Distributions of alternation rates in various forms of bistable perception

Jan W. Brascamp   Functional Neurobiology and Helmholtz Institute, Utrecht University, Utrecht, The Netherlands

Raymond van Ee   Physics of Man and Helmholtz Institute, Utrecht University, Utrecht, The Netherlands

Wiebe R. Pestman   Centre for Biostatistics, Utrecht University, Utrecht, The Netherlands

Albert V. van den Berg   Functional Neurobiology and Helmholtz Institute, Utrecht University, Utrecht, The Netherlands

Studying the temporal dynamics of bistable perception can be useful for understanding neural mechanisms underlying the phenomenon. We take a closer look at those temporal dynamics, using data from four different ambiguous stimuli. We focus our analyses on two recurrent themes in bistable perception literature. First, we address the question whether percept durations follow a gamma distribution, as is commonly assumed. We conclude that this assumption is not justified by the gamma distribution's approximate resemblance to distributions of percept durations. We instead present two straightforward distributions of reciprocal percept durations (i.e., rates) that both easily surpass the classic gamma distribution in terms of resemblance to empirical data. Second, we compare the distributions arising from binocular rivalry with those from other forms of bistable perception. Parallels in temporal dynamics between those classes of stimuli are often mentioned as an indication of a similar neural basis, but have never been studied in detail. Our results demonstrate that the distributions arising from binocular rivalry and other forms of bistable perception are indeed similar up to a high level of detail.

Introduction

When subjects are shown a visual stimulus that has two distinct interpretations, those interpretations will take turns reaching awareness, a phenomenon called bistable perception (for examples, see Figure 1). Bistable perception has over the years been studied by numerous researchers, because it is an interesting phenomenon in its own right but also because the partial decoupling of stimulus and percept might help to gain insight into the relation between awareness and brain function. Irrespective of these efforts, the neural mechanisms causing bistability are still under debate. One way to gain insight into these mechanisms is by studying the temporal dynamics of perceptual alternations. For a subject experiencing bistable perception, the duration of a percept commonly varies widely from one to the next, so speaking of a mean percept duration is of limited use. In contrast to this unpredictability, one of the most reproducible aspects of bistable perception is the shape of percept duration distributions. Invariably, these distributions are unimodal with a skew toward high durations (like the top curves in Figure 2). It has long been rec-

ognized that these distinctly shaped distributions may provide clues to understanding the neural processes involved.

In this study, we investigate the temporal dynamics of bistable perception, using data that were previously collected using four different ambiguous stimuli (Figure 1). We focus our analyses on two recurrent themes in the literature on the subject. First, we address the question whether distributions of percept durations follow a gamma distribution. As will be discussed shortly, there is reason to believe that this common assumption is not backed by empirical findings. Second, we compare distributions over different stimuli, because parallels and differences in temporal dynamics between stimuli can serve as an indication for parallels and differences in underlying neural mechanisms.

Gamma distributed percept durations?

An interesting aspect in the study of distributions of percept durations is their resemblance to known statistical distributions, because these are often associated with clearly defined mechanisms. One such theoretical distribution with the desired unimodal, right-skewed shape is the
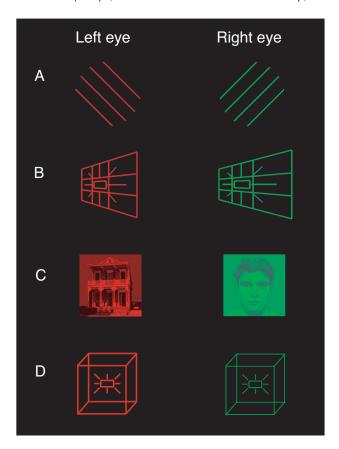
Figure 1. Left and right eye components of the four stimuli for which bistable perception was examined: orthogonal gratings (A), bistable slant (B), house-face (C), and Necker cube (D). See text for details. Adapted from Van Ee (2005).

gamma distribution, introduced in this context by Levelt (1967). Its probability density function is given by

$$f(t \mid k, \lambda) = \frac{1}{\lambda^k \Gamma(k)} t^{k-1} e^{-\frac{t}{\lambda}} \ .$$

By definition, $\Gamma(n)$ is the canonical continuous extension of $(n-1)!$, which itself is of course only defined for natural n. The parameters k and λ in the equation are usually called the shape parameter and the scale parameter, respectively.

Among other similarly shaped distributions, the gamma distribution is noteworthy because it has over the last decades become the standard for performing fits to empirical distributions of percept durations (e.g., Borsellino, De Marco, Allazetta, Rinesi, & Bartolini, 1972; Kovacs, Papathomas, Yang, & Feher, 1996; Logothetis, Leopold, & Sheinberg, 1996; Murata, Matsui, Miyauchi, Kakita, & Yanagida, 2003; Walker, 1975), even though Levelt himself acknowledged the fact that "other functions may fit as well." Indeed, four studies that statistically analyze the gamma distribution's fit performance do not univocally show a good fit to empirical data. The authors of two such studies (Borsellino et al., 1972; De Marco,

Penengo, & Trabucco, 1977) judged gamma distributions to fit their data acceptably well for their purposes, but their analyses leave considerable room for doubt. (Borsellino and coworkers stated that around 15–30% of their gamma fits have a chi^2 probability lower than 1%, and De Marco and colleagues mentioned two alternative theoretical distributions to fit equally well as the gamma distribution, although less favorable in the light of parsimony.) The two remaining studies (Cogan, 1973; Zhou, Gao, White, Merk, & Yao, 2004) show an unacceptable fit quality for the gamma distribution: In both cases more than half of the fitted distributions should be rejected at the 5% significance level. As a point in favor of the gamma distribution, it should be mentioned that on the basis of the above studies, one cannot identify an alternative distribution with better fit performance. Although Zhou et al. (2004) did show the lognormal distribution to fit their data better than the gamma distribution, Cogan (1973) rejected the lognormal distribution as an acceptable fit to her data (note that lognormal distributed percept durations were also proposed by Lehky, 1995).

What makes the question whether the gamma distribution is appropriate particularly interesting is the neural model that is associated with it. As Levelt (1967) pointed out, one process known to lead to gamma-distributed latencies between events is one in which every event occurs after a fixed number of consecutive ticks from a "Poisson clock" (i.e., randomly occurring ticks). Therefore, he proposed that "the summative effect of . . . successive spikes from the recessive [percept] is necessary and sufficient to re-establish dominance for that [percept]," without defining the exact nature of these Poisson "spikes" or ticks. In the case of a Poisson clock, λ in the above equation is the basic duration between two ticks, while the value of the shape parameter k is the number of ticks causing a perceptual alternation.

