Available online at www.sciencedirect.com

Vision Research 46 (2006) 3393–3402

Vision Research

www.elsevier.com/locate/visres

Endogenous influences on perceptual bistability depend on exogenous stimulus characteristics

Gijs Joost Brouwer, Raymond van Ee *

Helmholtz Institute, Utrecht University, The Netherlands

Received 13 December 2005; received in revised form 7 March 2006

Abstract

We investigated the influence of changing physical parameters and task on bistable perception of an ambiguously rotating sphere (SFM). Increasing dot-density and velocity decreased the duration of perceptual phases during both passive viewing and voluntary control exertion. Our main finding is that voluntary control of perception depends on the physical parameters constituting the stimulus. This dependency places important constraints on the mechanisms mediating voluntary control as these mechanisms cannot operate independently of stimulus characteristics. In addition, local asymmetries in dot-densities can trigger alternations towards the most salient direction, which is not necessarily associated with largest number of dots: competition between perceptual interpretations during SFM appears to occur between surface-based representations rather than between individual elements. Finally, we show that voluntary control remains effective, even when attentive tracking of individual stimulus elements is no longer possible.

Keywords: Bistability; Structure-from-motion; Psychophysics; Voluntary control

1. Introduction

Under certain conditions, the brain is confronted with visual information that is ambiguous; under such conditions, subjects experience only one interpretation at any moment and at seemingly random times, one perceptual state is replaced by its counterpart (Blake & Logothetis, 2002; Parker & Krug, 2003). Ambiguous stimuli have been investigated extensively using psychophysics (e.g., Levelt, 1965) and using both neurophysiology (Leopold & Logothetis, 1996; Logothetis & Schall, 1989) and functional imaging or fMRI (e.g., Brouwer, van Ee, & Schwarzbach, 2005; Kleinschmidt, Buchel, Zeki, & Frackowiak, 1998; Lumer, Friston, & Rees, 1998; Polonsky, Blake, Braun, & Heeger, 2000; Tong, Nakayama, Vaughan, & Kanwisher, 1998). One question is how the temporal dynamics of bistable perception depend on the physical characteristics of the ambiguous sensory input. Second, to what extent are we able to change the content of our perceptual experience by shear mental effort and how does this voluntary control depend on the physical characteristics of sensory input? If a dependency is found, it could shed light on how and where these processes interact at a neural level. Few studies have systematically investigated an observer's ability to voluntarily change perception, showing that it is possible in a number of distinct ambiguous stimuli, (e.g., Hol, Koene, & van Ee, 2003; Lack, 1978; Meng & Tong, 2004; Peterson, 1986; Suzuki & Peterson, 2000; Toppino, 2003; van Ee, van Dam, & Brouwer, 2005).[1] For example, voluntary control has been reported to have a multiplicative effect on bistable perception: when stimuli were physically biased towards one perceptual interpretation (by changing stimulus eccentricity), observers' intentional efforts to see that particular interpretation were more effective (Suzuki & Peterson, 2000).

---

* Corresponding author. Fax: +31 3025 22664.
E-mail address: r.vanee@phys.uu.nl (R. van Ee).
URL: www.phys.uu.nl/~vanee/ (R. van Ee).

[1] Note added after manuscript submission: Recently, it has also been reported that attending to stimulus features during binocular rivalry can prolong dominance durations (Chong, Tadin, & Blake, 2005).

0042-6989/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2006.03.016

A powerful example of perceptual bistability is the ambiguously rotating sphere (Fig. 1): dots following a path on an imaginary surface can evoke a strong percept of that surface in depth, a phenomenon referred to as structure-from-motion or SFM (e.g., Andersen & Bradley, 1998; Hol et al., 2003; Miles, 1931; Treue, Husain, & Andersen, 1991; Treue, Andersen, Ando, & Hildreth, 1995; Wallach & O'Connell, 1953). When all depth cues are removed from a sphere rotating about the vertical axis, its rotation becomes ambiguous: subjects perceive its front surface to be rotating either to the left (clockwise or *CW*) or to the right (counterclockwise or *CCW*). Neurophysiological studies using bistable SFM have reported a correlation between the activity of MT neurons and perceived rotational direction (Bradley, Qian, & Andersen, 1998; Dodd, Krug, Cumming, & Parker, 2001).

In this study, we use ambiguously rotating spheres to investigate the relationship between physical characteristics constituting the stimulus, the ability of observers to voluntarily control perception and the temporal dynamics of perceptual bistability in a series of three different experiments. First, we determined how average perceptual durations vary with both dot-density and velocity under both natural and voluntary control conditions. Second, we explored whether the bistability of the SFM stimulus can be explained in terms of competition between either stimulus elements or surface-based representations, as it has been shown that these surface-based representations play a vital role in depth perception (He & Nakayama, 1992, 1994a), motion perception (He & Nakayama, 1994a, 1994b) and structure-from-motion (Hildreth, Ando, Andersen, & Treue, 1995; Treue et al., 1995). Using spatial reverse correlation (Neri, Parker, & Blakemore, 1999; Neri & Heeger, 2002), introducing local asymmetries in dots moving in CW and CCW directions, we demonstrate that perceived rotation is determined by the saliency of motion directions and not by differences in the balance of dots moving in CW and CCW directions. This supports the notion that the visual system uses surface-based representations and that these representations are involved in the competitive mechanism underlying bistable perception. Third, we investigated a possible mechanism mediating voluntary control by examining its success during both static and dynamic stimuli. Our results indicate that voluntary control remains successful, even when individual dots can no longer be tracked attentively. Finally, eye movement recordings excluded these as a possible confound.

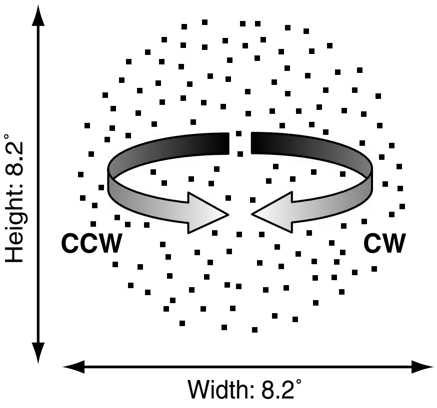

Fig. 1. Ambiguously rotating sphere. Experimental stimuli (8.2° × 8.2°), containing white dots on a black background. Spheres rotated about the vertical axis and were perceived to be rotating either clockwise (CW, from a top-view perspective) or counterclockwise (CCW).

