself (in green, identical in both rows) is applied to *V(T1,T2)*. The right column shows the predictions for the observer's performance after both STEP 1 and STEP 2. The red lines are described in the text.

fact include some of the response to the onset.) If an observer uses this quantity to make their decision on each trial in our experiment, we say the observer is using the OFFSET decision rule.

In the appendix, there are examples of $R1(t)$ and $R2(t)$ along with the three values of pooled visual response $V(T1,T2)$ that result from using the three different decision rules.

When there is no subscript at all, that is, $V(T1,T2)$, it means that the statement is true no matter which of the three specialized cases is considered.

STEP 2 - Apply monotonic S-function

We now need to get from the pooled visual response $V(T1,T2)$ to the predicted percent correct identification by an observer.

For STEP 2, we are going to make the following very simple assumption: A monotonic function – which we call the S-function – exists. It is shown in the box labeled "STEP 2" in Figure 11. This S-function is a simple stand-in for several stages of higher-level perceptual and cognitive processing.

This S-function takes as input the single number that is the value of pooled visual response V for a particular pair of test contrasts $(T1, T2)$. This number $V(T1,T2)$ ranges from 0 to something arbitrarily high.

The S-function produces as output a predicted percent correct identification. This is a number in the range from chance (50%) to perfect performance (100%). Thus,

$$\text{Predicted percent correct identiffication} = S[V(T1,T2)].$$

For the purposes of this paper, we do NOT need to postulate any particular form for the function $S(V)$ other than it being monotonic and generally S-shaped. In addition, accordingly, we do NOT present figures showing the predicted percent correct identification for different patterns. Instead, we will show plots of the values of $V(T1,T2)$ for the different patterns (as is done in Figures 13, 14).

Why do we NOT specify a particular function $S(V)$? It is possible – and to our minds preferable – to NOT specify any particular form for the function $S(V)$. The fact that $S(V)$ is monotonic and generally S-shaped will be sufficient to make the comparisons of model predictions to the experimental results (observers' percent-correct) that are needed to support the conclusions we draw here. Not specifying a particular $S(V)$ seems preferable because it produces a more transparent picture of the strengths and weaknesses of the input-side of our model (Figure 8). $S(V)$ obscures the action of the input side of the model by removing information from the data (by compressing it at the top and bottom of the scale). In addition, there seems to be no reason to add another few parameters to the model when we know that our data and arguments from it would not constrain these parameters in any interesting way. Indeed, the parameters would simply trade off.

Example illustrating effect of a monotonic S-function

Figure 12 illustrates the action of the monotonic S-function. The top row shows one example and the bottom row shows another example. The top row shows a case where the left end of the constant-difference series curve is high. The bottom row shows a case where the left end is low. The right ends are identical in the two rows.

The left column shows hypothetical examples of constant-difference-series curves *before* the action of the S-function. Average test contrast is plotted on the horizontal axis and the pooled visual response, $V(T1,T2)$, is plotted on the vertical axis. (As a reminder, this quantity $V(T1,T2)$ will be shown for the predictions.)

The middle column shows the S-function itself (which is the same in both rows). The pooled visual response $V(T1,T2)$ is plotted on the horizontal axis and percent correct identification, $S[V(T1,T2)]$, is plotted on the vertical axis.

The right column shows the constant-difference-series curve *after* the action of the S-function. Average test contrast is plotted on the horizontal axis and percent correct identification, $S[V(T1,T2)]$, is plotted on the vertical axis.

The reader can follow the action of the S-function by following the red lines in each row. In the upper left panel, the red lines show that at STEP 1 an average test contrast of 20% produces a pooled visual response of 0.7. In the upper middle panel, the red lines show that at STEP 2 a pooled visual response of 0.7 leads to a percent correct identification of 100%. The upper right panel plots the combined action of STEP 1 and STEP 2, showing that an average test contrast of 20% produces a percent correct identification of 100%. Similarly, one could follow the action in the bottom row by following the red lines there.

For a wide range of pooled visual responses $V(T1,T2)$ – from 0 to reasonably high – the S-function does not do much. But for high values of pooled visual response, the S-function compresses the output so that the output stays nearly constant. This can be seen in the top row of Figure 12 in the panel showing predictions *after* the S-function (upper right panel) where the left end of the curve has been flattened. Compare this case in the upper row with the case in the lower row. In the lower row, the curve after the S-function (lower right panel) maintains the general shape of the curve before the S-function (lower left panel).