One of our goals is to use our data to compare the fit quality of the gamma distribution and alternative distributions. The question is, what alternative distributions are viable candidates? We believe a step to finding such distributions could be to switch the focus of attention away from distributions of percept durations toward distributions of *reciprocal* percept durations, or rates. Because distributions of alternation rates are just as characteristic for bistable perception as those of percept durations, this provides no a priori reason to make fits to one or the other. We believe however that distributions of alternation rates can increase our understanding of underlying processes. This idea stems from saccadic (i.e., rapid eye movement) search literature, in which there is some tradition of using such an approach. There are several parallels between saccadic search and bistable perception, one of them being the timing of saccades: The distribution of periods that a subject focuses between saccades has a shape that is very similar to that of distributions of percept durations. Analogous to the practice in bistability literature, much research has been aimed at linking saccadic latency distributions to the neural mechanisms involved (for an overview, see Smith &

Ratcliff, 2004). However, Carpenter (1981) proposed that these mechanisms might be more directly represented in distributions of reciprocal latencies instead of latencies, and since then various studies have made use of this idea (Carpenter & Williams, 1995; Reddi & Carpenter, 2000; Van Loon, Hooge, & Van den Berg, 2002). The key element underlying these authors' points of view is the notion of a decision network that in the presence of an extrafoveal target accumulates information about that target, and causes a saccade to be made toward it as soon as this rising decision signal exceeds a certain threshold level. If one now assumes this threshold level to be fixed over trials, then variations in the rate of information accumulation are directly reflected in the reciprocal latency, or rate, of saccades. (Specifically, Carpenter showed that if this rate of accumulation is subject to Gaussian perturbations, a cumulative latency distribution as a function of reciprocal time will form a straight line on probability paper, which is the same as to say that reciprocal latency will have a Gaussian distribution.) What we propose in analogy regarding bistable perception is that distributions of *perceptual alternation rates* (i.e., reciprocal percept durations), being more directly related to underlying neural processes than distributions of percept durations, could give insight into neural processes. This especially makes sense in light of the idea that the numerous phenomenological similarities between saccadic search and bistable perception are not merely coincidental, but instead reflect a more fundamental link between the two (Leopold & Logothetis, 1999). Completing the analogy between rate models in saccadic search and in bistable perception, we therefore suggest a decision signal that starts rising at the beginning of a dominance period of percept *A*, and causes an alternation to percept *B* as soon as it reaches its threshold, so that characteristics of the rate of information accumulation are reflected in distributions of perceptual alternation rates.

In this study we pit two different rate distributions against the standard gamma fit. First, we explore the possibility that alternation rates have a gamma distribution, which means we take the reciprocal of our percept durations and perform gamma fits to the resulting distributions. We will hereafter refer to these fits as "gamma rate" fits, not to be confused with the classical "gamma duration" fits. Because Van Loon et al. (2002) showed gamma rate fits to be applicable to saccade data, there is some reason to expect an acceptable fit to our data, too. The second distribution we fit to our rate data is the beta' (pronounce: "beta prime") distribution, a two-parameter distribution related to the gamma distribution,[1] and similar in shape. Its probability density function is given by

$$f(r \mid S, F) = \frac{\Gamma(S+F)}{\Gamma(S)\Gamma(F)} \frac{r^{S-1}}{(1+r)^{S+F}} \ .$$

This distribution arises from a model that assumes two rising decision signals instead of one: one in favor of maintaining the present percept, the other in favor of switching.

These signals both increase toward a threshold over time, but compete for the same resource, so that at any given time only one of them can gain an increment (i.e., the model is based on a Bernoulli process). In terms of this model, the parameters *S* and *F* of the beta' distribution are determined by the thresholds for switching percepts and maintaining the current one, respectively. The parameter *r* is proportional to the observed alternation rate via a constant *R*: rate =*Rr*. In the present work, we fix *R* at 1 s^{-1}, performing our fits with a two-parameter distribution. The main reason for this is that the full three-parameter version caused divergence in our fitting algorithm, while much of the distribution's flexibility can already be achieved by varying only the two remaining parameters. It should be kept in mind however that removing this constraint on the third parameter would provide room for an improved fit quality compared to what we present here. A complete derivation of the beta' distribution from the model assumptions is given by Van den Berg and Van Loon (in press).

Apart from the neural model, there is one particular characteristic that makes the beta' distribution attractive in the present context; namely, it is invariant under inversion. That is to say that if variable *A* has a beta' distribution (with parameters *S* and *F*), then 1/*A* has a beta' distribution, too (with parameters *F* and *S*). The important implication here is that the beta' distribution will by definition fit our distributions of percept durations equally well as our distributions of alternation rates, thereby in a sense bridging the gap between the two approaches. Figure 2 gives an impression of the overall shape of the probability density functions (PDFs) and the cumulative distribution functions (CDFs) associated with the gamma duration, the gamma rate, and the beta' rate model. It shows both the overall similarity between them and the more subtle aspects in which they differ.