2. Methods

2.1. Visual stimuli

The ambiguous spheres (created using custom software) were presented on a LaCie monitor (1600 × 1200 pixels, 100 Hz, distance 60 cm.). Head movements were constrained using a chin-rest. We used a constant stimulus size (8.2°) with single white dots 5.8′ in width and height. All stimuli contained a central fixation dot of 11.7′. For Experiment 1, we varied dot-density (100, 300, 500, 700 or 900 dots) and angular velocity (16, 32, 48, 64 and 80°/s). For Experiment 2 we used a sphere of 500 dots containing a patch of increased dot-density (high-density patch) or no dots at all (gap-patch). For the high-density patch condition, we added 100 dots in a circular patch on the sphere. For the gap-patch condition we removed all dots within a similar circular patch. Patch location was chosen randomly every trial. Since the patch followed the curvature of this sphere, patch-size was not constant. Maximum patch diameter (occurring when the patch was exactly on the fixation dot) was 2°. For Experiment 3 we used both 'static' (unlimited lifetime of the dots) and 'dynamic' spheres (each dot had a lifetime of 250 m before it was relocated at a random position on the sphere). Note that 'static' here refers to the lifetime of the dots and not to their movement: all dots rotated with an angular velocity of 48°/s.

2.2. Procedure for the psychophysical experiments

In Experiment 1 subjects ($n = 6$) viewed stimuli for 60 s. During *natural viewing*, subjects were instructed to view the stimulus passively while fixating on the fixation dot and report alternations using two buttons to indicate changes the direction of the front surface. Subjects were instructed to press one button when the front surface reversed from a rightward to a leftward direction (*clockwise or CW*) and to press another button when the opposite occurred (*counterclockwise or CCW*). Although the leftward and rightward motion of the front surface correspond to clockwise and counterclockwise rotations of the sphere, respectively, two additional percepts are possible: subjects can also perceive the stimulus as two convex surfaces or two concave surfaces (Hol et al., 2003). However, subjects do still perceive one surface to be in front of the other. Therefore, we instructed subjects to report the direction of the front surface. Twenty-five stimuli were presented four times, with each stimulus containing a unique combination of density and velocity. During the *voluntary control condition*, subjects were instructed to switch their perceptual state as quickly as possible, while keeping strict fixation. Experimental stimuli were identical to the *natural condition*. During Experiment 2, each trial (30 s) consisted of a sphere with a newly randomized location of either the gap-patch or the high-density patch. Each run consisted of 36 stimuli and subjects ($n = 3$) performed four runs each (two high-density patch runs, two gap-patch runs). In Experiment 3, subjects ($n = 6$) viewed stimuli during 60 s trials. During each trial subjects either viewed the stimulus passively or attempted to switch as fast as possible between perceptual states. Spheres were either dynamic (with dots having a limited lifetime) or static (unlimited life time). Each particular combination of task (passive/voluntary) and sphere (static/dynamic) was repeated four times during one run and all six subjects performed two runs each.

2.3. Analysis

For Experiment 1 we calculated perceptual durations for each density and velocity for both conditions (natural/voluntary control). To compare the effect of stimulus parameters (velocity/density) on mean durations, we analyzed each condition (natural/voluntary control) separately as there is no a priori reason to assume that these effects must be similar. Combining the data from both conditions, we tested for condition-specific effects and interactions between condition on the one hand and stimulus parameters on the other using a N-way analysis of variance. Finally, as an interaction effect between condition and stimulus parameters does not guarantee that voluntary control depends directly on stimulus parameters (see results), we analyzed both the difference and ratio between the natural and voluntary control condition for both stimulus parameters using a two-way analysis of variance. In a separate analysis, we fitted the durations to a two-parameter gamma distribution

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

By definition, $\Gamma(n)$ is the canonical continuous extension of $(n-1)!$, which is only defined for natural n. k and λ are usually referred to as the shape and scale parameters, respectively. The shape parameter characterizes the skewness of the distribution, while the scale parameter scales the distribution along the abscissa. The quality of the fit was quantified through the Kolmogorov–Smirnov test, which determines the deviation between the cumulative distribution function (CDF) of the raw data and the fit. We use the CDF, rather than the PDF (probability distribution function) to prevent arbitrary binning of the data. Any fit with a Kolmogorov–Smirnov value greater than 0.05 is considered to be of high quality (see for a detailed description of the fitting procedure: Brascamp, van Ee, Pestman, & van den Berg, 2005). A paired-sample t test was used to estimate the change in fitted parameters as a function of task (natural and voluntary control), a two-way analysis of variance was used to estimate changes as a function of density and velocity. For Experiment 2 we used spatial reverse correlation (see below). For Experiment 3 we used durations in each condition (natural/voluntary control × static/dynamic spheres) and assessed the differences between conditions using a two-way analysis of variance.

2.4. Reverse correlation

We reconstructed the location of the patch (high-*density or gap*) on the sphere at the time of each alternation. Although there is a delay between the perceptual alternation and the button press indicating this alternation, we argue that this delay, due to the relatively low angular velocity of the sphere and therefore its stimulus elements (dots) cause only a minor shift in the dots relative to their actual location at the time of the perceptual alternation. We then identified those parts of the patch moving CW from the parts of the patch moving CCW (*Note that if the patch is close to the fixation point, it only contains one motion direction but if the patch is on an edge, part of it might move CW while the remaining part moves CCW*). At the location of those parts of the patch that were moving in the same direction as the newly perceived rotation of the sphere we added one to a summation image associated with the direction of the alternation (CW/CCW), for those parts that were associated with the opposite direction we subtracted one from this same summation image. After summation, the image was normalized by dividing it by the number of alternations. Positive values at a particular location indicate an increased probability that part of the patch was moving in same direction as the newly perceived rotation at that particular location, negative values indicate increased probability of movement in the opposite direction as the newly perceived rotation. After this, we test the statistical significance of each location using a t test.

2.5. Eye movements

Eye movements were recorded using a head mounted infrared camera based eye tracker (EyeLink II, sample rate 250 Hz) under three conditions: (1) passive viewing under strict fixation, (2) voluntary control under strict fixation and (3) voluntary control with eye movements allowed. Stimuli were identical to the psychophysical experiment (500 dots, 32°/s angular velocity). Each condition consisted of six 210-second trials. Eye movement data were preprocessed by removing blinks and detecting saccades by identifying samples in which the velocity exceeded and returned below 18°/s. Smooth pursuit eye movements were identified as samples in which the velocity exceeded and returned to 2°/s and were required to last at least 80 m. An analysis of variance was used to compare the number of smooth pursuit eye movements, saccades and mean perceptual durations between the three conditions.