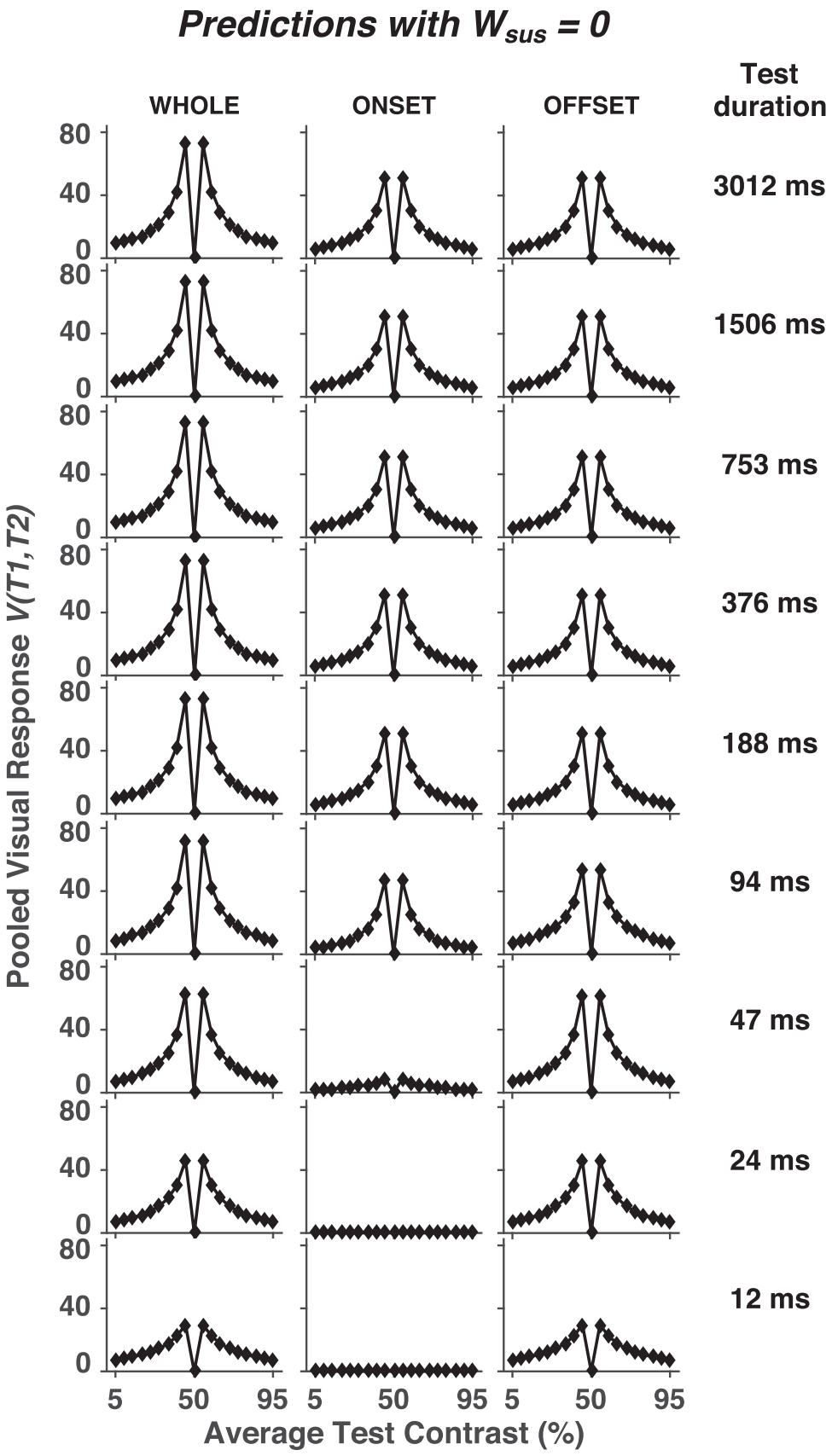

Figure 13. Predictions of the model for the pooled visual response *V(T1,T2)*. This stage is before the S-function has acted. Shown here are the predictions when the weight on the key sustained inhibitory mechanism (W_{sus}) equals zero, that is, the key sustained inhibitory mechanism is NOT active. Constant-difference series curves (for a test contrast difference of 10%) are shown for three decision rules (in different columns) and various test durations (in different rows). Values for the other model parameters are given in Figure 9. The predictions shown here cannot account for the observers' performances.

