Measuring perception without introspection

David Alais  
School of Psychology, University of Sydney, Australia

Mirjam Keetels  
Department of Cognitive Neuropsychology, Tilburg University, Netherlands

Alan W. Freeman  
Sydney Medical School, University of Sydney, Australia

Binocular rivalry, the perceptual alternation between incompatible monocular stimuli, is conventionally measured by asking the subject which percept is currently visible. This is problematic because the response is unverifiable, open to response bias, and falsely assumes that the perceptual experience is binary. We overcame these limitations in a new approach that does not require subjective reporting of perceptual state. A brief test stimulus was added to one eye’s inducing stimulus at random times and contrasts. The test was presented at one of two spatial locations, the subject indicated which alternative had been shown, and the correctness of the response was recorded as a function of test contrast. Given the random timing of the test stimulus, it was sometimes delivered when the tested eye was dominant and, at other times, suppressed. Accordingly, the psychometric function recorded during rivalry should be a mixture of the dominance and suppression forms of the function. This was indeed the case: The probability of a correct response during rivalry was significantly less than that obtained with a binocularly congruent stimulus. The psychometric function during rivalry was well modeled as a weighted sum of the congruence curve with an assumed suppression curve. Optimal fitting provided estimates of both suppression depth and percept predominance that corresponded closely with estimates obtained with the conventional method. We have therefore characterized rivalry without the uncertainties introduced by the subject’s perceptual report. This provides a model that may be applicable to the broader field of perceptual ambiguity.

Introduction

Visual perception is easy to experience but difficult to measure. The psychophysical tools available for measuring perception have evolved steadily since Fechner’s seminal work (Fechner, 1860/1966; Kingdom & Prins, 2010; Thurstone, 1959) but remain somewhat primitive. One useful approach is to use ambiguous stimuli, that is, patterns that support two interpretations. Viewing ambiguous stimuli evokes a bistable percept, with an irregular alternation between the two competing percepts: Conscious awareness alternates even though the sensory inputs are constant (Blake & Logothetis, 2002). In binocular rivalry, for example, the stimulus presented to one eye is incompatible with that presented to the other eye and perception switches between the monocular stimuli every few seconds in a never-ending cycle (Levelt, 1966). The seen and unseen stimuli are termed dominant and suppressed, respectively. The loss of visibility during suppression is not absolute, and visual sensitivity in the suppressed eye can be measured using a monocular test stimulus to compare sensitivity during suppression relative to dominance. The drop in sensitivity when the tested eye’s image is invisible (Alais, Cass, O’Shea, & Blake, 2010; Fox & Check, 1966; Nguyen, Freeman, & Alais, 2003) is called suppression depth and provides a useful metric for quantifying how much attenuation is needed to suppress an image from conscious awareness.

The standard method for measuring suppression depth requires the subject to continuously report which rivalry stimulus is currently visible so that the experimenter can deliver the test stimulus in the appropriate rivalry state—either dominance or suppression, depending on condition. When in the desired perceptual state, a test stimulus is triggered and the subject makes a forced-choice response about the stimulus (typically a contrast increment threshold task). This method of measuring suppression depth brings a host of problems. First, the experiment is complicated by a dual-task design, as the percept tracking is required in parallel with the forced-choice sensitivity task. Second, transitions between dominance and
suppression are often difficult to define (Blake, O’Shea, & Mueller, 1992), and percept tracking therefore introduces a degree of response bias. The transition period may elicit mixed percepts in which parts of each eye’s stimulus are visible, further complicating percept categorization. A third problem is that attention to the percept alters the dynamics of the system being measured (Lack, 1974; Meng & Tong, 2004; Paffen, Alais, & Verstraten, 2006).

Together, these problems add unnecessary variability to the data and even question the reliability of sensitivity measures in rivalry. Here we introduce a new method for measuring binocular rivalry suppression that overcomes these limitations: The subject responds to a test stimulus delivered at random times and with random contrast. Our new approach simplifies the traditional dual-task design to a single forced-choice task with a more objective response. Moreover, it allows us to measure visual sensitivity during rivalry dominance and suppression without asking subjects to report their subjective perceptual fluctuations. This overcomes the fundamental subjectivity of the standard approach and also avoids the problem of response bias. The work described here has been previously published in abstract form (Alais, Keetels, & Freeman, 2011).