3. Results

3.1. Experiment 1: Stimulus and task dependent changes in perceptual durations

In Experiment 1, we investigated the influence of angular velocity and dot-density of the ambiguous sphere on the temporal dynamics of perception during both natural viewing and a voluntary control condition (in which subjects were instructed to switch between perceptual interpretations as fast as possible). Five different densities and angular velocities were used, creating a total of 25 different ambiguous spheres. During natural viewing, perceptual durations decreased as a function of increased density ($F_{4,599} = 29.70$, $p = 0.000$) and increased velocity ($F_{4,599} = 54.02$, $p = 0.000$). An interaction between these two parameters was not significant, suggesting independence between velocity and density in their influence on the duration of perceptual phases (Figs. 2A and C). A similar, but somewhat weaker relationship was found for the voluntary control condition: mean durations decreased as a function of increased density ($F_{4,599} = 10.63$, $p = 0.000$) and increased velocity ($F_{4,599} = 6.96$, $p = 0.000$), see Figs. 2B and D; *note that the axes are not equal when comparing* Figs. 2A and C *with* Figs. 2B and D. An interaction between these parameters was not significant. Overall, the mean durations during voluntary control were significantly shorter than during natural viewing (*natural viewing*: 10.3 s, *voluntary control*: 4.09 s; $F_{1,599} = 705.02$, $p = 0.000$), demonstrating the success of voluntary control. In addition, it can be observed that the slopes associated with the natural condition are steeper compared to those associated with the voluntary control condition, indicative of an interaction between condition and stimulus parameters.

We indeed found a significant interaction between condition and velocity ($F_{4,599} = 29.97$, $p = 0.000$) and condition and density ($F_{4,599} = 11.38$, $p = 0.000$): as density and velocity are increased, the difference in mean duration between conditions decreases (Fig. 3A). From this difference it cannot be concluded that the success of voluntary control depends directly on stimulus parameters. A similar result could have been obtained if voluntary control merely decreases the durations of perceptual phases during natural

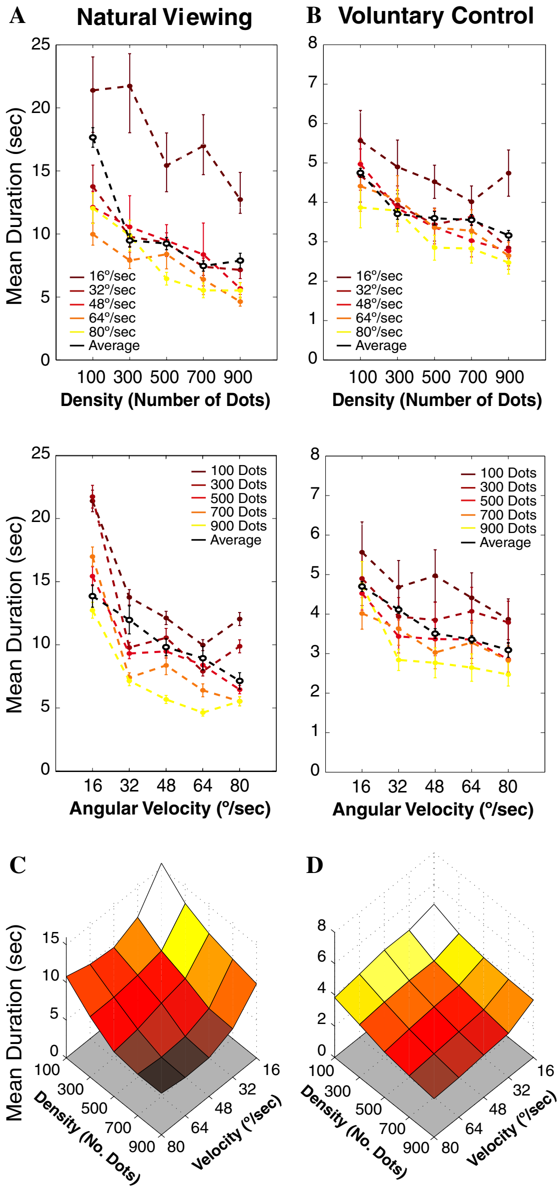

Fig. 2. Increased dot-density and angular velocity decrease perceptual durations for both natural (A) and voluntary control conditions (B). Mean durations for angular velocity and density combined for natural viewing (C) and voluntary control (D). Comparing (C) and (D), exerting voluntary control decreases durations for all densities and velocities. Error bars in (A) and (B) depict SEM. Note that the axes are not equal when comparing (A and B) and (C and D).

viewing by a fixed ratio or gain (instead of as a function of stimulus parameters): as the durations decrease when we increase density or velocity during natural viewing, durations during the voluntary control decrease even further (due to the multiplicative nature of applying a fixed ratio or gain). True evidence for direct dependence of voluntary control on stimulus parameters comes from examining the ratio of durations between conditions. If constant, voluntary control merely decreases the durations by a fixed factor, independent of stimulus characteristics. If not, voluntary control does depend on stimulus characteristics directly. Analysis revealed that this ratio is not constant:

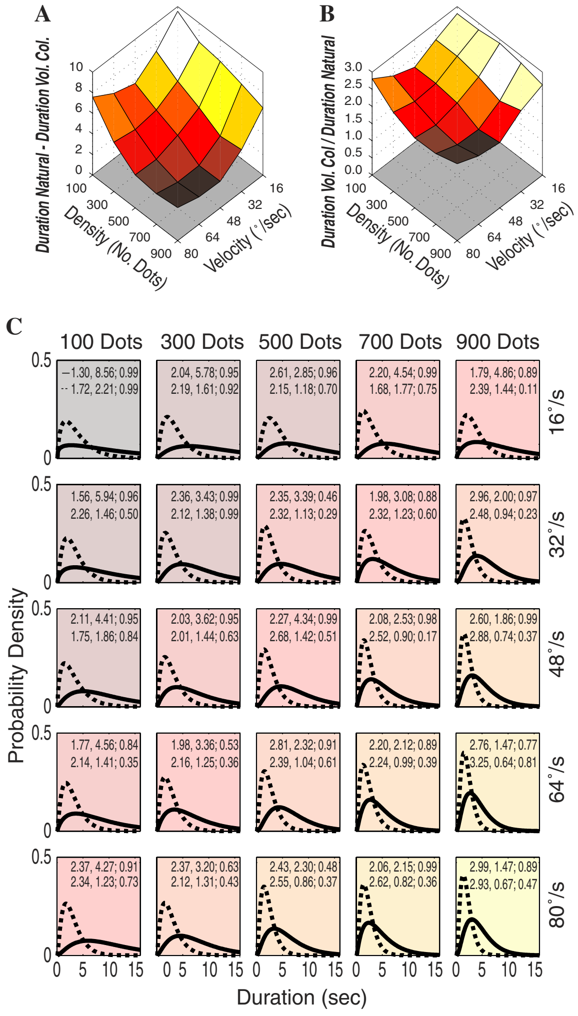

Fig. 3. Comparison of the natural and voluntary control conditions in terms of differences between durations (A) and ratios between durations (B). (A) As densities and velocities increase, the difference in mean durations between natural viewing and voluntary control decreases. (B) Furthermore, the decrease seen in (A) is not constant: the ratio in durations between natural and voluntary control conditions also decreases as a function of increased densities and velocities, demonstrating that the success of voluntary control is directly influenced by characteristics constituting the stimulus and that its effectiveness decreases for increased velocities and densities. (C) Fitted gamma distributions of durations for both conditions (natural, solid line; voluntary control, dashed line) across different densities and velocities. Inset show fitted parameters (shape/scale) and the associated fit qualities. All fits were found to be of significant quality (Section 2). The fits also illustrate that at higher densities and velocities, the maximum (or mode) of distributions are displaced to shorter perceptual durations.

it decreased for increased velocities ($F_{4,24} = 22.33$, $p = 0.000$) and to a lesser extend increased densities ($F_{4,24} = 6.47$, $p = 0.002$). This means that voluntary control decreases in its efficiency to shorten durations for higher velocities and densities (Fig. 3B).

3.2. Analysis of underlying distributions

We fitted the perceptual durations associated with each condition to a Gamma distribution with two free parameters: shape and scale (see Eq. (1) and Section 2). These gamma distributions (Fig. 3C) reproduce our earlier finding: the peaks for voluntary control exertions are displaced to the left, indicative of shorter mean durations. This shift is also observable for increased velocities and densities. Included within each plot are the fitted shape and scale parameters and a measure of the goodness-of-fit between data and fit, respectively. Quality of the fit between data and distribution was determined using a Kolmogorov–Smirnov test (Brascamp et al., 2005),[2] revealing that all fits could be accepted in terms of quality. We analyzed the influence of the three experimental manipulations (density, velocity, and natural vs. voluntary control) in terms of changing shape and scale parameters (van Ee, Noest, Brascamp, & van den Berg, 2006), extending our above-presented analysis of mean durations.

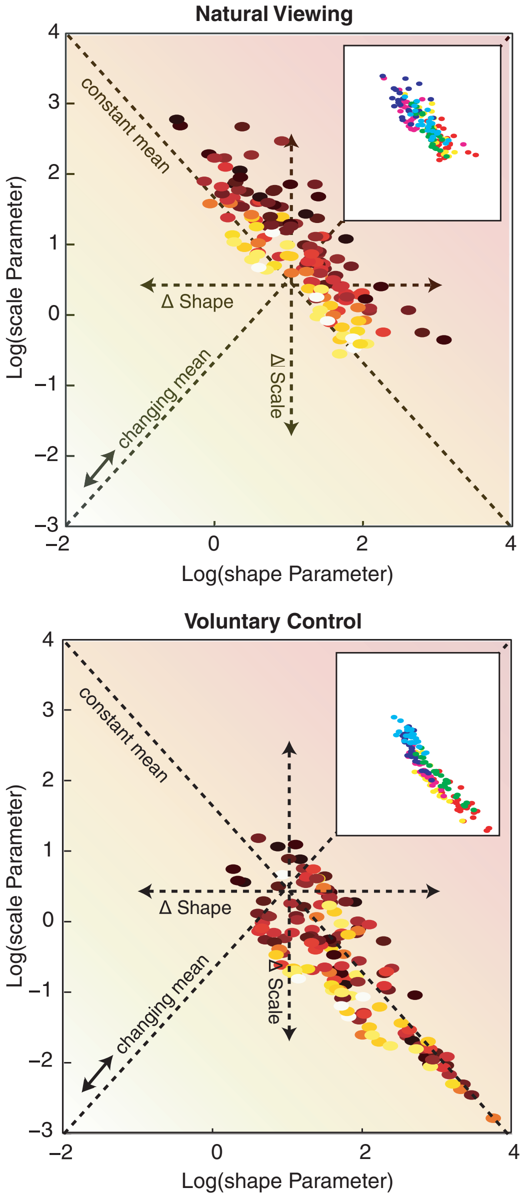

Fig. 4. Relationship between scale and shape parameters of the fitted Gamma distributions of durations for all velocities and densities for natural viewing (top) and voluntary control (bottom) in logarithmic space. Color coding of data points represents velocity and density combined (black, low-densities/velocities; white, high-densities/velocities). Background color gradient represents equal means (similar color equals similar mean). The difference in mean durations of perceptual phases between natural and voluntary control conditions is realized by a decrease in scale and increase in shape parameters. Insets, fits for individual subjects (different colors) tend to cluster together. In addition, the relative positioning of the subjects in terms of scale and shape remains fairly constant across tasks.

Fig. 4 shows the relationship between shape and scale parameters (in logarithmic space) for each parameter (color coding) and for both conditions (each data point represents a fit of the phases given a particular velocity/density combination and a particular subject). We used logarithmic space because it linearises changes in the scale and shape parameters and pronounces the effects of the experimental manipulations (van Ee et al., 2006). Since the gamma distribution's mean is defined by the shape parameter times the scale parameter (Levelt, 1965; van Ee et al., 2006), decreasing mean durations can be realized by decreasing scale, shape or both. Fig. 4 shows the effect of increasing velocity and density: high-densities/velocities (*white points*) are closer to the origin and thus have lower means than low-densities/velocities (*black points*). Furthermore, in the voluntary control condition the means are lower than in the natural viewing experiment (*compare the location of the individual points relative to the constant mean diagonal in Fig. 4*). The change in mean durations can be explained by specific changes in both scale and shape parameters as the data appears to be shifted downward (decrease in scale) and rightward (increase in shape) from natural to voluntary control experiments, which were both significant (*paired-sample t-test*, shape: $t_{149} = 10.33$, $p = 0.000$, scale: $t_{149} = 19.81$, $p = 0.000$). For the natural condition, the scale parameter showed a significant decrease for higher velocities and densities (velocity: $F_{4,149} = 6.67$, $p = 0.000$, density: $F_{4,149} = 5.91$, $p = 0.000$), while no such significant changes were observed for the shape parameter. For the voluntary control condition, the change in both parameters failed to reach significance as a function of both velocity and density, although the same trends in parameter-changes were observed. Interestingly, the greatest variation occurs in the direction perpendicular to the changes in the mean (Fig. 4). This underscores the benefit of using the here presented analysis in that it reveals one extra dimension of data changes: a dimension that was not revealed by examining only changes in mean durations. The found changes in scale parameter suggest that a change in the mean duration of perceptual phases (for different densities and velocities and between tasks) appear to be accomplished by an overall decrease in the individual

