



Out of sight, out of mind: Occlusion and eye closure destabilize moving bistable structure-from-motion displays

Alexander Pastukhov^{1,2} · Johanna Prasch¹ · Claus-Christian Carbon^{1,2}

Published online: 20 March 2018
© The Psychonomic Society, Inc. 2018

Abstract

Our brain constantly tries to anticipate the future by using a variety of memory mechanisms. Interestingly, studies using the intermittent presentation of multistable displays have shown little perceptual persistence for interruptions longer than a few hundred milliseconds. Here we examined whether we can facilitate the perceptual stability of bistable displays following a period of invisibility by employing a physically plausible and ecologically valid occlusion event sequence, as opposed to the typical intermittent presentation, with sudden onsets and offsets. To this end, we presented a bistable rotating structure-from-motion display that was moving along a linear horizontal trajectory on the screen and either was temporarily occluded by another object (a cardboard strip in Exp. 1, a computer-generated image in Exp. 2) or became invisible due to eye closure (Exp. 3). We report that a bistable rotation direction reliably persisted following occlusion or interruption only (1) if the pre- and postinterruption locations overlapped spatially (an occluder with apertures in Exp. 2 or brief, spontaneous blinks in Exp. 3) or (2) if an object's size allowed for the efficient grouping of dots on both sides of the occluding object (large objects in Exp. 1). In contrast, we observed no persistence whenever the pre- and postinterruption locations were nonoverlapping (large solid occluding objects in Exps. 1 and 2 and long, prompted blinks in Exp. 3). We report that the bistable rotation direction of a moving object persisted only for spatially overlapping neural representations, and that persistence was not facilitated by a physically plausible and ecologically valid occlusion event.

Keywords Bistable perception · Multistable perception · Predictive perception · Visual memory · Tunnel effect · Structure from motion · Ambiguity · Persistence

Our brain is a prediction machine. Its function is not only to build a useful representation of an outside world, in order to guide our behavior but also to anticipate its future states to optimize the use of limited resources. The predictive nature of perception has been extensively studied using multistable displays, such as binocular rivalry, the Necker cube, or structure from motion (SFM; see the [supplementary videos](#)). These

displays are compatible with two or more comparably likely and easily distinguishable perceptual outcomes; a combination that makes their perception unstable and perceptual changes very noticeable. When multistable displays are presented intermittently, their perceptual dominance at the display onset depends on numerous factors, such as attention (Mossbridge, Ortega, Grabowecky, & Suzuki, 2013), the relative strength of competing interpretations (Hupé, Lamirel, & Lorenceau, 2009; Song & Yao, 2009), or prior perceptual experience (Klink et al., 2008; Kornmeier & Bach, 2004; Orbach, Ehrlich, & Heath, 1963).

In the latter case, multistable perception is stabilized by neural persistence or by the sensory memory of multistable displays. Neural persistence is a continued response of neurons after stimulus offset (Coltheart, 1980). It is most effective for brief interruptions (e.g., <200 ms) when bistable displays remain constant and at the same retinotopic location. The highly specific nature of neural persistence, in combination with the sensitivity to masking, confines its influence to very brief interruptions such as fast blinks or small eye movements.

Electronic supplementary material The online version of this article (<https://doi.org/10.3758/s13414-018-1505-z>) contains supplementary material, which is available to authorized users.

✉ Alexander Pastukhov
Pastukhov.Alexander@gmail.com;
<https://alexander-pastukhov.github.io/>

¹ Department of General Psychology and Methodology, University of Bamberg, Markusplatz 3, D-96047 Bamberg, Bavaria, Germany

² Forschungsgruppe EPÆG (Ergonomics, Psychological Æsthetics, Gestalt), Bamberg, Bavaria, Germany

In contrast, the sensory memory of multistable displays is longer lasting (Leopold, Wilke, Maier, & Logothetis, 2002) and is highly resistant to intervening visual stimulations (Maier, Wilke, Logothetis, & Leopold, 2003), making it a better candidate for a predictive memory (Pearson & Brascamp, 2008). However, its utility is severely limited by the very weak influence it exerts, much weaker than the influence of either neural persistence or perceptual adaptation (Pastukhov & Braun, 2013). Here we asked whether this weakness might reflect possible shortcomings of a common intermittent-presentation design.