Interstimulus comparison

There are many stimuli that lead to bistable perception. One particularly well-studied class of stimuli is the one causing binocular rivalry. Binocular rivalry is elicited by presenting dissimilar images to corresponding areas of the two retinas, resulting in a percept that alternates between the two images (e.g., stimulus A and C in Figure 1). In recent literature on bistable perception, much attention has been aimed toward answering the question whether binocular rivalry, on the one hand, and perceptual rivalry of ambiguous figures (e.g., stimulus B and D in Figure 1), on the other, reflect distinct or similar neural mechanisms (for reviews, see Blake, 2001; Blake & Logothetis, 2002; Tong, 2001). Advocates of the former point of view argue that in binocular rivalry, alternations arise as the two monocular channels converge in the primary visual cortex, as a direct result of the incongruence between the retinal images. Researchers favoring the latter position, on the other hand, claim that competition in binocular rivalry is between stimulus representations rather than eyes, generally locating
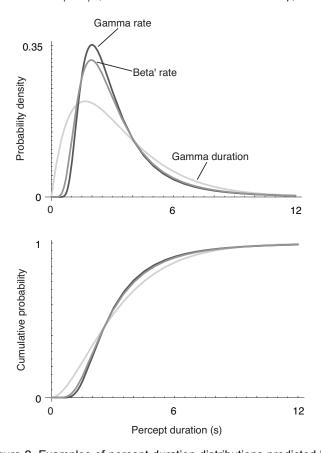
Figure 2. Examples of percept duration distributions predicted by the three models we compare in this study. Top. Probability density functions (PDFs) of percept durations, under the assumption of gamma distributed alternation rates (darkest shade), beta' distributed alternation rates (intermediate shade), and gamma distributed percept durations (lightest shade). To facilitate comparison, the parameters were chosen such that both mean and variance are the same for the three curves. Bottom. The corresponding cumulative distribution functions (CDFs). The PDFs show that all three theoretical distributions satisfy the basic criteria of unimodality and rightward skew, but that the classic gamma duration distribution differs from the other two by having a shorter (in this example, hardly any) initial phase of low slope, and a lower peak. The CDFs, on the other hand, emphasize that any difference between the three distributions is modest in light of the overall similarity. Clearly, these graphs depict distributions of percept durations, but note one subtlety: If one changes the quantity on the horizontal axis to "alternation rate," the graphs can just as easily be read as distributions of alternation rates. The darkest shade then becomes associated with gamma distributed durations and the lightest shade with gamma distributed rates. The intermediate shade remains associated with beta' distributed rates, due to the beta' distribution's invariance under inversion.

it in the higher visual areas and placing it in the same category as reversals in ambiguous figure perception. An often repeated argument in this discussion is that percept durations have similarly shaped distributions in both binocular rivalry and ambiguous figure rivalry (e.g., Logothetis et al., 1996; Logothetis, 1998; Lumer, Friston, & Rees, 1998; also see Carter & Pettigrew, 2003). However, this argument is based on no more than a broad resemblance and a fair fit to the gamma duration distribution. Our second goal in this study is to test whether the claim holds in the face of a detailed comparison of distributions across stimuli.

Data collection

A detailed account of the conditions in which data were gathered is given by Van Ee, Van Dam, and Brouwer (2005). In short, a total of six subjects were tested observing four different stimuli (Figure 1):

Orthogonal grating stimulus. Orthogonal grating rivalry is a classical form of binocular rivalry, where orthogonal gratings are presented to the two eyes. In our case the stimulus consisted of four parallel lines at 45 deg with the vertical that had orthogonal orientations in the two eyes.

Bistable slant stimulus. This is a stimulus first systematically studied by Van Ee, Van Dam, and Erkelens (2002), consisting of a trapezoid that is viewed binocularly. Bistability arises from the fact that linear perspective and binocular disparity specify opposite slants (e.g., in Figure 1B perspective information corresponds to a rectangle seen with the right side in front, whereas disparity information corresponds to a trapezoid seen with the left side in front).

House-face stimulus. This is a stimulus developed by Tong, Nakayama, Vaughan, and Kanwisher (1998). Like the orthogonal gratings described above, it gives rise to a form of binocular rivalry. In house-face rivalry, however, the conflicting images are not orthogonal lines but pictures of a house and a face.

Necker cube stimulus. This is arguably the best known ambiguous stimulus. It is an image that can be interpreted as a cube seen from either of two viewpoints.

The rate of perceptual alternations for all but the slant stimulus typically lies in the range of 0.3 to 1 alternations per second; the slant stimulus generally causes slower alternations, averaging in the order of 0.2 per second. All stimuli were depicted in red and green and viewed through red and green anaglyph glasses to separate the left and right eye image where necessary. Even though there is no need for such separation in Necker cube rivalry, anaglyph glasses were used here as well for consistency. During 3-min trials, a stimulus was viewed on a computer screen, leading to an

unstable percept alternating back and forth between the two alternatives. Alternations were reported by the subjects using button presses. Stimuli covered an angular width of 1.2 deg, and were surrounded by a background of fusible squares to maintain alignment of the eyes. The prevailing notion on the influence of eye movements on bistable perception is that fixation position influences the alternation process, even though eye movements are not necessary for perceptual alternations to occur (Van Dam & Van Ee, in press; Toppino, 2003). Therefore, subjects were instructed to maintain central fixation during all trials, making use of a sunburst figure in the cases of the slant and the Necker cube stimulus. It was established that drift in the alternation rates was restricted to the first 30 s of a session, during which it never exceeded a few percentages of the mean rate, so our data were reasonably stable both across small data chunks and across experimental sessions.

To check whether the subjects were reliably reporting the occurrence of perceptual alternations, rather than pressing keys at random, so-called "catch periods" were included in trials using the slant and orthogonal grating stimuli. During these semi-randomly placed periods stimulus ambiguity was removed respectively by adjusting the disparity-specified slant to become congruent with the perspective-specified slant, or by fading out one of the sets of parallel lines. This was done without the subject being aware of what was happening. Button presses recorded during catch periods were not included in the data analysis; instead, they were used to increase the data's reliability by removing data in case the responses during catch periods were not satisfactory (i.e., if a subject continued reporting perceptual alternations while the stimulus was in fact unambiguous). All button presses in a 20-s period after a catch period had ended were excluded from analysis as well, because examination of the data indicated that in some cases it took some startup time after each catch period for subjects to reach the phase of stable perceptual alternations we were interested in.

Because the data were originally gathered, among other reasons, for investigating the influence of voluntary control on the dynamics of bistable perception, subjects were given one of four different instructions: either to view a stimulus in a natural way without attempting to control the alternation rate, or to try to keep percept A for as great a fraction of the time as possible, or to keep percept B for as great a fraction of the time as possible, or to speed up the alternation rate as much as they could. For each experiment, we treated the durations of the one percept (e.g., the house) separately from the durations of the other percept (e.g., the face). All variables taken together led to 6 (subjects) × 4 (stimuli) × 4 (instructions) × 2 (percepts) = 192 distributions. After removal of the 2% largest and 2% smallest values, the distributions had an average size of 243 data points.

Gamma distributed percept durations?