Methods

Subjects

Seven subjects participated in these experiments (two male, five female; age range 26–45 years). All had normal or corrected-to-normal visual acuity and normal stereovision. Two of the subjects (DA and MK) were authors, while the remaining five were not aware of the experiment’s aims. The Human Ethics Committee of the University of Sydney approved the study.
experiment the radial grating had a contrast of 0.55 while the concentric grating had a four-fold lower contrast of 0.14. For the congruent control condition, both monocular contrasts were 0.14 (to prevent rivalry). In both experiments, contrast sensitivity was measured by adding a small contrast increment to the concentric grating, either in the upper or lower half. This test stimulus, shown in Figure 1b, was a Gaussian function of space and time where the standard deviations were 0.8° for horizontal distance, 0.4° for vertical distance, and 55 ms for time. The eye receiving the concentric grating (and thus the test stimulus) was the behaviorally dominant eye, as determined by a sighting test (Porac & Coren, 1976).

Procedure

Subjects viewed the rivalry stimuli through the stereoscope, using a chinrest to minimize head movements, in a dimly lit and sound-attenuated booth. For experiments using the random-sampling method, the interval between the start of a trial and the test stimulus was 2 s on average; the interval’s duration was drawn from a uniform probability density from 1.5 to 2.5 s. Subjects used the keyboard to indicate the apparent location of the test stimulus, and the next trial began with the subject’s response. A pilot study was used to estimate the contrast threshold for the test stimulus, and then the contrast increment was varied randomly from trial to trial according to a Gaussian probability density centered near this threshold. Because some tests were inevitably below threshold, or occurred when the tested eye was suppressed, a brief tone pip sounded after each test to indicate that a test had just been presented and a response was required. Each session comprised 200 test stimuli and a total of 1,000 tests were presented.
for each experimental condition (rivalry and control); rivalry trials were interleaved with control trials in blocks of 10 trials.

Rivalry was also measured using the conventional self-triggered method. This involved discrete trials in which subjects waited for the concentric circles to become dominant or suppressed, depending on the condition. They then used a key-press to trigger the test stimulus and indicated their response (upper or lower). Test contrast was varied using an adaptive staircase and six QUEST staircases of 40 trials each were run per condition. We also used a single-task version of the conventional method: The subject continuously signaled his or her percept, but no test stimulus was delivered. The purpose of this experiment was to measure predominance, that is, the proportion of total viewing time (excluding mixed percepts) that the concentric circles were visible. In this case each subject completed $3 \times 3$ min blocks of tracking rivalry alternations.

Analysis

Psychometric functions were constructed in three steps. First, all test contrasts delivered in a specific condition were binned into a frequency histogram. Second, using the same bin boundaries, test contrasts yielding correct responses were binned into a second histogram. The second histogram was then divided by the first.

Model

Let $c$ be $\log_{10}$ contrast, $g$ be the curve fitted to the monocular or congruent data, and $p$ be the probability of a correct choice. The model fitted to the rivalry data is then

$$p(c) = wg(c, \mu) + (1 - w)g(c, \mu - s)$$

where

$$g(c, \mu) = .5 + \frac{5}{\sqrt{2\pi} \sigma} \int_{-\infty}^{c} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

$\mu, \sigma$ = mean, standard deviation and the parameters of interest are

$s$ = suppression depth (log units)

$w$ = predominance of tested eye's rivalry stimulus

The parameter $w$ takes values between 0 and 1.
shifted copy of the monocular curve, an assumption the suppression curve we assumed that it was a laterally rivalry data fall significantly below the monocular data, (Blake & Camisa, 1978; Fox & Check, 1966). olds measured in the dominance phase of rivalry monocular viewing are indistinguishable from thresh- Figure 3b. In our new procedure, illustrated on the left in dominance and suppression curves, as shown in Figure 2a, the test is delivered at random times: equal contrasts (0.55 vs. 0.55) and unequal contrasts (0.14 vs. 0.55). Error bars give 95% confidence intervals.