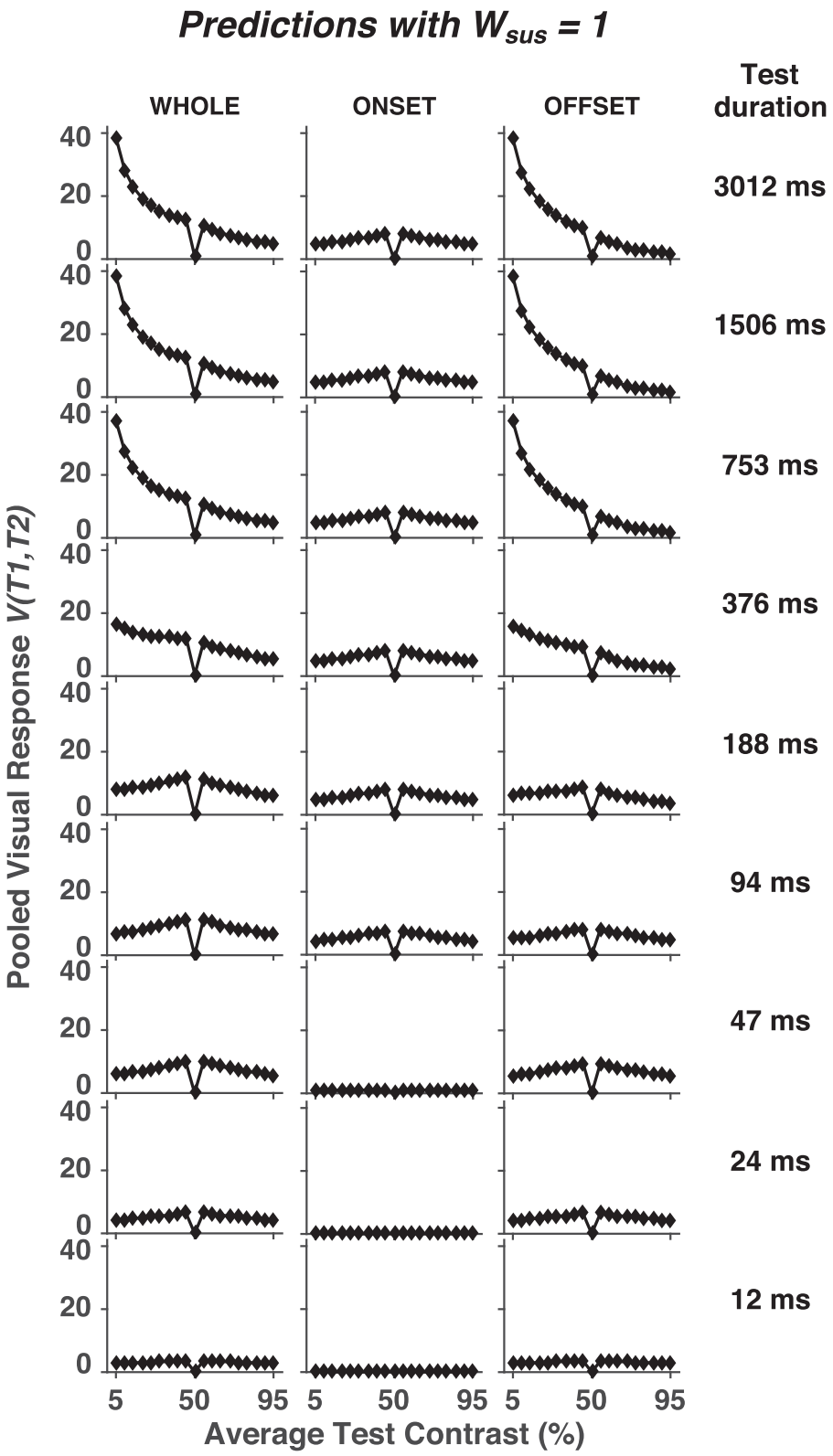

Figure 14. Predictions of the model for the pooled visual response *V(T1,T2)* with W_{sus} =1. Other parameters are the same as for Figure 13. This stage is before the S-function has acted. The predictions shown here (in Figure 14) with the WHOLE and OFFSET decision rules can account for the observers' performances.

Discussion: Comparing predictions to data

Predictions from the model are shown in Figures 13 and 14. These two figures show predictions from the model with one critical parameter changed, namely, W_{sus} is 0 in Figure 13 and W_{sus} is 1 in Figure 14. Each of the panels in these figures shows a constant-difference-series curve. The three columns show predictions from the three different decision rules (WHOLE, ONSET, and OFFSET). The rows show predictions for different test durations.

Let us first look at Figure 13, which shows predictions when $W_{sus} = 0$, in other words, when only two of the three key mechanisms – transient excitation and transient inhibition – are active. (The key sustained inhibitory mechanism is NOT active.) All these predictions are symmetric around an average test contrast of 50%, that is, the predictions are symmetric around the adapt contrast which is 50%. The symmetry around the adapt contrast of 50% will not be altered by the action of any S-function. Now compare the experimental results in Figures 5 and 6 to these predictions in Figure 13. That comparison makes it clear that the predictions are dramatically at odds with the experimental data: they cannot account for the rise of the left end of the observer's performance curve as test duration lengthens.

Now let us look at the predictions when all three key mechanisms are active (Figure 14 for which $W_{sus} = 1$). The predictions in the middle column use the ONSET decision rule. These predictions are always symmetric and therefore do not show the rise of the left end of the observer's performance curve (Figures 5, 6). Logically the model using only onset information could not possibly predict the effect of test duration.

The predictions in the left and right columns of Figure 14 (labeled WHOLE and OFFSET) do use test-pattern offset information. The WHOLE predictions use both test-pattern onset and offset information, whereas the OFFSET predictions use only the offset information. Both sets of predictions show a rise at the left end of the curve for long test durations. Once the S-function has exerted its compressive action, these predictions would be a good account of the observers' performance. Consider what happens with short test durations (bottom row in Figure 12 and bottom few rows of Figure 14). Even after the S-function, the predictions will remain symmetric and thus closely resemble the experimental data (Figures 5, 6). Now consider what happens with long test durations (top row in Figure 12 and top few rows of Figure 14). After the S-function, the predictions will now closely resemble the experimental data (Figures 5, 6) in having high flat left ends.

At a broad-brush level, one can understand the long test duration predictions with $W_{sus} = 1$ (i.e. the key sustained inhibitory mechanism is active) as follows: once the test-pattern duration gets long enough, the contrast of the test pattern (rather than that of the preceding adapt pattern) dominates the inhibitory signal. For low test contrasts (the left end of the constant-difference-series curve), by the time the test pattern has been on for a while, it will be like having had a very low-contrast gray adapt pattern on for a while. An example of this for a very low test contrast is shown in Figure 10. The sustained inhibition (right column, third row from the top) rises to the level of the adapt contrast during the adapt pattern but then drops down to the level of the test contrast during the long duration test pattern.

In the results of other dynamic experiments that we are not presenting in this paper but plan to present subsequently, we will find it necessary to sometimes use $W_{sus} = 1$ (as in this paper) and at other times to use a much smaller value of W_{sus} (such as $W_{sus} = 0$). We currently view this as meaning that there are two different subsystems (or channels) of the visual system – one with sustained inhibition and one without – that get recruited for the observer's performance in our full set of dynamic experiments.

Summary

The work presented in this paper is a first step in trying to characterize the dynamic properties of the contrast-comparison process as they show up in human visual perception. This work also leads to some further characterization of the contrast-normalization process.

We present a simple dynamic model to account for our experimental results. The model contains a key excitatory mechanism that is transient, and responds positively to both increases and decreases of contrast. The model also contains two key inhibitory mechanisms, one that is transient and one that is sustained. The contrast-comparison process is carried out by the key excitatory mechanism. The contrast-normalization process is carried out by the two key inhibitory mechanisms.

The experimental results show that as the duration of the test pattern increases, performance on the straddle pattern improves little (which initially surprised us). In the model, this result is a direct consequence of the transient nature of the key excitatory mechanism.

In addition, the experimental results show that, as the test pattern duration increases, the left end (low average test contrast) of the constant-difference-series curve rises up. In the model, this rise is a consequence of the combined action of the transient and sustained key inhibitory mechanisms.