Specifically, the display in a typical intermittent-presentation study behaves like a Cheshire cat, with a sudden appearance being followed by an equally sudden disappearance. This presentation schedule has little in common with normal visual event sequences in our daily lives, as objects typically become invisible because they are gradually occluded or because we close our eyes. Moreover, prior research indicates that a physically plausible and consistent visual sequence facilitates the persistence of a temporarily invisible object. For example, in the tunnel effect, the persistence of an object that passes behind an occluder is facilitated by the visibility of an occluding object, a predictable trajectory, a consistent deletion/accretion sequence, and so forth (Flombaum & Scholl, 2006; Kawachi & Gyoba, 2006). Similarly, during endogenously generated saccades that briefly interrupt the normal flow of sensory evidence, perceptual stability is ensured by predictive remapping of the object's features to the future location (Melcher, 2007).

Here, we investigated whether we can facilitate the persistence of bistable displays following a period of invisibility by employing physically plausible and ecologically valid occlusion event sequences. Specifically, we employed a bistable rotating structure-from-motion display that moved along a linear trajectory on the screen and either was temporarily occluded by another object (Exps. 1 and 2) or became invisible due to eye closure (Exp. 3).

An occlusion by another object had been used previously in conjunction with the binocular rivalry display (Blake, Sobel, & Gilroy, 2003). The latter consisted of two incompatible dichoptically presented patterns, so that only one of them tended to be perceived at a time and the inputs from the other eye are suppressed. The binocular rivalry display moved along the circular trajectory and was temporarily occluded by a stationary object. That study reported strong destabilization by occlusion. However, because binocular rivalry relies on interocular suppression, the eye dominance rather than the object representation was what failed to persist. Accordingly, in Experiments 1 and 2 we sought to address this issue by using a bistable structure-from-motion display that relied on a distributed representation in extrastriatal regions (Orban, 2011).

Eye closure has also been used previously to study the intermittent perception of bistable displays (Leopold et al., 2002). Here, several authors have reported a strong perceptual stabilization, which is thought to reflect an influence of the sensory memory of multistable displays (Adams, 1954; Leopold et al., 2002; Orbach, Ehrlich, & Vainstein, 1963). However, the previous multistable displays had remained at the same spatial location throughout the entire presentation session. Thus, for this measurement it was hard to disentangle lower-level, location-specific from potential higher-level predictive effects, which might be trajectory- rather than location-specific. In our Experiment 3, we sought to alleviate these potential confounding factors by combining eye closure with the moving bistable structure-from-motion display.

Method

Participants

All procedures were in accordance with the national ethical standards on human experimentation and with the Declaration of Helsinki of 1975, as revised in 2008, and were approved by the University of Bamberg. The observers had normal or corrected-to-normal vision and showed normal color vision, with the exception of observer SDA95m, who had a red–green deficiency. Apart from one of the authors (observer SKL94w), all observers were naïve as to the purpose of the experiments. Informed consent was obtained from all participants prior to the experimental session.

In total, 14 participants took part in the experiments. Five of the participants, including the second author, took part in all three experiments; three participants took part in Experiments 1 and 2; one took part in Experiments 2 and 3; and five took part in only one experiment. Ten observers (eight females, two males; ages 16–28 years) participated in Experiment 1. Nine observers (six females, three males; ages 16–28 years) participated in Experiment 2 (observer BPM97w was excluded from the analysis due to a strong perceptual bias in favor of the downward direction of rotation, since it constituted more than 95% of this participant's total clear perception reports). Nine observers (four females, five males; ages 21–28) participated in Experiment 3.

Apparatus

In the first two experiments, displays were presented on a 24.5-in. EIZO CG245W screen (size of the visible area 51.7 cm × 32.3 cm, resolution 1,920 × 1,200, refresh rate 59 Hz, viewing distance 50 cm; the head was stabilized with a chin rest). A single-pixel subtended approximately 0.029°. In Experiment 1, observers listened to the panning sound using Sennheiser HD-202 headphones.