---

[2] It has been shown (Brascamp et al., 2005) that for some bistable stimuli, fit quality is increased when one uses the rate (1/duration) of alternations as opposed to durations. Here, we found no significant differences in quality between fitting durations or rates, and because most of our analyses are in terms of duration we present duration distributions.

durations and not by reducing the number of long durations or increasing the number of short durations. Finally, it can be observed from the insets of Fig. 4 that fits for individual subjects tend to cluster together. In addition, the relative positioning of the subjects in terms of scale and shape remains fairly constant across tasks.

3.3. Experiment 2: Reverse correlation

In Experiment 2, we explored whether we can explain the bistability of the SFM stimulus in terms of local competition between stimulus elements or in terms of global competition between two surface-based representations (CW and CCW). Subjects were presented with spheres containing a localized patch consisting either of a higher number of dots (*high-density patch*) or no dots at all (*gap-patch*). This patch was placed randomly on the surface of the sphere and rotated at the same speed as the dots. We recorded alternations and reversed correlated the position and direction of these patches at the time of alternations with the newly perceived rotation (Fig. 5A). If local asymmetries influence perception in favor of the motion direction associated with the largest number of elements, we expect that the high-density patch will cause alternations towards the patches' motion direction and that the gap patch will cause alternations opposite the patches' motion direction (because through it, only dots in the opposite direction are seen). Figs. 5B and C show the relationship between patch location and direction (color coding) as a function of the direction of the alternations (CW or CCW) in terms of percentages (*top row*) of alternations during which part of the patch was at a particular location and moving in a particular direction. Figs. 5B and C also shows a statistical significance of this percentage value (*bottom row*). To be more specific: a positive value of 50% on the percentage map means that during as much as half of all alternations, part of the patch was located at that position and was moving in the same direction as the newly perceived rotation. We associated these percentages with a statistical value by determining whether these percentages deviated significantly from zero across all alternations (see Section 2). We found that both high-density and the gap patches have a strong effect on perception. More importantly, subjects will alternate towards the direction of the patch regardless as to whether it is of higher density or void of any motion. Figs. 5B and C shows that for both types of patches it is more likely that the patch was moving in the same direction as the newly perceived direction after the alternation. These results demonstrate that having a larger number of dots moving in a certain direction does not guarantee that perception is biased towards that perceived rotational direction. Instead, it is more likely that the motion direction that is most salient (due to a distinguishing feature such as a patch) has a much greater influence on perception.

Although in principle our findings demonstrate a correlation between the perceived rotation of the patch and sphere, there is reason to believe that the patches cause (or trigger) perceptual alternations: for all three subjects, the location of the patch matters: alternations of the sphere are most frequently experienced when the patch is at a particular location (around central fixation for S1 and S2 and moving towards central fixation for S3, see Fig. 5). A second property that can be deduced from Fig. 5 is that the perceptual phases become more regular when a patch is included: if the sphere switches rotation primarily when the patch is at a certain location and the rotation of both sphere and patch is constant in velocity (and equal) this means that most perceptual phases last as long as it takes the patch to complete roughly one (or multiples of one) full rotation. These two properties are obviously not independent but merely reflect the two possible ways of demonstrating the influence of the patch on the sphere: in both temporal and spatial terms. It does suggest that the perceived rotation of the sphere is at least highly depended on patch location in space and time and that, in all probability, patches can trigger perceptual alternations of the sphere in the direction of the rotation of the patch.

3.4. Experiment 3: Static vs. dynamic stimuli and voluntary control

In a final experiment, we examined whether voluntary control is mediated by attentive tracking of individual dots. Using spheres in which the dots had limited lifetimes we compared those 'dynamic' spheres with the more conventional 'static' spheres containing dots with an unlimited lifetime. If tracking dots facilitates voluntarily switching perceptual states, subjects ought to show reduced ability in switching dynamic spheres since it is not possible to track dots. The data shows that dynamic spheres are not associated with significant shorter durations compared to the static spheres in both conditions (Fig. 6). In addition, replicating our previous findings, we found that voluntary control significantly decreases mean durations for both the static (*natural condition*: 4.54 s, *voluntary control condition*: 2.75 s, $F_{1,51} = 12.25$, $p = 0.0019$) and the dynamic sphere (*natural condition*: 4.34 s, *voluntary control condition*: 2.49 s, $F_{1,51} = 11.30$, $p = 0.0014$). Furthermore, the ratio (see insets in Fig. 6) between natural and voluntary control conditions did also not change significantly between the static and dynamic spheres. These results demonstrate that attentive tracking of individual dots is not essential for voluntary control, suggesting that subjects track the surface implied by the dots, rather then the dots themselves.

3.5. Eye movement recordings

Fig. 7 depicts the number of saccades, smooth pursuit eye movements, and alternations for natural and voluntary control conditions, revealing three important findings. First, subjects decreased the durations of the perceptual phases for voluntary control compared to natural viewing. Second, when eye movements were allowed during volun-