Keywords: contrast adaptation, dynamics, straddle effect, contrast comparison, contrast normalization

Acknowledgments

The authors thank all of our observers for their many hours of work.

Supported in part by the National Eye Institute Grant EY08459. Some of the results were presented at the Vision Sciences Society annual meeting (Graham, Wolfson, & Patterson, 2013).

Commercial relationships: none.
Corresponding author: Norma V. Graham.
Email: nvg1@columbia.edu.
Address: 406 Schermerhorn Hall, Department of Psychology, Columbia University, New York, NY 10027, USA.

References

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience, 13*(1), 51–62.

Chariker, L., Shapley, R., & Young, L. (2020). Contrast response in a comprehensive network model of macaque V1. *Journal of Vision, 20*(4):16, 1–19.

Cleland, B. G., Dubin, M. W., & Levick, W. R. (1971). Sustained and transient neurones in the cat's retina and lateral geniculate nucleus. *Journal of Physiology, 217*(2), 473–496.

Foley, J. M. (1994). Human luminance pattern-vision mechanisms: masking experiments require a new model. *Journal of the Optical Society of America A, 11*(6), 1710–1719.

Foley, J. M. (2011). Forward-backward masking of contrast patterns: The role of transients. *Journal of Vision, 11*(9): 15, 1–24.

Graham, N., & Hood, D. (1992). Modeling the Dynamics of Light Adaptation: the Merging of Two Traditions. *Vision Research, 21*(7), 1373–1393.

Graham, N., & Sutter, A. (2000). Normalization: contrast-gain control in simple (Fourier) and complex (non-Fourier) pathways of pattern vision. *Vision Research, 40*(20), 2737–2761.

Graham, N., & Wolfson, S. S. (2007). Exploring contrast-controlled adaptation processes in human vision (with help from *Buffy the Vampire Slayer*). In Laurence R. Harris, & Michael Jenkin (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. Dosher, Z. Lu, & R. Schiffrin (Eds.), *Human Information Processing: Vision, Memory, and Attention* (pp. 13–27). Washington, DC, US: American Psychological Association.

Graham, N. V., & Wolfson, S. S. (2018). Is the straddle effect in contrast perception limited to second-order spatial vision? *Journal of Vision, 18*(5):15, 1–43.

Graham, N., Wolfson, S. S., & Patterson, C. A. (2013). Temporal characteristics of the Straddle Effect (Buffy contrast adaptation) and modeling with on-off neurons. *Journal of Vision, 13*, 309.

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience, 9*, 181–197.

Heeger, D. J., & Zemilianova, K. O. (2020). A recurrent circuit implements normalization, simulating the dynamics of V1 activity. *Proceedings of the National Academy of Sciences USA, 117*(36), 22494–22505.

Kulikowski, J. J., & Tolhurst, D. J. (1973). Psychophysical evidence for sustained and transient detectors in human vision. *Journal of Physiology, 232*, 149–162.

Legge, G. E., & Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America, 70*(12), 1458–1471.

Motoyoshi, I., & Kingdom, F. A. A. (2007). Differential roles of contrast polarity reveal two streams of second-order visual processing. *Vision Research, 47*(15), 2047–2054.

Pokorny, J., & Smith, V. C. (1997). Psychophysical signatures associated with magnocellular and parvocellular pathway contrast gain. *Journal of the Optical Society of America A, 14*(9), 2477–2486.

Sawada, T., & Petrov, A. A. (2017). The divisive normalization model of V1 neurons: a comprehensive comparison of physiological data and model predictions. *Journal of Neurophysiology, 118*(6), 3051–3091.

Shooner, C., & Mullen, K. T. (2020). Enhanced luminance sensitivity on color and luminance pedestals: Threshold measurements and a model of parvocellular luminance processing. *Journal of Vision, 20*(6):12, 1–1.

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

Teo, P. C., & Heeger, D. J., (1994). Perceptual image distortion. In B. E. Rogowitz, & J. P. Allebach (Eds.), *Human Vision, Visual Processing, and Digital Display V. Proc. SPIE 2179* (pp. 127–141). Retrieved from https://spie.org//Publications/Proceedings/Volume/2179?SSO=1.

Watson, A. B., & Solomon, J. A. (1997). Model of visual contrast gain control and pattern masking. *Journal of the Optical Society of America A, 14*(9), 2379–2391.

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.

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.

Appendix

This appendix gives a fuller explanation of why the model with $W_{sus} = 1$ predicts the raising of the left