In the third experiment, displays were presented on a Samsung SyncMaster 2233 (size of visible area 47.5 cm × 29.5 cm, resolution 1,680 × 1,050, refresh rate 120 Hz, viewing distance 50 cm; the head was stabilized with a chin rest). A single pixel subtended approximately 0.0302°. The observers listened to the auditory signal over the loudspeakers. Eye movements were monitored binocularly with a desk-mounted eyetracker (EyeLink 1000, SR Research) at a frequency of 1,000 Hz.

Displays

Observers viewed a moving, ambiguously rotating, structure-from-motion (SFM) display (see Fig. 1b and Videos 1–11, described in the Appendix). The width of the SFM object was systematically varied in Experiment 1 (4.3°, 5°, 5.7°, 7.1°, and 8.5° of visual angle, respectively; 1:1, 1 $\frac{1}{6}$:1,

1 $\frac{1}{3}$:1, 1 $\frac{1}{3}$:1, and 2:1 ratios between the width of the SFM object and the occluding strip), but was kept constant at 4.3° of visual angle in Experiments 2 and 3. The individual dots subtended 0.03° and were semitransparent in order to exclude bias from the occlusion cues.

The SFM object rotated around the horizontal axis (90°/s) while moving in a horizontal trajectory at a constant speed of 5.4°/s. The trajectory endpoints were at 10.8° of eccentricity. The presentation duration of a single block was 48 s. We split the continuous time series for each block into trials, with a single trial being defined as the time for the SFM object to traverse from one trajectory limit to the other (see Fig. 1e for the schematic schedule representation). Accordingly, each block consisted of 12 trials, and each trial was 4 s long.

In Experiment 1, the central portion of the screen was occluded by a rectangular piece of cardboard (4.3° wide; see Fig. 1a and Videos 1–6). A red fixation point was drawn on it at the location that corresponded to the center of the screen. The