Methods

We determined the best-fitting gamma and beta' distributions to our empirical data in two different ways. First, we employed the classical least squares method, minimizing the sum of squared residuals *SSE*. For the fits we obtained in this manner, we calculated the fit quality as $SSE/(n-2)$ (where *n* is the number of data points in a given distribution), but also using the Kolmogorov-Smirnov test for goodness of fit. The Kolmogorov-Smirnov test involves the largest overall deviation between empirical and fitted distribution, and the associated probability *pKS*. In addition to these least squares fits, we performed maximum likelihood fits on the same data, providing an alternative estimate of the best-fitting parameter values, as well as of the fit quality: the likelihood *L*. Note that likelihood fits by definition involve PDFs, whereas for the least squares fits we used continuity corrected CDFs, for reasons of robustness and objectivity (contrary to the PDF, the CDF does not involve an arbitrary bin size). All fitting algorithms were implemented in the software package Scilab (http://scilabsoft.inria.fr/).

Results

Figure 3 shows examples of the three distributions fitted to an empirical distribution (using the least squares fitting algorithm). As expected, at first glance all seem in fair agreement with the data, but there are some deviations. A detailed analysis of these deviations is given in the next section, but for now we can state that for the gamma duration fit, the fitting algorithm's tendency to align model and experimental data on the CDF's steeply ascending flank results in the fit overshooting the empirical CDF at both inflection points. For the two fits to rate data, the deviations between empirical and fitted distributions seem smaller and less consistent.

This impression of the rate distributions fitting our data better is confirmed by quantitative analysis. Figure 4 displays the fit quality of the three distributions, in terms of three different measures. It can be seen that, irrespective of the specific fitting algorithm or goodness measure used, both rate fits perform better than the gamma duration fit, and the gamma rate fit even slightly outperforms the beta' rate fit.

Figure 4 clearly shows how the fit qualities of the three distributions relate to each other, but it does not tell us whether or not particular fits should be rejected. To assess that question, we visualized the fit qualities in a different way in Figure 5, plotting the fraction of fits that would be accepted, as a function of the critical Kolmogorov-Smirnov *p* level that one chooses. An advantage of this presentation method is that it summarizes the acceptance of fits, without
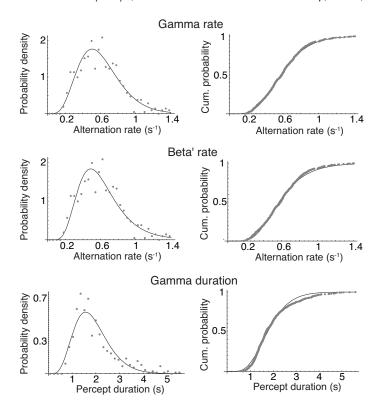
Figure 3. Gamma rate, beta' rate, and gamma duration distributions (drawn lines) fitted to one of our data sets (dots). Probability density functions (left) are shown for illustration, but cumulative distribution functions (right) were used for fitting. The number of points in this data set is 541; the PDF bin size was chosen such that 35 bins span the entire data range. The *pKS* values associated with these fits, purposefully low for illustrating their weaknesses, are 0.06, 0.009, and 0.0002 for the gamma rate, the beta' rate, and the gamma duration distribution, respectively. Note that the top four graphs have alternation rates on the horizontal axis, whereas the bottom two have percept duration.

Figure 4. Quality of the fits to our 192 empirical distributions, measured using the Kolmogorov-Smirnov test (top), the sum of squared residuals *SSE* (middle), and the likelihood *L* (bottom). Each box with whiskers summarizes 192 values, displaying median, 25% and 75% quantile, and extreme values (but note that these fall outside the plot range in some cases). All three figures indicate a low fit quality for the gamma duration distribution compared to the rate distributions, and a slightly better fit for the gamma rate than for the beta' rate distribution.

pinpointing the critical *pKS* level in advance, revealing that both rate distributions have a higher acceptance than the gamma duration distribution at any critical *p* level. For instance, in case one should decide to choose a level of 0.1 (vertical line in Figure 5), fractions of 0.92, 0.86, and 0.69 of the gamma rate, beta' rate, and gamma duration fits pass, respectively.

We can conclude that this analysis positively demonstrates an inferior fit for the gamma duration distribution, compared to the two rate distributions.

Control experiment for motor bias

The shown difference in fit quality, however consistent, is based on fairly subtle differences between distributions (Figure 2). We should therefore be wary of any bias introduced by our experimental design. Particularly, because we relied on key presses to infer our subjects' perceptual state, it is important to rule out the possibility that our results reflect characteristics of the motor system rather than the
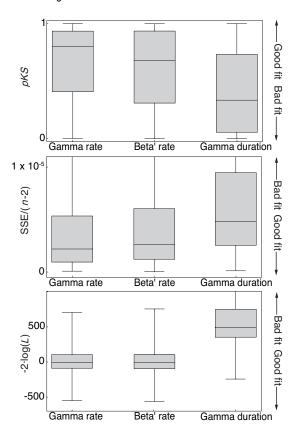
perceptual system. As Figure 2 shows, the main difference between the gamma duration distribution and the other two is there is less probability mass near short percept durations (or near high rates). Because we count the time difference between two successive key presses as one duration, our distributions are biased if the manual response time is not independent of percept duration (e.g., if it takes more time to respond to a short duration than to a long one). It would then be conceivable that subjects do experience percepts that have gamma distributed durations, but that a longer lag in their manual response at short percept durations causes the two rate distributions to fit the reported distributions better after all. To investigate this possibility, we performed a control experiment determining the relation between perceived and reported durations. In this experiment we showed subjects a stimulus that resembled our orthogonal grating stimulus, but that did not cause bistable perception. Instead, "perceptual alternations" were induced by presenting the green and the red set of lines in an alter-

nating fashion on the computer screen, and subjects were asked to press a key when they saw the stimulus change. Importantly, we drew the durations of these successive presentations from a gamma distribution, so if all was well it would be the gamma duration distribution that fitted the subjects' reported distributions best. We found that percept durations shorter than some hundreds of milliseconds did tend to be underrepresented in the reported distributions, and that this did benefit the two rate fits. However, this effect was not nearly strong enough to explain our findings. This is illustrated in the Figure 6 display of the results of one crucial experiment in which we presented six subjects with quite fast alternations (200 durations from a gamma distribution with a mean of 1.2 s and a standard deviation of 0.7 s, comparable to the fastest distributions from the actual experiments). The left panel shows (in the same way as Figure 5) the quality of fits to the distributions we presented to our subjects. Because we drew the presented durations from a gamma distribution, it comes as no surprise that the gamma duration distribution fits best. The more interesting finding comes from the right panel that displays fit quality to the distributions that were inferred from the subjects' key presses. Even though there seems to be some improvement in the rate fits, the gamma duration fit has retained its superior quality, and the gamma rate fit (the best fit to our actual data) fits worst by far. Therefore, our results cannot be explained by a bias introduced by the motor response, but should instead be interpreted in terms of the perceptual process itself.