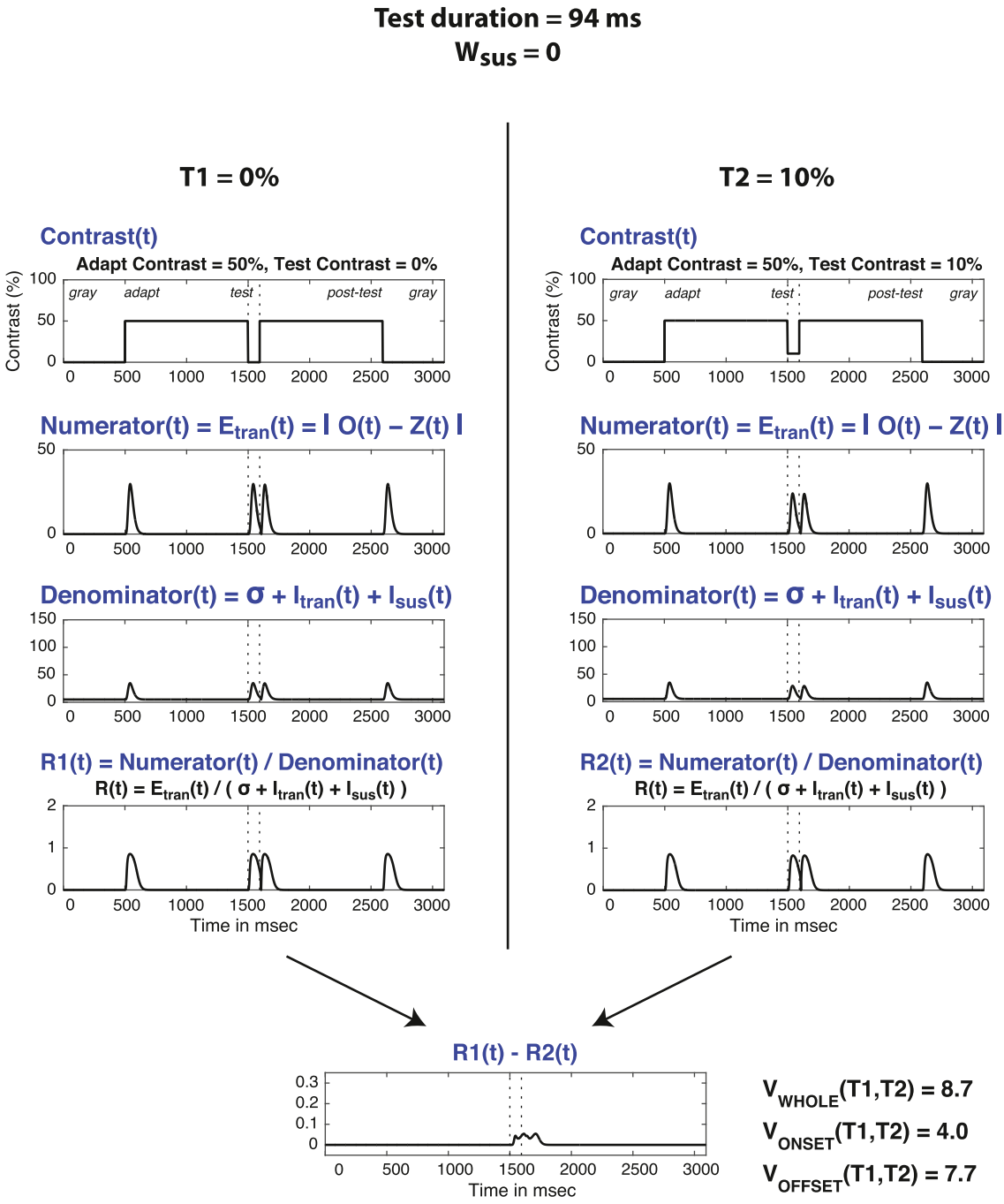

Figure 15. Illustration of the calculation of the model's predictions for a test duration of 94 ms with $W_{sus} = 0$.