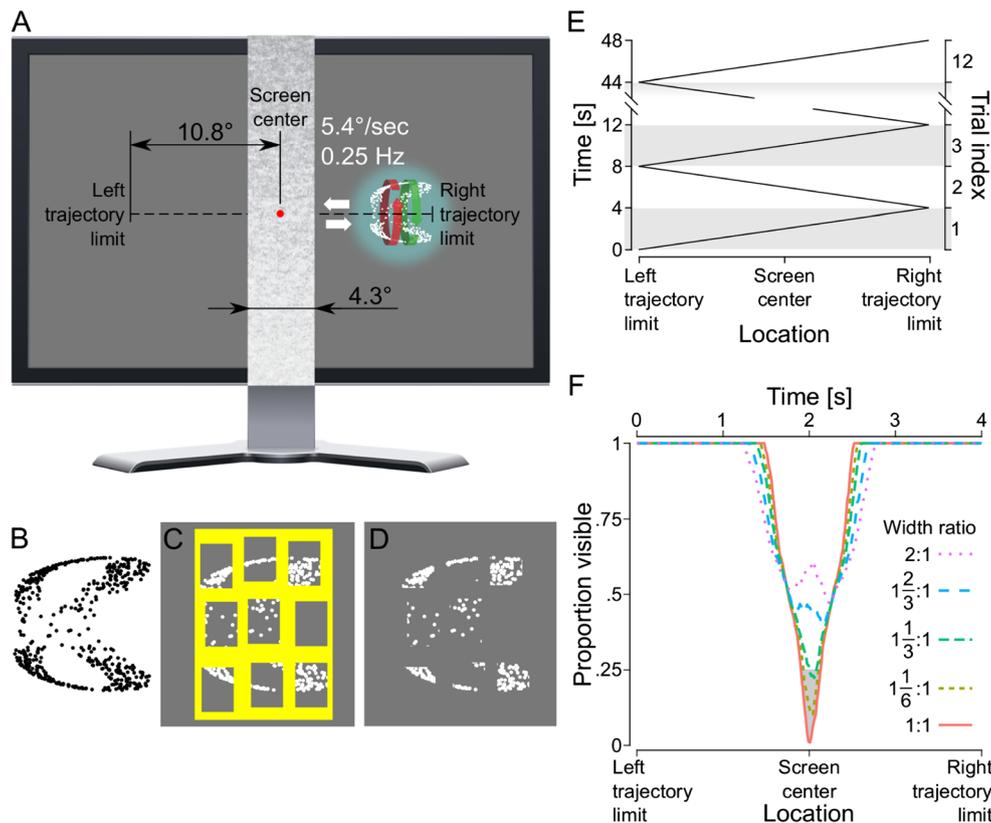


Fig. 1 Visual stimuli and procedure. (a) Schematic display and procedure for Experiment 1, not drawn to scale. A structure-from-motion (SFM) object, which rotated ambiguously around the horizontal axis, repeatedly traversed the screen in a horizontal trajectory (marked by the dashed line). The central portion of the screen was occluded by a rectangular piece of cardboard. Observers were instructed to fixate on a red dot drawn on the cardboard occluder and to report on the direction of the rotation. (b) The SFM object used in the study; see also Videos 1–11. (c, d) Experiment 2: The SFM object behind a visible (c) and the same but camouflaged (d) occluder with a 50% aperture area. (e) Schematic

presentation schedule. The continuous time series was split into trials, with a single trial being defined as the time for the SFM object to traverse from one trajectory limit to the other. (f) Proportions of visible dots for SFM objects of various widths as a function of their location on the trajectory. Only the object with a 1:1 width ratio (relative to the width of the occluding strip) is fully occluded for a single frame. Less than 25% of all the dots were visible for 83 ms for the 1 $\frac{1}{3}$:1 width ratio, 166 ms for the 1 $\frac{1}{6}$:1 width ratio, and 233 ms for the 1:1 width ratio

proportion of dots visible at each location is plotted in Fig. 1f. Please note that only the object with the 1:1 width ratio was completely invisible for just a single frame. However, fewer than 25% of all the dots were visible for 83 ms for the $1\frac{1}{3}$:1 width ratio, 166 ms for the $1\frac{1}{6}$:1 width ratio, and 233 ms for the 1:1 width ratio.

In Experiment 2, the occluder was a computer-generated image ($4.3^\circ \times 6.0^\circ$) that completely or partially occluded the moving SFM object. The total occluded area was systematically manipulated and was set at 0% (no occluder, labeled as *full visibility*), 5%, 10%, 25%, 50%, and 100% (complete occlusion). The occluder was either colored yellow (Fig. 1c; *visible* occluder condition) or was the same gray color as the background (Fig. 1d; *camouflaged* occluder condition; see also Videos 7–10). In Experiment 3, the moving SFM object was never occluded (see Video 11). Instead, observers were instructed to shut their eyes in response to an auditory signal. This is detailed in the “**Procedure**” section below.

Pilot measurements for Experiment 1 indicated that complete invisibility strongly destabilized the rotation (first and last authors only and 1:1 object-to-occluder width condition only, informal viewing session; these data were not included in the analysis or the online dataset). Accordingly, to facilitate the persistence of the fully occluded SFM object in Experiment 1, it was overlaid on a “halo” image accompanied by a stereo sound that panned congruently from left to right, or both. The halo image was a blue circle with a gradual decrease in color (Gaussian spatial transparency profile, $L^*a^*b^*$: 62.46, –28.19, –8.47; see Fig. 1a). It was centered on the SFM object and was never completely occluded. The panning sound was constructed using the Audacity 2.0.6 software. It was localized to the left or the right of the listener using the interaural time difference.

To distinguish between persistence and the spatially specific biases, the direction of the rotation of the SFM object was systematically perturbed. Specifically, at the beginning of each trial, the experimental software attempted to induce an exogenously triggered perceptual reversal by inverting the vertical on-screen motion (see Pastukhov, Vonau, & Braun, 2012, for further details). This gave us a better opportunity to compare the effect of persistence with that of the spatially specific memory influences. The exogenous trigger was effective in 82% [44%, 97%] (mean, range) of the trials in Experiment 1, 90% [70%, 97%] of the trials in Experiment 2, and 89% [75%, 100%] of the trials in Experiment 3.