Interstimulus comparison

Methods

As a first step in comparing the distributions between stimuli, we reexamined the previously determined acceptance levels of our fits, now treating the results of the four stimuli separately. We were particularly alert to differences in the results of binocular rivalry versus other forms of rivalry that might reflect the alleged differences in neural basis. One should keep in mind however that in such an analysis of the *extent* to which the empirical distributions deviate from the fitted ones, one neglects all information on the *manner* in which they deviate. In other words, two empirical distributions might fit a given model equally well, but if one of them consistently undershoots where the other overshoots, this similar fit quality does not reflect a similar shape. We therefore performed an additional analysis in which we compared the distributions' shapes in a more direct way, by looking at the nature of the deviations between empirical and fitted distributions. For this purpose, we calculated for all three fits the mean fit residual as a function of position along the CDF, averaging over all 48 distributions associated with a particular stimulus. Because every CDF spans a different domain of rates (or durations), the data had to be standardized to combine them and cal-
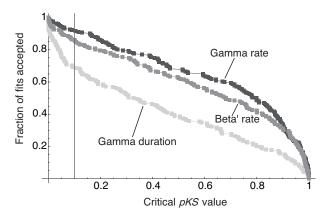

Figure 5. Fit quality as measured by the Kolmogorov-Smirnov *p* value. The fraction of fits that would be accepted is plotted as a function of the critical *pKS* level. Irrespective of the particular critical *p* level that is chosen, the gamma duration distribution, represented here by the lightest shade, clearly fits the data less well than the beta' rate distribution (intermediate shade) and the gamma rate distribution (darkest shade), which fits best. The vertical line at *p* = .1 crosses the three curves at 0.69, 0.86, and 0.92: the fractions of accepted fits when testing at the 10% level.
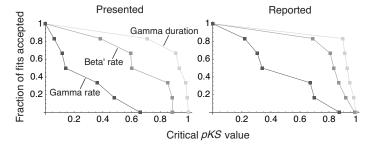

Figure 6. In this control experiment we presented our subjects with a sequence of two alternating images that resembled the two percepts of binocular grating rivalry, and asked them to report every alternation by means of a button press. By drawing the durations between successive alternations from a gamma distribution, we mimicked the time course of a Poisson-driven bistable perception experiment. The left panel shows (in the manner introduced in Figure 5) the quality of the gamma rate, the beta' rate, and the gamma duration fits to the distributions we presented to our subjects in this way. The gamma duration distribution fits best, as was expected, because we drew the presentation durations from a gamma distribution. In the panel on the right, showing fit quality to the distributions of alternations that our subjects *reported*, this superior fit quality is maintained. This experiment indicates that the inferior gamma duration fit in our real experiments is probably not due to a bias introduced by the motor response.

culate this average. A simple but effective way to achieve this is to consider the residuals as a function of the value of the fitted CDF, instead of as a function of the rate (or duration). Our procedure is illustrated in Figure 7.
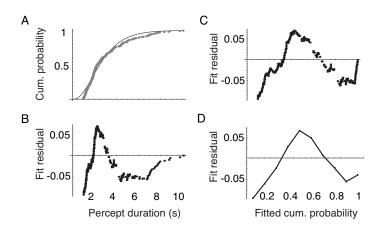
Figure 7. Analyzing fit residuals. To compare the distributions' shapes between stimuli, we analyzed the deviations between our empirical and fitted distributions. A. An example of an empirical CDF of percept durations (gray dots) and the corresponding gamma duration fit (drawn line). B. Residuals of the same fit, as a function of percept duration. The values plotted here are simply the differences between the empirical and the fitted distribution depicted in A. C. For each distribution, panel B may span a different domain of percept durations, preventing us from averaging residuals over distributions. Therefore, in panel C we standardized the *x*-axis in panel B by replacing each *x*-value with the corresponding value of the fitted CDF. Note that in statistics, plots such as these, depicting residuals as a function of cumulative probability, are referred to as "detrended Q-Q plots." D. The same data as in panel C, after averaging over bins spanning 0.1 unit on the *x*-axis. After this final step, residuals could be combined over fits. Note that this step of binning over 0.1 unit slightly decreases the amplitude of the curve.

Results

[Figure 8](#) depicts the acceptance of fits in the same way as [Figure 5](#), except that now the results of the four stimuli were treated separately. This procedure reveals that the isolated results for the bistable slant stimulus (top right) deviate from the general pattern, showing hardly any difference in acceptance between the three distributions. However, the results for the remaining stimuli do conform to the overall finding that the rate distributions outperform the gamma duration distribution. Therefore, contrary to what one might predict, no dichotomy between the binocular rivalry stimuli (grating and house face) and the other stimuli can be made on the basis of this figure. On the contrary, any separation would be between the isolated results for the bistable slant stimulus, and those for the other stimuli.

The results of our second analysis, comparing fit residuals between stimuli, are depicted in [Figure 9](#). This figure agrees with our preliminary impression that gamma duration fits deviate from the empirical distributions in a highly consistent way (bottom left panel; see also figure caption). Because the gamma duration distribution has a fairly bad fit quality, it is not so surprising to see that this pattern of deviations is similar over all four stimuli: This simply
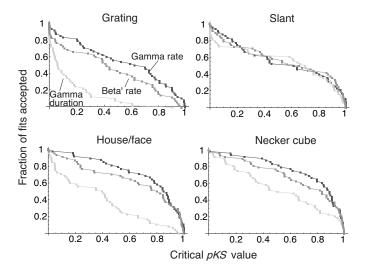

Figure 8. Fit quality as measured by the Kolmogorov-Smirnov *p* value, separated by stimulus. The fraction of fits that would be accepted is plotted as a function of the critical *pKS* level. For grating, house/face, and Necker cube rivalry, the results conform to the overall pattern, with the gamma duration distribution (lightest shade) fitting our data less well than the beta' rate distribution (intermediate shade) and the gamma rate distribution (darkest shade). The results for the bistable slant stimulus form an exception, with no clear difference in quality between any of the three fits. There is however no indication of a dichotomy between the binocular rivalry stimuli (grating and house/face) and the other two. All curves are built up of 48 data points.

means that any subtle differences that may exist between the four groups of distributions are overshadowed by the gamma duration distribution's large structural fit error. In terms of the interstimulus comparison, the more interesting panels are the top two, depicting residuals for the two (better fitting) rate distributions. As is apparent from these panels, the pattern of deviations is comparable for all but the slant stimulus (second lightest shade). Regarding the remaining three stimuli, we can therefore state not only that they are similar in their fit quality to the tested distributions, but also that the minute fit residuals that do exist follow a similar pattern. This again provides no evidence for a dichotomy between the binocular rivalry stimuli and the remaining stimuli, but rather between the slant stimulus and the other three stimuli. On a side note, it is worth mentioning that the pattern of residuals is similar for the gamma rate and the beta' rate distribution.