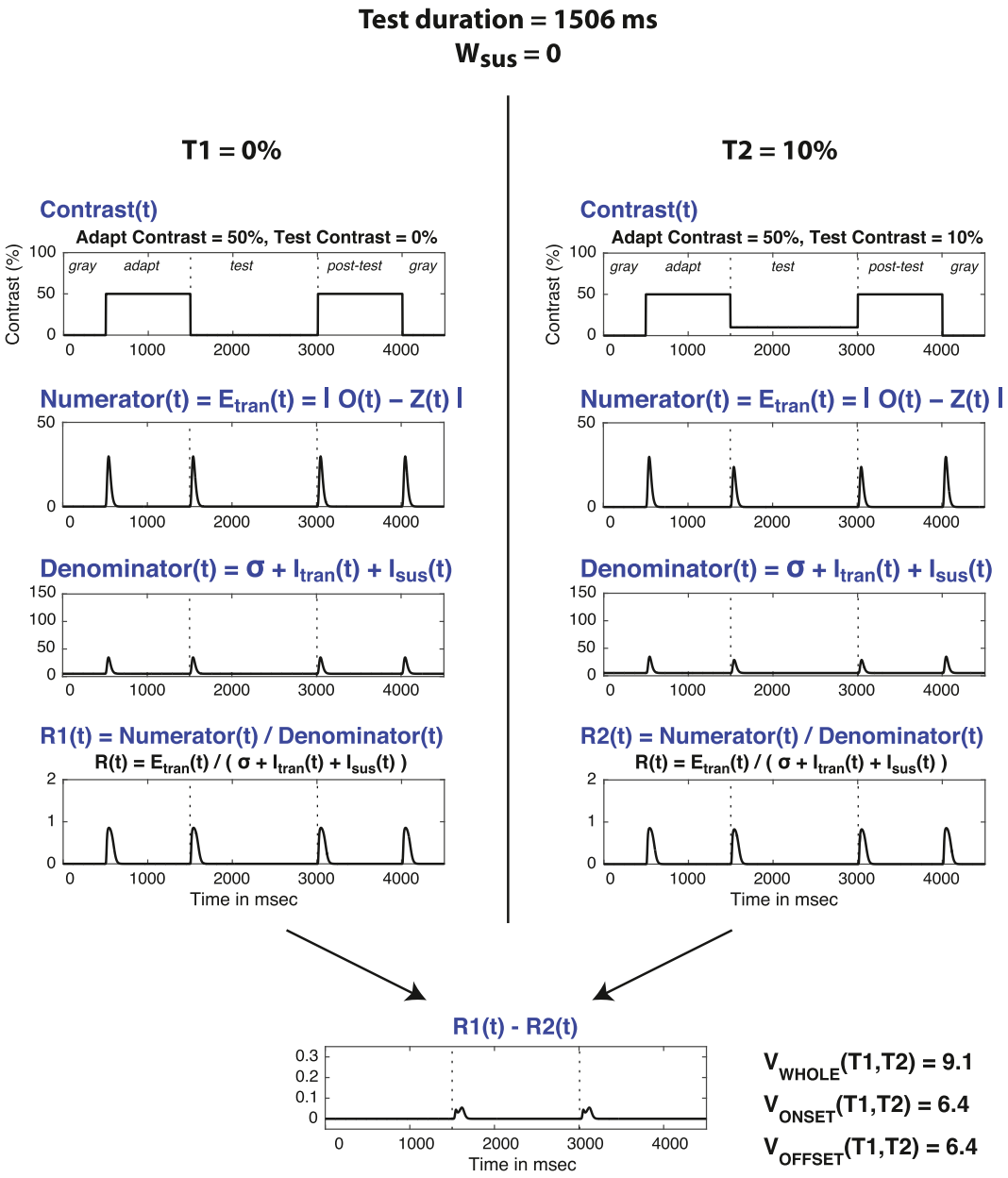

Figure 16. Illustration of the calculation of the model's predictions for a test duration of 1506 ms with $W_{sus} = 0$.

side of the constant-difference-series curve when test duration increases. We will assume the reader is familiar with the description of the model given in the Modeling section. The following equation gives the response of the model to a Gabor patch as a function of time:

$$R(t) = \frac{E_{tran}(t)}{\sigma + I_{tran}(t) + I_{sus}(t)}$$

The figures in this appendix illustrate how the model works in making predictions for the far left end of constant-difference-series curves for two test durations (94 ms and 1506 ms) and for two values of W_{sus} (0 and 1).

The first four figures in this appendix (Figures 15–18) have the following format: The left column shows the case for a test contrast of 0% (T1 = 0%), and the right column for a test contrast of 10% (T2 = 10%). Each column shows, as a function of time, the pattern contrast (top row), the numerator of the model equation (second row), the denominator of the model equation (third row), and the quotient (fourth row). This quotient is labeled $R1(t)$ in the left column and $R2(t)$ in the right column. At the bottom of the figure is the difference $R1(t)$-$R2(t)$. Also shown there is the value of the pooled visual response $V(T1,T2)$ for each of the three observer decision rules we explored (WHOLE, ONSET, and OFFSET).

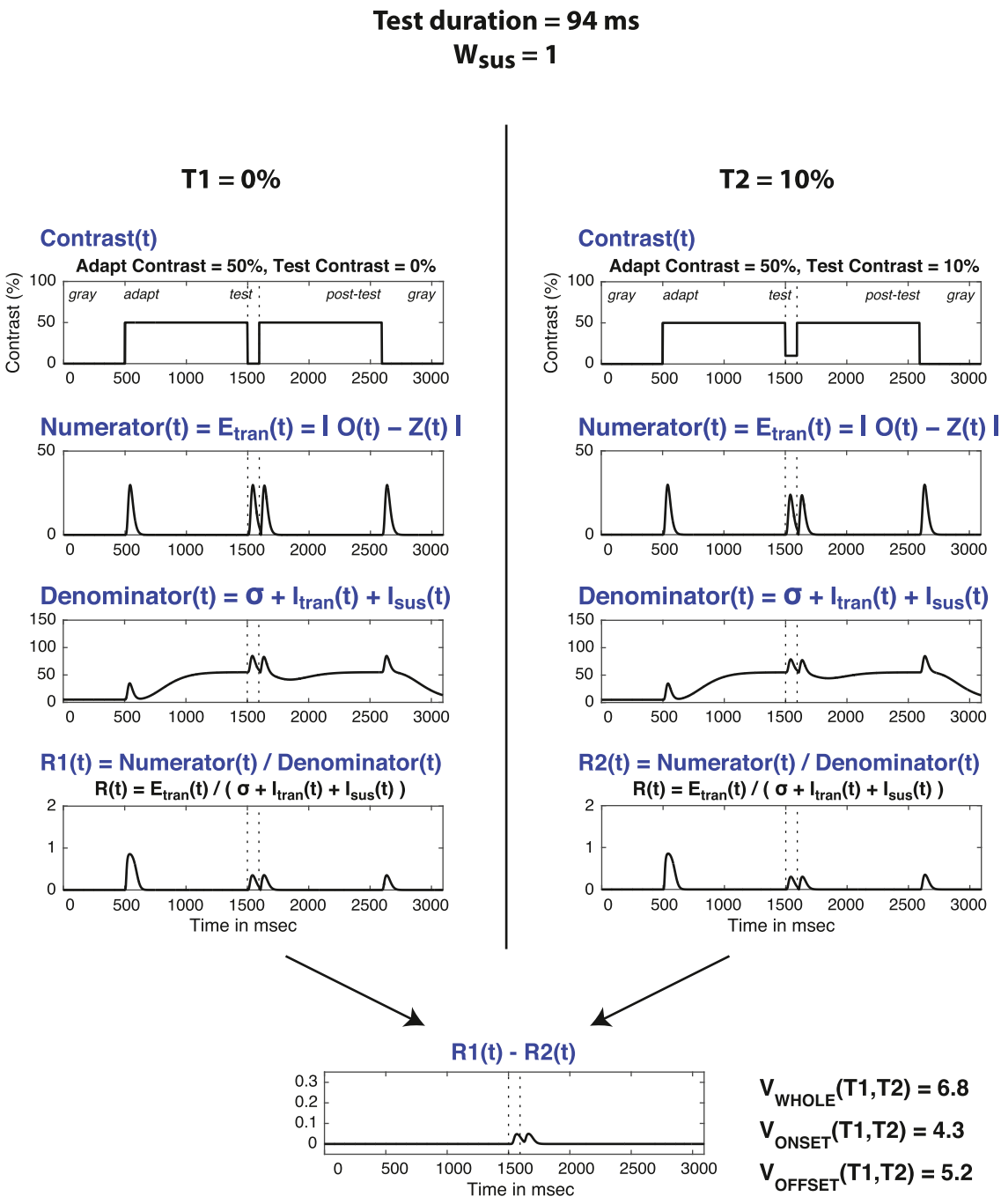

Figure 17. Illustration of the calculation of the model's predictions for a test duration of 94 ms with $W_{sus} = 1$.