Procedure

Observers reported on the perceived direction of rotation of the SFM object around the horizontal axis using the *up* and *down* arrow keys. They were instructed to fixate either on the

dot drawn on the cardboard occluder (Exp. 1) or on a computer-generated red circle (Exps. 2 and 3). A single block lasted 48 s in Experiments 1 and 2, and 24 s in Experiment 3. Experiments 1 and 2 were measured during a single experimental session. Experiment 3 was measured separately two months later.

Experiment 1 contained 20 conditions: five widths of the SFM object, combined with a present or absent “halo” and a present or absent panning sound. The presentation order was randomized, and the randomized sequence of blocks was presented first in a forward and then in a backward order (ABBA design, 40 blocks in total). See also Videos 1–6.

Experiment 2 contained 12 conditions: six variants of the occluder and two occluder colors. As in Experiment 1, the presentation order was randomized and the blocks were repeated in the ABBA order (24 blocks in total). See also Videos 7–10.

Experiment 3 also consisted of 12 blocks. All blocks and trials had the same visual display sequence (see Video 11). However, on every second and third trial, observers heard a tone played over the computer speakers that lasted either 1,000 or 1,500 ms and started, respectively, 600 or 850 ms before the SFM object reached the center of the screen. During the first four blocks, observers were instructed to ignore the sounds. In the following eight blocks, they were instructed to shut their eyes for the duration of the tone. Their fixation and blinking was monitored via an eyetracker. Observers reliably closed their eyes in response to the tone: Five observers missed only a single tone; four observers never missed a tone. They were also highly consistent in the duration of their eye closure, which was 955 ± 71 ms (mean \pm *SD*) for the 1,000-ms tone and $1,355 \pm 71$ ms for the 1,500-ms tone.

Statistical analysis

The statistical analysis was performed in R (R Core Team, 2016) using the BayesFactor package (Morey & Rouder, 2015) for Bayesian linear mixed models, packages lme4 (Bates, Mächler, Bolker, & Walker, 2015) and lmerTest (Kuznetsova, Bruun Brockhoff, & Haubo Bojesen Christensen, 2016) for linear mixed models analysis, and package ggplot2 (Wickham, 2009) to generate the figures. Eye movement data were processed using the edfImport toolbox (Pastukhov, 2017).

Data availability

All data files, along with the code used to perform the statistical analyses and produce the figures, are available under a Creative Commons Attribution 4.0 International Public License at <https://osf.io/qqrzp>.

Results

Experiment 1: Solid occluder

In our first experiment, we investigated whether a representation of a moving object persisted when it was partially or completely occluded. To this end, we used an ambiguously rotating SFM object and examined whether the bistable rotation would be destabilized after the object passed behind the occluder. To compare the persistence of fully and partially occluded objects, we systematically manipulated the width of the SFM shape from an object-to-occluder ratio of 1:1 (the SFM object was fully occluded for a single frame) to 2:1 (the occluding object covered no more than half the width of the moving object; see also Fig. 1f for information on the duration of the partial-occlusion episodes). Because prior work had indicated that the object’s persistence is facilitated by a physically plausible and congruent event sequence (Flombaum & Scholl, 2006), we used a cardboard strip rather than a computer-generated image to occlude the central part of the screen. To ensure the variability of perceptual states, we attempted to reverse the rotation by inverting the vertical component of 2-D motion at the beginning of each trial. The exogenous trigger was effective in $82\% \pm 17\%$ (mean ± 1 SD) of the trials (see Pastukhov et al., 2012, for details about the method).

To further facilitate the persistence of the occluded object, we added two additional cues that indicated its continued presence. The first was a colored “halo” around the object. This halo was wider than the occluding strip and was therefore always at least partially visible. The second was a panning sound that moved congruently from side to side with the SFM object. Neither cue was informative about the dominant direction of rotation. In total, we used four conditions: no cues, a single cue, or both cues present. For further details, please see Fig. 1a and b and Videos 1–6.