Discussion

As addressed in the [Introduction](#), we are not the first to show a fairly bad fit quality for the gamma distribution fitted to distributions of percept durations. Previous authors to do so have proposed a number of alternative duration distributions, but none considered taking the recipro-
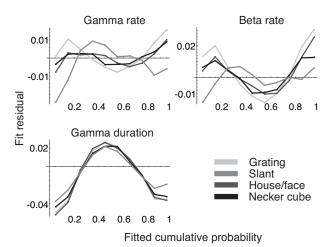
Figure 9. Average residuals of our three fits, separated by stimulus. Each curve represents the binned residuals averaged over 48 distributions, as a function of fitted CDF value (see Figure 7). A positive residual means that the empirical data overshoot the fit; a negative value means the data undershoot the fit. Standard deviations on the points in these graphs (not shown for the sake of clarity) are in the order of 0.02. The clearest panel is the one depicting the results for the gamma duration fits (bottom left). For all four stimuli, the curves show the same pattern of the empirical data initially undershooting the fit, then overshooting it near the mean, and again undershooting it toward the end. The two rate distributions (top panels) give rise to more noisy residuals, as was to be expected in light of their superior fit quality. Still, for both rate fits, the pattern of residuals is fairly consistent for all but the slant stimulus. We can therefore state that, in agreement with our previous analysis, the distributions associated with the bistable slant stimulus are slightly different from those associated with the other stimuli.

cal of their percept durations. This makes it all the more remarkable that we find this simple operation to improve fit quality so greatly. The distributions that previous authors compared to the gamma distribution never showed a notably better fit, except in one particular case. This one case was provided by Zhou et al. (2004), who showed the lognormal distribution to fit their data dramatically better than the gamma distribution. To find out how these findings relate to our results, we performed a lognormal fit to our duration distributions as well, and we can confirm an improved fit quality compared to the gamma duration distribution. However, this improvement was modest, and not as large as what we have shown for the two rate distributions. Using, as we did earlier, the fraction of accepted fits at a 10% significance level as a brief indication of fit quality, the lognormal distribution, with a fraction of 0.78, falls right in between the values of the gamma duration distribution (0.69) and the two rate distributions (0.86 and 0.92). These rate distributions should therefore be preferred to the lognormal distribution when judging merely on the basis of fit quality, but a few additional notes on the log-

normal distribution are required. First of all, one might wonder why Zhou and coworkers found such a large difference in fit quality between the gamma and the lognormal distribution, while we find only a small one. We think a possible explanation for this is that their comparison is biased by their use of an inferior fitting routine for their gamma fits (namely parameter estimates based on their analytical relationship with a distribution's first and second order moment) compared to their lognormal fits (for which they used a more sophisticated, numerical fitting routine). Zhou and colleagues therefore performed their analysis of fit quality on suboptimal gamma fits. Although this might account for the big difference they showed between their gamma and lognormal fits, it cannot explain why their lognormal fits are better than the ones we performed. Therefore, it does appear that their data for some reason conform to the lognormal distribution better than ours. As a second note we think it is worth mentioning that, like the beta' distribution, the lognormal distribution is invariant under inversion, so a lognormal duration fit is equivalent to a lognormal rate fit (also see Introduction).

We consider it beyond the scope of this study to present an exhaustive comparison of fit qualities of the numerous distributions that have over the years been proposed in this context. As mentioned previously, many can be discarded because other studies showed their fit qualities not to exceed that of the gamma duration distribution. One last distribution that we do wish to discuss here however is the "dual Gaussian rate distribution." Because this is a rate distribution of some renown in the field of saccadic search, it seems appropriate to include it here. When using this distribution, which is actually a composite of two separate distributions, in the conventional way, one assumes that a part A of an empirical rate distribution, the part containing low rates, can be described by one truncated Gaussian distribution Φ_A, and the remaining part B by another one Φ_B (Carpenter & Williams, 1995; Reddi & Carpenter, 2000). The raw data are plotted on probability paper, so the distinction between data section A and B can be made by eye, after which the parameters of Φ_A and Φ_B can be estimated (see Figure 10, left panel). If we want to compare the fit quality of the dual Gaussian distribution to that of the ones we investigated, this approach will not do, mainly because it involves visual inspection. Because it does seem that the combination of two Gaussians describes rates of saccadic eye movements well, we adjusted the dual Gaussian distribution to investigate its fit quality to our data on alternation rates. According to this adaptation, the probability density function at rate r is given by

$$f(r \mid \mu_1, \sigma_1, \mu_2, \sigma_2, a) = a \cdot \varphi(r \mid \mu_1, \sigma_1) +$$

$$(1-a) \cdot \varphi(r \mid \mu_2, \sigma_2).$$

Here, φ is a Gaussian distribution. Note that instead of fitting two truncated Gaussians to separate parts of the empirical distribution, we fit a mix of two Gaussians to the
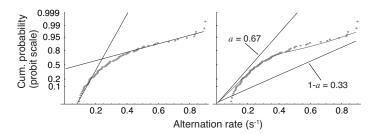
Figure 10. Dual Gaussian fits to one of our distributions. Both plots depict cumulative probability as a function of alternation rate, on probit scale. Left. The conventional method. The data (dots) are divided into two sections, each of which can be described by a Gaussian distribution: the two straight lines. This method does not readily lend itself for a comparison of fit quality. Right. An adapted method, applied to the same data. The two straight lines depict two Gaussian distributions that together comprise the fitted distribution (curved line). The relative contributions of the two Gaussians, determined by weighing factor a, are displayed next to the corresponding lines.

entire distribution, and introduce a free scaling parameter a to ensure a total cumulative probability of 1. This mixed Gaussian distribution was introduced over a century ago (Pearson, 1894), and similar mixed distributions are presently used in numerous fields of research (McLachlan & Peel, 2000). This dual Gaussian model produced excellent fits to our data (all were accepted at a critical *pKS* level of 0.1, after removal of nonconvergent fits), but a direct comparison to the other models is not possible because these contain only two free parameters instead of five (the dual Gaussian distribution is over-parameterized, as witnessed by the convergence problems we experienced). In addition, it is worth mentioning that our adapted dual Gaussian fits often reach quite different results than the conventional ones (see Figure 10). A final conceptual problem one might have with these fits is that both in the original form and in our adaptation, they allow for the occurrence of negative alternation rates.