Results

The test stimulus, illustrated in Figure 1b, was a contrast increment in the upper or lower half of one eye’s rivalry stimulus. The subject’s task was to indicate test location. Figure 2a represents a conventional rivalry experiment, in which the subject waits until the required rivalry stimulus is dominant, triggers a test, and responds to it. This yields two psychometric curves, one measured during dominance and the other during suppression as shown in Figure 2b. In our new procedure, illustrated on the left in Figure 3a, the test is delivered at random times: sometimes during dominance, sometimes during suppression, and in the remaining cases during transitional or mixed states. We therefore expected the resulting psychometric curve to be a weighted sum of the dominance and suppression curves, as shown in Figure 3b. To test this idea we also measured the psychometric function in the absence of rivalry, with one eye’s rivalry stimulus blanked (Figure 3a, right). Previous work has shown that test thresholds during monocular viewing are indistinguishable from thresholds measured in the dominance phase of rivalry (Blake & Camisa, 1978; Fox & Check, 1966). The results, in Figure 4, match expectations: The rivalry data fall significantly below the monocular data, as shown by the 95% confidence intervals. To recover the suppression curve we assumed that it was a laterally shifted copy of the monocular curve, an assumption backed up by previous work (Green & Luce, 1975; Nachmias, 1981; Roufs, 1974) and by our own results (Figure 6). The suppression curve, shown dashed, was located so that a weighted sum of it and the monocular curve provided an optimal fit to the rivalry curve. Both models in Figure 4a fit the data well in that each explains at least 96% of the variance. The lateral shift between the monocular and suppression curves provides a measure of suppression depth. Figure 5 shows suppression depth measured conventionally and by the random sampling method. Mean suppression depths match well across the two types of experiment, $t(3) = 0.33, p = 0.77$, but the estimates are less variable with the new method.

The model fitted to the rivalry data has two parameters. One is suppression depth, and the other is the relative weighting of the dominance and suppression functions. What does the second parameter represent? Because the test stimulus is delivered at random times, it will sample the dominance and suppression states in proportion to their durations. The weight parameter therefore estimates the fraction of viewing time for which the tested eye’s rivalry stimulus is the dominant percept, also known as predominance. Predominance can be manipulated by adjusting the relative contrasts of the monocular rivalry stimuli (Levelt, 1966; Mamassian & Goutcher, 2005), as illustrated in Figure 6a. The psychometric function was measured after lowering the contrast of the rivalry stimulus for the tested eye to a quarter of the contrast for the fellow eye. Compared with Figure 4, the rivalry curve is shifted further right of the red curve, and modeling indicates that the best-fitting suppression curve now lies close to the rivalry curve. Predominance of the tested eye’s stimulus was estimated from the weight assigned to the dominance curve during model fitting. Figure 7 compares predominance values estimated in this way with the values obtained by the conventional method. The conventional method in this case was a single-task version in which subjects signal their percepts but in which no test stimulus was delivered. The data show a close correspondence between the two experiments, $t(3) = 1.1, p = 0.35$ for equal contrasts; $t(3) = 0.75, p = 0.51$ for unequal contrasts.

We can now make a stronger comparison between the random sampling and conventional methods by asking two questions. First, do the indices of rivalry measured with the random-sampling method match those for the conventional method? Second, is the new method more efficient than the older one? To answer the first question, we combined the results of all the experiments in an analysis of variance. The analysis had three predictor variables—monocular contrast difference (same or different), experimental method (conventional or random sampling), and subject (seven.
in total)—and two response variables (suppression depth and predominance). Using a 5% significance level, suppression depth did not depend on the experimental method, $F(1, 6) = 0.16$, $p = 0.70$, and nor did predominance, $F(1, 13) = 0.60$, $p = 0.45$.

Second, does the random-sampling method have efficiency advantages over its predecessor? To resolve this issue we calculated the variance explained by the fitted psychometric functions versus data collection time. The result, in Figure 8, shows that for any given percentage of variance explained, data collection time is shorter for the random-sampling method. Further, at a data collection time of 30 or more minutes, the new method explains almost all, 95%, of the variance. The conventional method, by contrast, explains at most 70% of the variance. The residual variance in this case could well be due to such variable factors as the subject’s perceptual criterion. The psychometric functions for the conventional method were obtained with a dual-task (perceptual judgment and test response) experiment. Estimating predominance requires an additional (single-task) experiment and therefore extra data collection time; this further reduces the efficiency of the conventional method. It is clear, therefore, that the random-sampling method of measuring binocular rivalry is more efficient that the conventional method.

### Discussion

Our new method for measuring binocular rivalry suppression depth and predominance has several advantages over conventional procedures. First, it is less demanding because subjects perform only one task—responding to the test stimulus. As a result, mean trial duration was shortened from 6 s in the conventional experiment to 3 s for the random-sampling experiment. Second, it removes the criterion-dependent variability of judging perceptual state. Third, it overcomes the circular confound that monitoring perceptual state alters the rivalry dynamics being measured (Lack, 1974; Meng & Tong, 2004; Paffen et al., 2006). Fourth, the conventional method requires two types of experiment: a dual-task version to measure suppression depth and a single-task version to measure predominance. The random-sampling method, by contrast, requires a single experimental design. The new method is therefore simpler and better controlled, improving the reliability of the data.