To quantify the persistence versus destabilization of rotation, we computed the probability of participants reporting the change in direction of rotation shortly after the object had passed the center of the screen (the time point of minimal visibility; please see Fig. 1f):

$$P_{\text{switch}} = \frac{N_{\text{switch}}}{N_{\text{total}}}, \tag{1}$$

with N_{total} as the total number of trials and N_{switch} as the number of trials in which observers reported a perceptual switch between 200 and 800 ms after the object had passed the center of the screen. We picked this response time interval because prior work had indicated that it should contain the most reports on perceptual changes (Pastukhov et al., 2012). Extending this interval would increase the observed destabilization without qualitatively altering the results. However, for those longer

periods the perceived rotation also depended on spatially specific biases, making it harder to disentangle their relative contributions (see below).

To assess the influences of individual factors, we performed both the multilevel linear mixed-effects models (Bates et al., 2015) and a mixed-model Bayesian analysis of variance (ANOVA; Bolstad & Curran, 2016; Morey & Rouder, 2015), with object width, the presence of the halo and the sound, and their interaction as independent factors, and participants as a nested random effect. Maximum likelihood was used as an estimation method, to allow for between-model comparisons via ANOVA.

The results of Experiment 1 are presented in Fig. 2 and Table 1. We found that the presence of neither the halo nor the sound affected the perceptual stability of rotation. In contrast, the width of the SFM object had a strong and significant effect on persistence ($R^2 = .533$ for the linear mixed-effect model with object width as a single independent factor, assuming correlated random intercepts and slopes). For most observers, perceptual destabilization was strongly and negatively correlated with the object’s size (i.e., smaller width ratios led to stronger destabilization): Spearman’s ρ was $-.77$ [$-.81, -.64$] (median [1st, 3rd quartiles]).

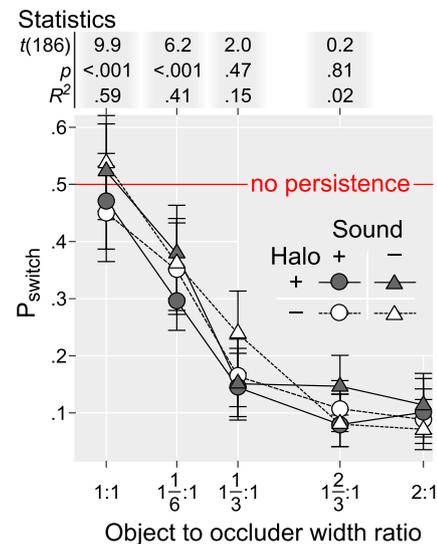


Fig. 2 Experiment 1: Effects of the object-to-occluder width ratio and the presence of the halo and/or the panning sound on the perceptual stability of rotation. Error bars depict a 95% binomial confidence interval assuming the group’s mean performance and the total number of trials. $P_{\text{switch}} = .5$, labeled “no persistence,” shows the probability of the switch that corresponds to no persistence. Values below that line indicate persistence, and values above that line indicate consistent switching. The values above the plot depict, from top to bottom, the t statistics (Satterthwaite approximations to the degrees of freedom), the corresponding p values, and the effect sizes when comparing P_{switch} for the corresponding width ratio and that for the width ratio of 2:1. The comparison was performed using a linear mixed-effect model with object width as an independent factor and observer identity as a nested random effect.

Table 1 Multilevel linear mixed-effect models and the mixed-effect Bayesian ANOVA with observer identity as a nested random effect, and object width, the presence of the halo or the sound, as well as their interactions as independent factors

Model	<i>df</i>	AIC	BIC	Log-Likelihood	χ^2	<i>p</i> Value	Bayes Factor
Observer	3	22.3	32	− 8.46			
+ Width	6	− 56.0	− 43	32.01	80.97	<.001	221,475 ± 0.9%
+ Sound	7	− 55.7	− 39	32.82	1.62	.20	35,288 ± 1.4%
+ Halo	8	− 53.7	− 34	32.84	0.02	.88	18,500 ± 3.1%
+ Sound × Width	9	− 52.4	− 29	33.22	0.76	.38	1,973 ± 3.1%
+ Halo × Width	10	− 50.8	− 24	33.39	0.35	.56	238 ± 2.9%
+ Sound × Halo	11	− 48.9	− 19	33.45	0.11	.74	14 ± 2.3%