As discussed in the Introduction, the gamma duration distribution is associated with a so-called "Poisson clock" model. Because we showed a fairly bad fit quality for this distribution, it is important to address a recent study by Murata et al. (2003), presenting data that seem to provide strong support for this model. According to the Poisson model, the gamma duration distribution's shape parameter k reflects the number of Poisson ticks after which a perceptual transition takes place (see Introduction). On the basis of the model, one might therefore predict that this shape parameter should be a natural number. What Murata and coworkers showed is that the gamma duration distributions they fitted to their data indeed had shape parameters that grouped around natural numbers. It would seem that this result cannot be explained unless by accepting the idea of a Poisson clock and the associated gamma duration distribution, but we see two reasons why such a conclusion is pre-

mature. First of all, Murata and coworkers did not support their claim with a statistical analysis. A straightforward way of statistically testing for grouping around natural numbers, using estimated shape parameters from gamma duration fits, would be to subtract the nearest natural number from each of these estimates. This operation would produce a distribution of *residual* shape parameters ranging from –0.5 to 0.5 that should be peaked around 0 in case of natural k values. One can statistically test for the presence of such a peak using a standard test. We have performed such an analysis on our data without finding any evidence for natural shape parameters; however, Monte Carlo simulations show that a data set as large as ours (192 distributions of 243 points on average) does not provide enough statistical power to demonstrate natural k values in this way, even if they are present. This is because the parameter estimates from any fitting procedure have only a limited accuracy, and we do not have enough data to detect a signal in this noise. Similar simulations show that Murata and colleagues' data set (227 distributions of 350 data points) might be just large enough to successfully perform such an analysis, but it should be noted that these simulations were performed without adding any noise.

On top of this statistical issue, we have a second reservation regarding the alleged finding of natural shape parameters. This reservation is centered around possible neural interpretations for the gamma *rate* distribution. At present, we do not have any concrete neural model, but we might speculate that a perceptual alternation might occur as a threshold is reached by some decision signal that is itself the sum of a number of rising signals. If these rising signals would, in analogy to the classic Poisson model, have rates that are drawn from a Poisson distribution, perceptual alternation rates would follow gamma distributions, and these would have natural k values. Clearly, this is only speculation, and Poisson distributed rates are less intuitive than Poisson distributed latencies, but we think mentioning this interpretation is worthwhile because of a particular characteristic of the gamma shape parameter: When a data set reasonably fits both a gamma duration distribution and a gamma rate distribution, the k values of these two distributions are equal. This relation can be proven mathematically[2], and we can confirm it for our fits (when we plotted the k values from the fits to our 192 distributions against each other, linear regression gave $k_{rate} = -0.02 + 1.02k_{duration}$, with an r^2 of 0.998). The important implication of this characteristic is that should one be able to support the Poisson clock model by finding the k values of gamma duration distributions to take on natural values, this would automatically mean the same for a gamma rate model.

The distributions associated with the bistable slant stimulus seem to differ from the ones produced using other stimuli. They are exceptional in that they fit the gamma duration distribution just as well as the two rate distributions, and also in their pattern of fit residuals. Because another notable characteristic of the bistable slant stimulus is

the relatively low rate of perceptual alternations it elicits, it is tempting to associate both features. As previously mentioned, and illustrated by Figure 2, the gamma duration distribution mainly differs from the other two in the sense that there is less probability mass near short percept durations (or near high rates). Our findings might therefore be in agreement with some extension of the classic Poisson model, in which the addition of some low-pass filter results in the Poisson-generated gamma distributions being deformed at shorter percept durations. This idea would explain the improved gamma duration fit for the slower alternating slant stimulus, and seems somewhat in agreement with Figure 9, which shows that the difference in fit residuals between the slant stimulus and the other stimuli, is partly concentrated in the region of short durations (high rates), although also at the other extreme.

Our analyses did not reveal any difference between binocular rivalry and other forms of rivalry. In fact, differences among nonbinocular rivalry stimuli (Necker cube and slant) were larger than between binocular and nonbinocular rivalry stimuli. Should we now conclude that the two phenomena have a shared neural basis? Obviously, this would not be justified. After all, absence of proof is not the same as proof of absence, so there might be differences in aspects of temporal dynamics that we have not addressed. In fact, such differences have been found in the extent to which subjects can voluntarily control the rivalry process (Meng & Tong, 2004; Van Ee et al., 2005). In addition, as Blake (2001) put it ". . . comparable temporal fluctuations may reflect a fundamental property of neural dynamics, but not necessarily a common neural mechanism." Nevertheless, there are two conclusions that we think are justified. First, the claim of similar temporal dynamics of binocular and other rivalry, hitherto based on no more than a broad resemblance of the distributions, still holds after detailed analysis. Our results therefore render it more convincing as one of the pieces of converging evidence linking the two phenomena. Second, any model placing binocular and other rivalry in different brain "modules" should be able to account for the highly similar distributions of alternation rates.

Conclusion

In summary, the prominent position in the literature of the gamma duration distribution is not supported by its fit quality to empirical data. Instead, our data show two straightforward rate distributions, the beta' rate distribution and the gamma rate distribution, to conform to our data much better. This, together with the notion that they form a more direct representation of the neural decision process, makes rate distributions more appropriate than duration distributions for studying the time course of bistable perception. Furthermore, our results demonstrate that the temporal dynamics of binocular rivalry and ambiguous figure perception are similar up to a high level of detail.

Footnotes

[1]The gamma distribution and the beta' distribution are linked by the fact that dividing two gamma distributions of equal scale parameter produces a beta' distribution. The beta distribution (a scaled version of the beta' distribution) previously made its appearance in bistable perception literature when Borsellino et al. (1972) made use of this feature in their analysis of supposedly gamma-distributed percept durations.