Most importantly, our new approach has the advantage of measuring rivalry in its free-running state, rather than artificially dividing it into periods of dominance and suppression. Previous studies have generally discarded periods of mixed percept, which may amount to as much as 60% of viewing time (Blake et al., 1992). By probing continuously, our method samples rivalry during dominance, suppression, and mixed states, and therefore better characterizes the underlying processes. This result can be stated in graphical terms. Figure 4a shows curves for the conventional states, monocular and suppression. We obtained these curves in order to compare our data with conventional studies. The main result of our study, however, is the rivalry curve because that is the one that shows perception in its free-running state.

Recent work has quantified the transitional state between one percept and the other in binocular rivalry. Fahle, Stemmler, and Spang (2011) equipped their subjects with a joystick with continuously variable output to indicate their current percept and any intermediate states. The authors showed that the transition between the two percepts required an average of 0.9 s in both directions. Transitions were much faster when viewing was binocularly congruent and one stimulus was physically swapped for the other, showing that the slow transition during rivalry was perceptual rather than in the motor response. Naber, Frässel, and Einhäuser (2011), also using a joystick response, showed that subjects spend significant periods of rivalry time stuck between the two percepts.

A number of previous studies have attempted a more objective measurement of perceptual status during binocular rivalry (see for example Fahle et al., 2011;
Fox, Todd, & Bettinger, 1975; Frässle, Sommer, Jansen, Naber, & Einhäuser, 2014; Naber et al., 2011). Two of these studies (Fahle et al., 2011; Naber et al., 2011) used monocular stimuli of differing luminance and showed that pupil size decreased shortly before the stimulus of higher luminance became perceptually dominant. Naber et al. (2011) also used a horizontally drifting grating as one monocular stimulus and a grating moving in the opposite direction for the other stimulus. The slow phase of the resulting optokinetic nystagmus provided a good predictor of the dominant percept. Indeed, the authors found a strong correlation between the nystagmus velocity and the perceptual report as measured by a joystick response. The correlation coefficient peaked at 0.7, for nystagmus measurements preceding the perceptual report by 0.8 s. Pupillary and nystagmus measurements have the advantage of a continuous readout of perceptual status. While our method cannot provide this continuous report, it has other advantages. First, there is no need for measurement of oculomotor responses. And second, unlike the previous studies, our method does not require the monocular stimuli to differ in luminance or motion.

Our aim of measuring perception more objectively has important precedents using a number of methodologies. Magnetic resonance imaging of neural activity has been used to predict the perceptual responses of human subjects to binocular rivalry and motion stimuli (Brouwer & van Ee, 2007; Haynes & Rees, 2005; Serences & Boynton, 2007). Single-unit neural studies have used choice probability to quantify the extent to which the response of a single neuron can predict an animal’s behavior (Britten, Newsome, Shadlen, Celebrini, & Movshon, 1996). Our approach complements these studies and has the advantages that it minimizes the demands on the subject and is not invasive.

The methodology we have described here has potential applications beyond the field of ambiguous perception. One example is the currently active field of visual crowding in which a peripheral target becomes perceptually suppressed when embedded within a surrounding array of elements (Whitney & Levi, 2011). Our method could be used to validate the visual sensitivity loss associated with the phenomenal disappearance of the target without requiring subjective perceptual reports. Another example is visual attention, in which subjects are asked to shift the focus of their cognitive processing between differing objects and locations in the visual scene (Chun, Golomb, & Turk-Browne, 2011). Incomplete or slow attention shifts would add uncontrolled variance to the experiment and our method could be used to objectively verify that the shift has been accomplished.

Conclusions

Binocular rivalry suppression and predominance can be measured by delivering test stimuli at random times and contrast during both rivalry and binocular congruence. This prevents the requirement that the subject make a perceptual judgment, thereby simplifying the subject’s task and removing an uncontrolled source of variance.

Keywords: binocular rivalry, suppression depth, predominance

Acknowledgments

This research was supported by grant DP130102336 from the Australian Research Council.

Commercial relationships: none.
Corresponding author: Alan Freeman.
Email: Alan.Freeman@sydney.edu.au.
Address: Sydney Medical School, University of Sydney, Australia.

References