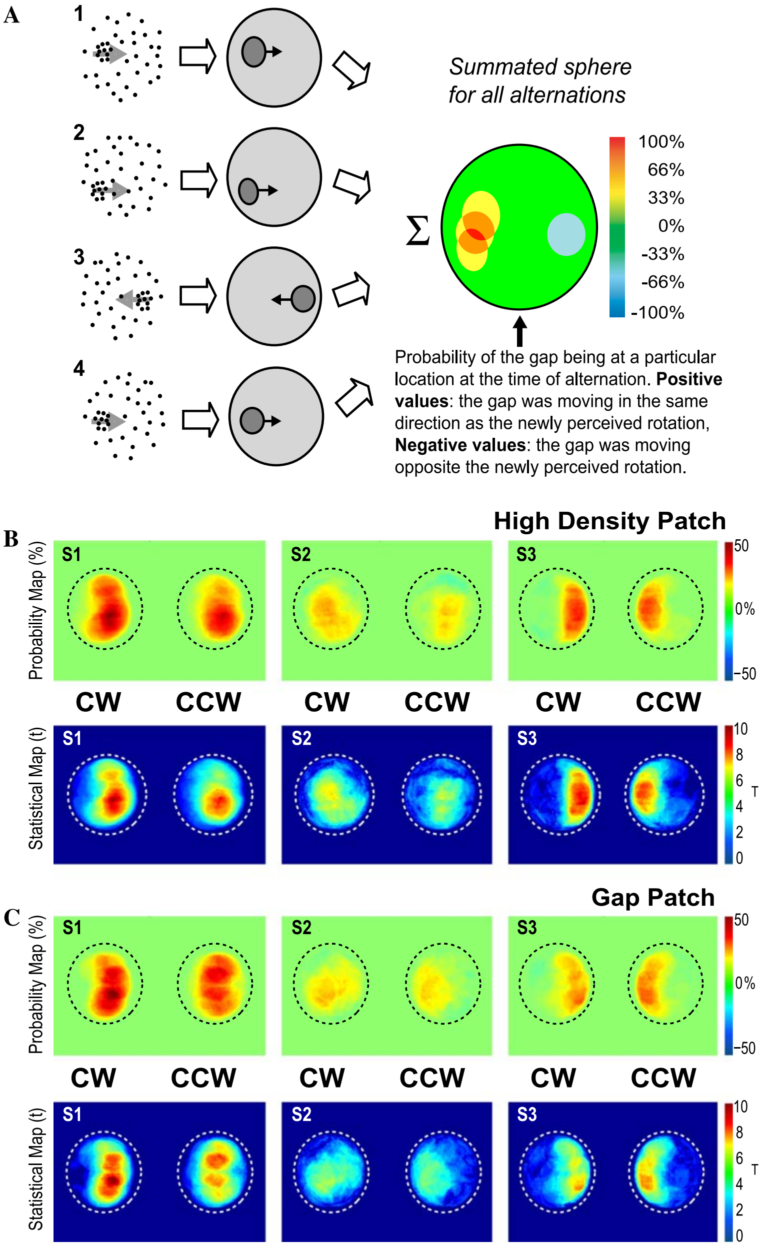

Fig. 5. Spatial reverse correlation. (A) Schematic representation of the method. Shown are (hypothetical) reconstructed spheres (containing a local patch of increased dot-density) at the time of four alternations towards a CCW perceived rotation. For alternations 1, 2, and 4, the patch moves in the direction of newly perceived rotation (CCW) and are added to an average image. In alternation 3, the patch moves in the opposite direction and is subtracted from the same image. After summation, this average image is divided by the total number of alternations towards the associated perceived direction (in this example: CCW), yielding a probability map. Positive values at a particular location indicate a large probability that part of the patch was moving in the same direction as the newly perceived rotation at that particular location, negative values indicate movement in the opposite direction of the newly perceived rotation. Spatiotemporal correlation between the position of a high-density patch (B) or gap patch (C) and alternations in three subjects (columns). Top row, probability of the patch being at a particular position in a particular direction (color coding, green to red, direction of patch identical to direction of newly perceived rotation after an alternation; green to blue, direction of patch opposite to direction of newly perceived rotation after an alternation) at the time of an alternation to a CW rotating sphere (left) and an alternation to a CCW rotating sphere (right). Bottom row, associated statistical maps. For both types of stimuli there is a tight correlation between the position and the direction of the patch. As indicated by the surplus of high positive percentages (up to 50%, indicating a 50% chance of an alternation when the gap moves through a particular point), both types of patches evoke alternations in the direction of the patch itself.

tary control, the perceptual durations decreased even further. Third and most importantly: mean perceptual durations decreased significantly from the natural to voluntary control condition under strict fixation (*natural condition:* 10.1 s, *voluntary control condition with strict fixation:* 5.4 s; $F_{1,35} = 45.97$, $p = 0.000$) without significant increases in the frequency of saccades and smooth pursuit eye movements.

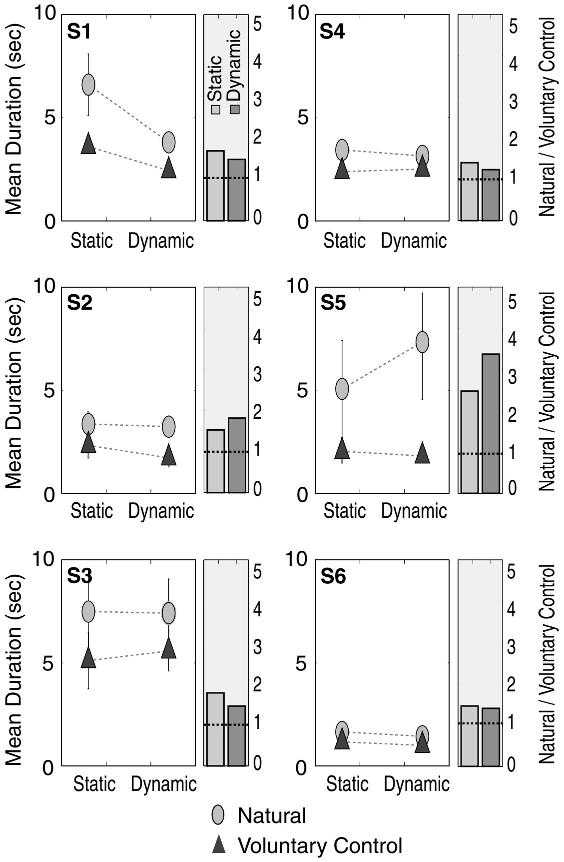

Fig. 6. Mean perceptual durations as a function of natural (gray discs) and voluntary control (black triangles) conditions for static and dynamic spheres for all six subjects. Bar graphs depict the ratio between the durations in the voluntary control and natural, for each type of sphere (static/dynamic). Replicating our previous findings, voluntary control conditions are associated with shorter mean durations than the natural condition for both static and dynamic spheres (compare gray discs/black triangles). More importantly, the ratio of durations (voluntary control/natural) did not change between the static and dynamic spheres (see insets), indicating that voluntary control is just as efficient during viewing of dynamic spheres as it is for static spheres. Error bars represent SD.