The fifth figure in this appendix (Figure 19) is a summary showing the values of the pooled visual response V for the three different decision rules, the two test durations, and the two values of W_{sus} used in the preceding four figures.

Let us start by looking at the two figures with $W_{sus} = 0$ (Figures 15, 16). In these figures, the curves showing the numerator, denominator, and the quotient of the equation (second, third, and fourth rows) are all of similar shape: they have four quick pulses of activity generally separated by horizontal line segments. In any one panel, the four quick pulses are of approximately equal height. The four quick pulses in each curve occur immediately after four events in the visual stimulus: the adapt pattern's onset, the test pattern's onset, the test pattern's offset, and finally the adapt pattern's offset. The quick pulses in the numerator are the response of the key excitatory mechanism, which is transient. The quick pulses in the denominator reflect the response of the key transient inhibitory mechanism. (Because

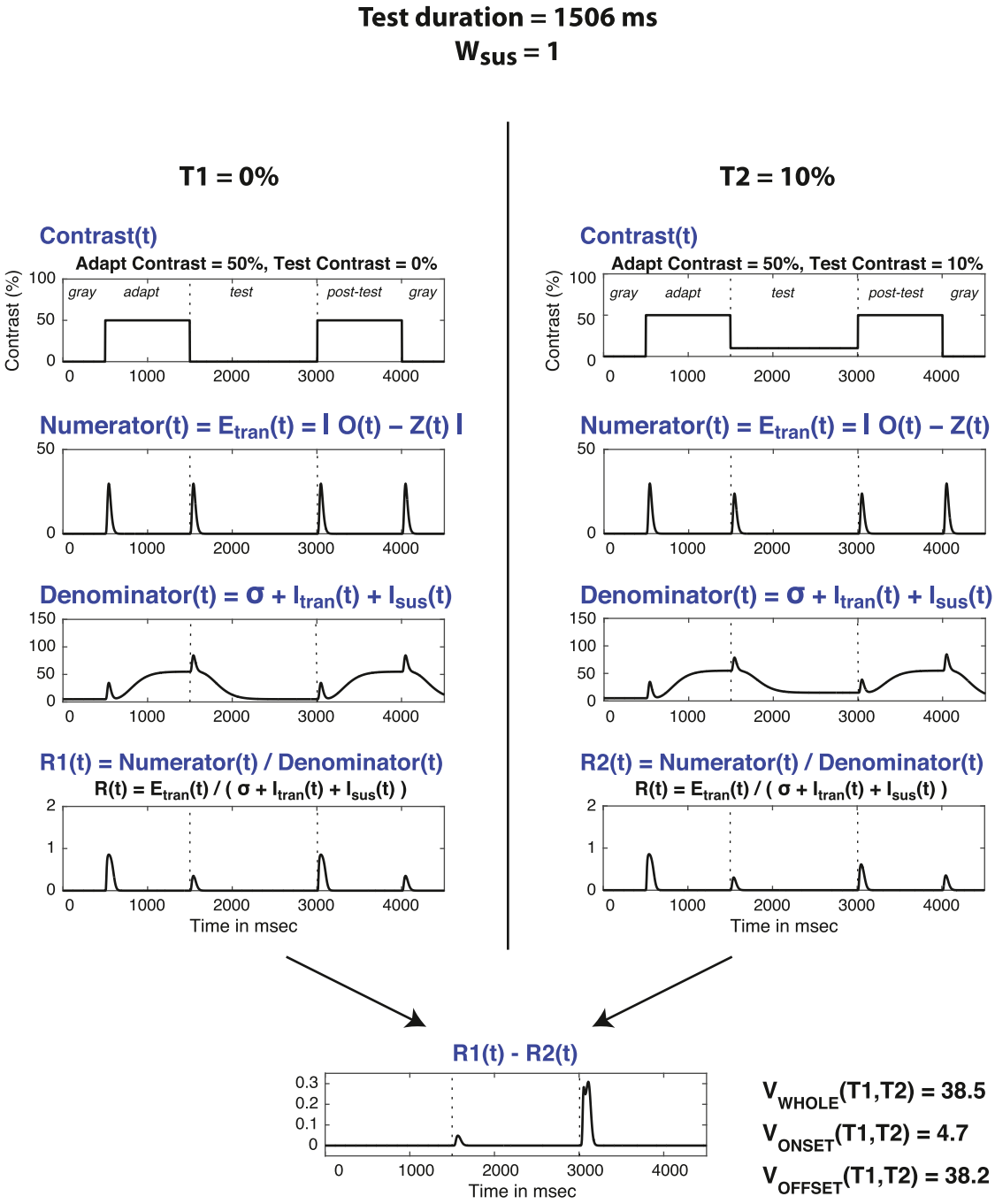

Figure 18. Illustration of the calculation of the model's predictions for a test duration of 1506 ms with $W_{sus} = 1$. The outputs in the right column of this figure are also shown in Figure 10. Figure 10 also includes panels which show the individual components of the numerator and denominator functions.

$W_{sus} = 0$ here, there is only transient inhibition, not sustained inhibition.)