The Bayes factor was computed relative to the model with random effects only. χ^2 was computed relative to the preceding simpler model. *df*: degrees of freedom; AIC: Akaike's information criterion; BIC: Schwarz's Bayesian information criterion

The reduced persistence for lower width ratios could be due to the general impoverishment of sensory evidence (i.e., fewer dots were visible during the occlusion episode; see Fig. 1f) or to the lack of grouping for the dots on the two sides of the cardboard strip. It is likely that for larger objects this grouping allowed the visual system to bridge the gap and extend the currently dominant perceptual state to the dots on the other side. For the smaller objects, such as those with the width ratio of $1\frac{1}{6} : 1$, the number of dots simultaneously present on both sides could have been too small to allow for effective grouping, leading to reduced stability even though the strip width was smaller than a typical receptive field of neurons in hMT+ (Amano, Wandell, & Dumoulin, 2009). Finally, for the object with a 1:1 width ratio, the dots never appeared simultaneously on both sides, possibly providing little evidence for the recruitment of spatially distant neural representations across the strip. Taken together, our results indicate that the persistence of illusory depth and, possibly, of an overall object representation critically depends on the presence of uninterrupted and reliable sensory evidence.

Experiment 2: Partial occluder

Our first experiment showed that the persistence of the rotation was proportional to the area of the object still visible when the cardboard maximally occluded it. As we noted above, this dependence could reflect impoverished evidence for the object (i.e., fewer dots visible throughout the occlusion) or possibly, in addition, reduced grouping between the individual dots on the two sides of the cardboard strip. To disentangle these two hypotheses, we repeated the experiment but used a fixed-size SFM object (object-to-occluder ratio of 1:1) and a computer-generated occluder that contained rectangular apertures. These apertures covered a certain fraction of the occluder. Six conditions were used in total: a solid occluder (no apertures; effectively, a replication of Exp. 1 with a computer-generated occluding object instead of the cardboard

strip), 5%, 10%, 25%, or 50% of the total area occluded, and no occluder. The apertures diminished the number of visible dots. However, the small distance between the individual apertures was designed to facilitate grouping, and hence stabilize the perception.

Because prior work had indicated that the visibility of the occluder has a profound effect on the grouping of individual motion components (McDermott & Adelson, 2004), we used the same occluder twice, once as a visible occluder (yellow color; Fig. 1c) and once as a camouflaged occluder (same color as the background, so that its presence was evident only while the SFM object was passing behind it; see Fig. 1d and Videos 7–10). Otherwise, the procedure, the computed observables, and the general analysis were identical to those aspects of Experiment 1.

As in Experiment 1, the complete occlusion strongly destabilized rotation, irrespective of whether the occluder was visible (Fig. 3). Aperture area had the strongest impact on the probability of survival (effect size $R^2 = .40$; see also Table 2). However, the lack of a significant decrease in perceptual stability for the aperture areas of 10% and above indicates that the large spatial separation in Experiment 1 played a crucial role in grouping, leading to the perceptual destabilization.

The visibility of the occluding object might have had an influence on the smaller (5% and 10%) aperture area conditions, in which a visible occluder appeared to facilitate the stability of rotation. The effect size of occluder visibility for 5% and 10% apertures was only moderate ($R^2 = .435$) and failed to reach significance, $\chi^2(1) = 1.7$, $p = .2$ (for a linear mixed model with aperture size and occluder visibility as fixed effects and observer identity as a nested random effect vs. a similar model without the occluder visibility factor).

To summarize, we found that a lack of spatial overlap between successive locations (0% aperture condition), as well as impoverished sensory evidence (5% aperture condition), strongly destabilized the perception of bistable rotation.