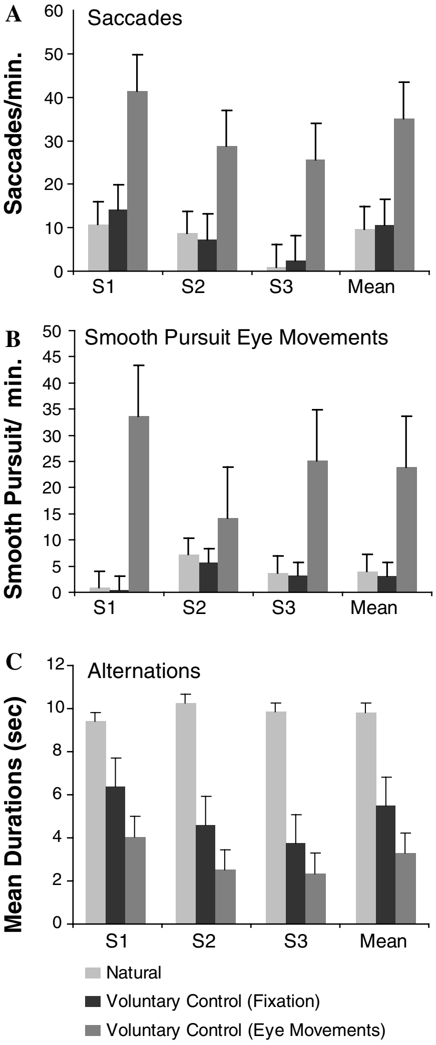

Fig. 7. The number of (A) saccades per minute, (B) smooth pursuit eye movements per minute and (C) mean durations as a function of the three conditions. Perceptual durations decrease significantly (see text) from the natural to the voluntary control under strict fixation condition, while no significant changes in the number of saccades and smooth pursuit movements are observed. This means that subjects can exert voluntary control without the need for eye movements. Furthermore, when eye movements are permitted (and made), mean perceptual durations decrease even further.

This demonstrates that eye movements are not required for (the success of) exerting voluntary control (i.e., decreasing the mean duration of perceptual phases) over perception during bistable SFM. This fits well with known observations: although eye movements do facilitate perceptual alternations, they appear not to be essential (Toppino, 2003; van Dam & van Ee, 2006).

4. Discussion

4.1. *Exogenous and endogenous influences on temporal dynamics*

Investigating the influence of varying physical parameters constituting the stimulus on the temporal dynamics of perception during both natural and voluntary control conditions, we found that increasing dot-density and angular velocity of the sphere decreases durations of perceptual phases during both conditions. In addition, the low interaction between density and velocity for both conditions points to a relatively independent influence of these measures on the temporal dynamics of bistable perception.

The decrease in durations from natural to voluntary control conditions provides evidence for the ability of observers to voluntarily decrease these durations, and therefore the ability to influence perception. This has been demonstrated previously for various distinct ambiguous stimuli (e.g., Hol et al., 2003; Lack, 1978; Meng & Tong, 2004; Peterson, 1986; Suzuki & Peterson, 2000; Toppino, 2003; van Ee et al., 2005). More importantly, our results indicate that the ratio of durations between the natural and voluntary control condition for different dot-densities and velocities is not constant, demonstrating that exerting voluntary control does not merely decrease the mean durations by a fixed ratio, but is itself also influenced by parameters constituting the stimulus. This dependency places important constraints on possible mechanisms mediating voluntary control: these mechanisms cannot operate freely regardless of stimulus characteristics and the success of a self-initiated change in perception is a function of the possibility of those changes occurring naturally.[3] To assess underlying mechanisms in more detail, we fitted perceptual durations to a two-parameter gamma distribution (Brascamp et al., 2005; Levelt, 1965; van Ee et al., 2006). As we have shown, this type of analysis can potentially reveal interesting effects that remain hidden when only variations in mean duration are examined. Fitted parameters revealed that the decreasing durations as a function of velocity, density and exerting voluntary control can be explained by specific changes in both parameters and that the scale parameter accounts for most of the variance. The similarity in these changes between velocity, density and voluntary control could suggest a common mechanism. Given the statistical relationship between the mean, scale and shape parameters we conclude that voluntary control results in a decrease of all individual perceptual durations, rather than an increase of short durations and/or decrease of long durations. This, too, constrains the possibly mechanism mediating voluntary control. Our results corroborate a related finding from literature: biasing a perceptual interpretation by changing stimulus eccentricity makes observers' intentional effort to perceive that interpretation more successful (Suzuki & Peterson, 2000). Here, we show that changing parameters constituting the stimuli can also have a profound effect on voluntary control, even without changing the relative dominance of each perceptual interpretation.

One possibility why increased velocity and density decrease perceptual durations is that these increased densities and velocities are associated with an elevated neural response. In turn, this elevated neural response could imply faster adaptation of neurons representing the current dominant percept. This could result in faster switching between states because of instability of the network.

4.2. Surface interpolation

It is natural to think that when the visual system is presented with ambiguous information, these information sources compete for awareness. To investigate the possible level of competition during ambiguous structure-from-motion, we introduced both global and local asymmetries on the sphere. The visual system uses the motion of individual elements to reconstruct a complete surface in depth (He & Nakayama, 1994b; Hildreth et al., 1995; Treue et al., 1995). This suggests that competition could occur between individual elements that are incongruent (leftward versus rightward moving dots) or between surface-based representations, each associated with a different rotational direction. Using reverse correlation, we demonstrated that differences in the number of dots moving in either direction can have a profound effect on perception. More importantly, we found that subjects alternate towards the direction of the patch regardless as to whether it is of higher density or void of any motion (and the asymmetry is thus associated with a larger number of dots in the opposite direction). To be more specific, we showed that when one removes dots locally (introducing a gap within the sphere) perception is not biased towards the motion visible through the gap (which is in the opposite direction as the gap itself) but instead it is biased in the direction of the gap. We suggest that the gap, even though it is not containing any stimulus elements, still enhances the saliency of its associated motion.

Finally, we have provided evidence that voluntary control over the temporal dynamics of perception during viewing of the sphere cannot be mediated completely by attending to the motion of individual dots. Instead, subjects are quite capable of switching perceptual states when these dots have a very limited lifetime, suggesting that voluntary control does not interact with perception at the level of individual elements but at a higher level of surface-based representations.

5. Conclusions

Taken together, our results constrain the mechanism underlying bistable perception in several ways. First, changes in physical parameters influence the temporal dynamics of bistable perception. Second, voluntary control is quite successful in changing (i.e., decreasing) the durations of perceptual phases. Most importantly, voluntary control did depend directly on the physical parameters constituting the stimulus. This dependency places important constraints on possible mechanisms mediating voluntary control: these mechanisms cannot operate independently of stimulus characteristics.