[2]Based on the characteristic alluded to in the previous footnote, we can mathematically prove that the shape parameters from our gamma rate fits should be equal to those of our gamma duration fits. This can be understood as follows. If X and Y show gamma distributions with different shape parameters p and q, but equal scale parameter r, then X/Y shows a beta' distribution with parameters p and q (e.g., Pestman, 1998). Now by choosing for X and Y two halves of some set of percept durations that reasonably fits a gamma distribution, we make certain that $p = q$. Consequently, taking either X/Y or Y/X will both yield the same beta' distribution with the two parameters equal. Clearly, if the same data set conforms to a gamma rate distribution, too, we might just as well divide $1/X$ by $1/Y$ to arrive at the same beta' distribution. We can conclude that if some data set shows a fair fit to both a gamma rate distribution and a gamma duration distribution, both have an equal shape parameter.

References

Blake, R. (2001). A primer on binocular rivalry, including current controversies. *Brain and Mind, 2*(1), 5-38.

Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience, 3*(1), 13-21. [PubMed]

Borsellino, A., De Marco, A., Allazetta, A., Rinesi, S., & Bartolini, B. (1972). Reversal time distribution in the perception of visual ambiguous stimuli. *Kybernetik, 10*(3), 139-144. [PubMed]

Carpenter, R. H. S. (1981). Oculomotor procrastination. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and visual perception*. Hillsdale, NJ: Erlbaum.

Carpenter, R. H. S., & Williams, M. L. L. (1995). Neural computation of log likelihood in control of saccadic eye movements. *Nature, 377*(6544), 59-62. [PubMed]

Carter, O. L., & Pettigrew, J. D. (2003). A common oscillator for perceptual rivalries? *Perception, 32*(3), 295-305. [PubMed]

Cogan, R. (1973). Distributions of durations of perception in the binocular rivalry of contours. *Journal of General Psychology*, *89*(2d Half), 297-304. [PubMed]

De Marco, A., Penengo, P., & Trabucco, A. (1977). Stochastic models and fluctuations in reversal time of ambiguous figures. *Perception*, *6*(6), 645-656. [PubMed]

Kovacs, I., Papathomas, T. V., Yang, M., & Feher, A. (1996). When the brain changes its mind: Interocular grouping during binocular rivalry. *Proceedings of the National Academy of Sciences U.S.A.*, *93*(26), 15508-15511. [PubMed][Article]

Lehky, S. R. (1995). Binocular rivalry is not chaotic. *Proceedings of the Royal Society London B, 259*(1354), 71-76. [PubMed]

Leopold, D. A., & Logothetis, N. K. (1999). Multistable phenomena: Changing views in perception. *Trends in Cognitive Sciences*, *3*(7), 254-264. [PubMed]

Levelt, W. J. (1967). Note on the distribution of dominance times in binocular rivalry. *British Journal of Psychology*, *58*(1), 143-145. [PubMed]

Logothetis, N. K. (1998). Single units and conscious vision. *Philosophical Transactions of the Royal Society of London B, 353*(1377), 1801-1818. [PubMed]

Logothetis, N. K., Leopold, D. A., & Sheinberg, D. L. (1996). What is rivalling during binocular rivalry? *Nature*, *380*(6575), 621-624. [PubMed]

Lumer, E. D., Friston, K. J., & Rees, G. (1998). Neural correlates of perceptual rivalry in the human brain. *Science, 280*(5371), 1930-1934. [PubMed]

Meng, M., & Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision, 4*(7), 539-551, http://journalofvision.org/4/7/2/, doi:10.1167/4.7.2. [PubMed][Article]

McLachlan, G. J., & Peel, D. (2000). *Finite mixture models*. New York: John Wiley & Sons.

Murata, T., Matsui, N., Miyauchi, S., Kakita, Y., & Yanagida, T. (2003). Discrete stochastic process underlying perceptual rivalry. *Neuroreport*, *14*(10), 1347-1352. [PubMed]

Pearson, K. (1894). Contributions to the mathematical theory of evolution. *Philosophical Transactions of the Royal Society of London A, 185*, 71-110.

Pestman, W. R. (1998). *Mathematical statistics: An introduction*. Berlin & New York: De Gruyter.

Reddi, B. A. G., & Carpenter, R. H. S. (2000). The influence of urgency on decision time. *Nature Neuroscience*, *3*(8), 827-830. [PubMed]

Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neuroscience*, *27*(3), 161-168. [PubMed]

Tong, F. (2001). Competing theories of binocular rivalry: A possible resolution. *Brain and Mind, 2*(1), 55-83.

Tong, F., Nakayama, K., Vaughan, J. T., & Kanwisher, N. (1998). Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron*, *21*(4), 753-759. [PubMed]

Toppino, T. C. (2003). Reversible-figure perception: Mechanisms of intentional control. *Perception & Psychophysics, 65*(8), 1285-1295. [PubMed]

Van Dam, L. C., & Van Ee, R. (in press). The role of (micro)saccades and blinks in perceptual bi-stability from slant rivalry. *Vision Research.*

Van den Berg, A. V., & Van Loon, E. M. (in press). An invariant for timing of saccades during visual search. *Vision Research.*

Van Ee, R. (2005). Dynamics of perceptual bi-stability for stereoscopic slant rivalry and a comparison with grating, house-face, and Necker cube rivalry. *Vision Research, 45*(1), 29-40. [PubMed]

Van Ee, R., Van Dam, L. C., & Erkelens, C. J. (2002). Bistability in perceived slant when binocular disparity and monocular perspective specify different slants. *Journal of Vision, 2*(9), 597-607, http://journalofvision.org/2/9/2/, doi:10.1167/2.9.2. [PubMed][Article]

Van Ee, R., Van Dam, L. C. J., & Brouwer, G. J. (2005). Voluntary control and the dynamics of perceptual bistability. *Vision Research, 45*(1), 41-55. [PubMed]

Van Loon, E. M., Hooge, I. T., & Van den Berg, A. V. (2002). The timing of sequences of saccades in visual search. *Proceedings of the Royal Society of London B, 269*(1500), 1571-1579. [PubMed]

Walker, P. (1975). Stochastic properties of binocular rivalry alternations. *Perception & Psychophysics*, *18*(6).

Zhou, Y. H., Gao, J. B., White, K. D., Merk, I., & Yao, K. (2004). Perceptual dominance time distributions in multistable visual perception. *Biological Cybernetics*, *90*(4), 256-263. [PubMed]