Now let us look at the two figures with $W_{sus} = 1$ (Figures 17, 18) and consider the effect of test duration:

- When the test duration is 94 ms (Figure 17), the amount of inhibition coming from the key sustained inhibitory mechanism – both at the beginning and at the end of the test pattern – will be determined primarily by the contrast of the adapt pattern that preceded the test pattern (and not affected by the contrasts T1 and T2 in the test pattern itself).
- When the test duration is 1506 ms (Figure 18), the amount of inhibition coming from the key sustained inhibitory mechanism at the

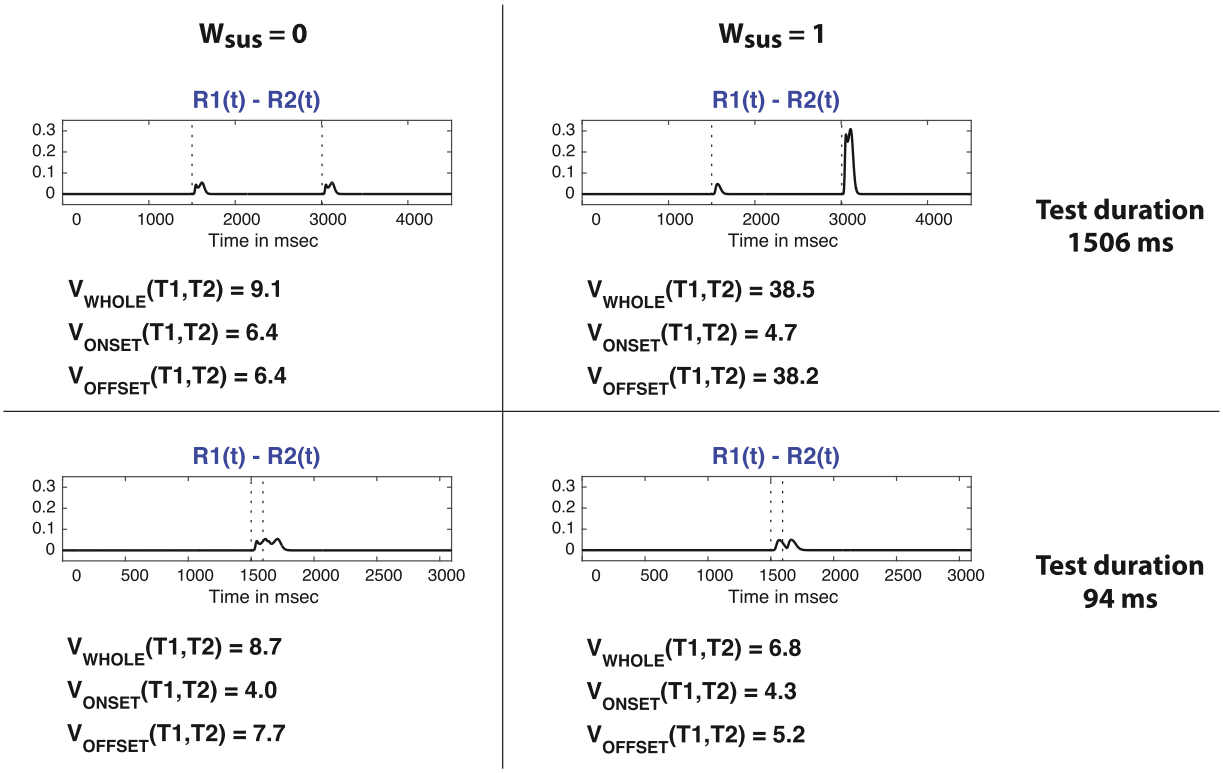

Figure 19. The pooled visual response *V* for the three different decision rules (WHOLE, ONSET, and OFFSET), the two test durations, and the two values of W_{sus} shown in Figures 15–18.

beginning of the test pattern is not affected by the test-pattern contrasts T1 and T2. However, the amount of inhibition at the end of the test pattern is heavily influenced by the test-pattern contrasts.

Next, let us look at the two figures with test duration of 1506 ms and consider the effect of changing from $W_{sus} = 0$ to $W_{sus} = 1$ (Figures 16, 18). This change does not affect the numerator (the output of the key excitatory mechanism) at all. But it does change the denominator (the output of the key inhibitory mechanisms) and hence the quotient. The denominator has four quick pulses (due to the key transient inhibitory mechanism) in both cases. However, for the case of $W_{sus} = 1$, there are slow rises or falls (due to the key sustained inhibitory mechanism) in between those four quick pulses rather than horizontal lines (as in the case of $W_{sus} = 0$).

At the bottom of each of the first four figures in the appendix (Figures 15–18), is a panel showing $R1(t)$-$R2(t)$ for a pattern in which the two test contrasts are 0% and 10%. The bottom panels from these first four figures are shown again in Figure 19.

As you can see by examining Figure 19, when $W_{sus} = 1$ and the test duration is long (upper right panel), the value of the pooled visual response $V(T1,T2)$ is large (approximately 38) when the observer uses the WHOLE or OFFSET decision rule. However, when $W_{sus} = 1$ and the test duration is short (lower right panel), $V(T1,T2)$ is always small (less than 7). Thus, when $W_{sus} = 1$, and the WHOLE or OFFSET decision rule is used, the left end of the constant-difference-series curve for a long test duration is predicted to be substantially higher than the left end for a short test duration (as in Figure 14). This model prediction is consistent with the experimental results (Figure 5).

However, when $W_{sus} = 0$ (Figure 19, left column), the value of the pooled visual response $V(T1,T2)$ is always small (less than 10). Thus, the left end of the constant-difference-series curve is not affected by test duration (as in Figure 13). This model prediction is inconsistent with the experimental results (Figure 5).