In addition, voluntary control during viewing of the bistable sphere appears to be mediated by attending to a particular motion direction, instead of tracking individual stimulus elements. Third, our experiments suggest that dominance of either perceptual state depends on the saliency of the motion, and not so much on a competition between stimulus elements (individual dots) constituting the stimulus. Finally, we demonstrated that eye movements are not essential for exerting voluntary control (i.e., decreasing durations) over perception during bistable SFM.

Acknowledgments

We thank reviewers for thoughtful and constructive comments. We also thank J.W. Brascamp and W. Pestman

---

[3] Note added after manuscript submission: It has recently been reported that both endogenous and exogenous attention can influence initial dominance (Chong & Blake, in press). Here, we extend these findings by demonstrating that endogenous influences depend on exogenous aspects.

for technical support. The authors were supported by a grant from the Netherlands Organization for Scientific Research (NWO) assigned to R.v.E.

References

Andersen, R. A., & Bradley, D. C. (1998). Perception of three-dimensional structure from motion. *Trends in Cognitive Neuroscience, 2*, 222–228.

Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience, 3*, 13–21.

Bradley, D. C., Qian, N., & Andersen, R. A. (1998). Encoding of three-dimensional structure-from-motion by primate area MT neurons. *Nature, 392*, 609–611.

Brascamp, J. W., van Ee, R., Pestman, W. R., & van den Berg, A. V. (2005). Distributions of alternation rates in various forms of bistable perception. *Journal of Vision, 5*, 287–298.

Brouwer, G. J., van Ee, R., & Schwarzbach, J. (2005). Activation in visual cortex correlates with the awareness of stereoscopic depth. *Journal of Neuroscience, 25*, 10403–10413.

Chong, S. C., Blake, R. (in press). Exogenous attention and endogenous attention influence initial dominance in binocular rivalry. *Vision Research*, doi:10.1016/j.visres.2005.10.031.

Chong, S. C., Tadin, D., & Blake, R. (2005). Endogenous attention prolongs dominance durations in binocular rivalry. *Journal of Vision, 5*, 1004–1012.

Dodd, J. V., Krug, K., Cumming, B. G., & Parker, A. J. (2001). Perceptually bistable three-dimensional figure evoke high choice probabilities in cortical areas MT. *Journal of Neuroscience, 21*, 4809–4821.

He, Z. J., & Nakayama, K. (1992). Surfaces versus features in visual search. *Nature, 359*, 231–233.

He, Z. J., & Nakayama, K. (1994a). Apparent motion determined by surface layout, not by disparity or three-dimensional distance. *Nature, 367*, 173–175.

He, Z. J., & Nakayama, K. (1994b). Surface shape, not features determines apparent motion correspondence. *Vision Research, 34*, 2125–2136.

Hildreth, E. C., Ando, H., Andersen, R. A., & Treue, S. (1995). Recovering three-dimensional structure from motion with surface reconstruction. *Vision Research, 35*, 117–137.

Hol, K., Koene, A., & van Ee, R. (2003). Attention-biased multi-stable surface perception in three-dimensional structure-from-motion. *Journal of Vision, 3*, 486–498.

Kleinschmidt, A., Buchel, C., Zeki, S., & Frackowiak, R. S. (1998). Human brain activity during spontaneously reversing perception of ambiguous figures. *Proceedings of the Royal Society of London. Series B, 265*, 2427–2433.

Lack, L. C. (1978). *Selective attention and the control of binocular rivalry*. The Hague, The Netherlands: Mouton.

Leopold, D. A., & Logothetis, N. K. (1996). Activity changes in early visual cortex reflects monkey' percepts during binocular rivalry. *Nature, 379*, 549–553.

Levelt, W. (1965). *On binocular rivalry*. Soesterberg, The Netherlands: Institute of Perception.

Logothetis, N. K., & Schall, J. D. (1989). Neuronal correlates of subjective visual perception. *Science, 245*, 761–763.

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

Meng, M., & Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision, 45*, 39–51.

Miles, W. R. (1931). Movement interpretations of the silhouette of a revolving fan. *The American Journal of Psychology, 42*, 392–405.

Neri, P., & Heeger, D. J. (2002). Spatiotemporal mechanisms for detecting and identifying image features in human vision. *Nature Neuroscience, 5*, 812–816.

Neri, P., Parker, A. J., & Blakemore, C. (1999). Probing the human stereoscopic system with reverse correlation. *Nature, 14*, 695–698.

Parker, A. J., & Krug, K. (2003). Neuronal mechanisms for the perception of ambiguous stimuli. *Current Biology, 13*, 433–439.

Peterson, M. A. (1986). Illusory concomitant motion in ambiguous stereograms: evidence for nonstimulus contributions to perceptual organization. *Journal of Experimental Psychology, 12*, 50–60.

Polonsky, A., Blake, R., Braun, J., & Heeger, D. J. (2000). Neuronal activity in human primary visual cortex correlates with perception during binocular rivalry. *Nature Neuroscience, 3*, 1153–1159.

Suzuki, S., & Peterson, M. A. (2000). Multiplicative effects of intention on the perception of bistable apparent motion. *Psychological Science, 11*, 202–209.

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

Toppino, T. C. (2003). Reversible-figure perception: mechanisms of intentional control. *Perception and Psychophysics, 65*, 1285–1295.

Treue, S., Andersen, R. A., Ando, H., & Hildreth, E. C. (1995). Structure-from-motion: perceptual evidence for surface interpolation. *Vision Research, 1*, 139–148.

Treue, S., Husain, M., & Andersen, R. (1991). Human perception of structure from motion. *Vision Research, 31*, 59–75.

van Dam, L. C. J., & van Ee, R. (2006). The role of saccades in exerting voluntary control in perceptual and binocular rivalry. *Vision Research, 46*, 787–799.

van Ee, R., Noest, A. J., Brascamp, J. W., van den Berg, A. V. (2006). Attentional control over either of the two competing percepts of ambiguous stimuli revealed by a two-parameter analysis: Means do not make the difference *Vision Research*, *46*, in press, doi:10.1016/j.visres.2006.03.017.

van Ee, R., van Dam, L. C. J., & Brouwer, G. J. (2005). Voluntary control and the dynamics of perceptual bi-stability. *Vision Research, 45*, 41–51.

Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology, 45*, 205–217.