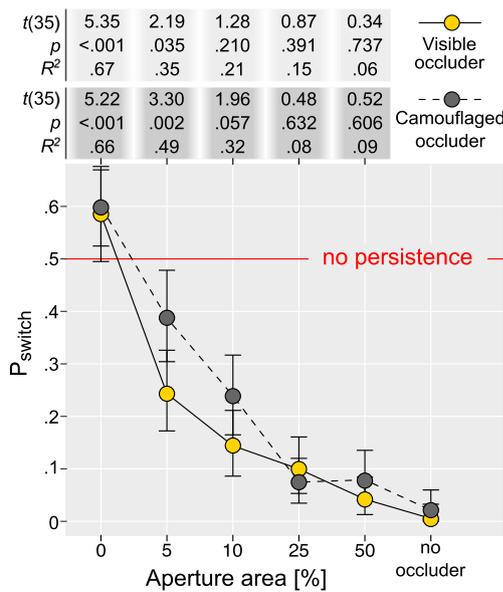


Fig. 3 Experiment 2: Effects of occluder visibility and of total aperture area. Error bars depict 95% binomial confidence intervals assuming the group’s mean performance and the total number of trials. $P_{switch} = .5$, labeled “no persistence,” shows the probability of a switch that corresponds to no persistence. Values below that line indicate persistence, and values above that line indicate consistent switching. The values above the plot depict t statistics, the corresponding p values, and the effect sizes when comparing P_{switch} for the corresponding aperture area versus the no-occluder condition (linear mixed-effect model with total aperture area and occluder visibility as independent factors and observer identity as a nested random effect)

Experiment 3: Blinking

Our first experiments demonstrated that complete occlusion strongly destabilized rotation. We wondered whether this reflected a lack of persistence mechanisms or a lack of their activation. To clarify this further, we repeated the measurement but relied on eye closure to render the moving object temporarily invisible. Blinking constitutes one of the most common causes for interruptions in sensory evidence (Volkman, Riggs, & Moore, 1980) and is an endogenously generated event. This means that the system has full knowledge of why and when evidence for the object’s presence is disturbed and has the best opportunity to employ a mechanism

for perceptual stabilization. On the one hand, this suggestion is supported by earlier studies that demonstrated a profound stabilizing effect of long eye closure on the perception of rotation, albeit for a stationary object (Leopold et al., 2002), as well as predictive remapping of an object’s features to a new spatial location before a saccade (Melcher, 2007). On the other hand, research on change blindness indicates the same lack of persistence for blinks as for blanks (O’Regan, Deubel, Clark, & Rensink, 2000).

The display and procedure were identical to those of Experiments 1 and 2. However, the moving object was never exogenously occluded, and the observers heard a tone played on every second and third trial (the tone duration was either 1,000 or 1,500 ms). In the first four blocks, the observers were told to ignore the tone, whereas in the following eight blocks the participants were instructed to keep their eyes shut while the tone was playing. Blinking and the accuracy of fixation were monitored via an eyetracker, and the observers proved to be highly reliable. Just five out of nine observers failed to close their eyes for a single tone, and their timing was highly consistent (the eye closure time was 955 ± 71 ms (mean \pm SD) for the 1,000-ms tone and $1,355 \pm 71$ ms for the 1,500-ms tone).

For the analysis, we divided the trials into four types (please see the upper table in Fig. 4a for information about their relative frequencies). *Control* trials contained neither sounds nor blinks. Here, the postevent time window was between 200 and 800 ms after the object had passed the center of the screen, making it identical to the no-occluder condition in Experiment 2. *Sound* trials included the tone but no blinks, controlling for the potentially destabilizing effect of the sound alone. The tone was considered to trigger the switch if a report occurred between 200 and 800 ms after the sound onset. *Prompted blinks* were trials in which the observers shut their eyes in response to the tone, whereas *spontaneous blinks* were all other trials that contained unprompted blinks (i.e., blinks in the absence of the tone). Approximately 7% of the trials contained multiple blinks and were excluded from the analysis. For both types of blinks, the postevent time window was set between 200 and 800 ms after the participant had eyes open. We found a small but statistically significant systematic

Table 2 Repeated measures Bayesian ANOVA and a linear mixed-effect model with total aperture area, occluder visibility, and the interaction between total aperture area and visibility as independent factors, and with observer identity as a nested random effect

Model	df	AIC	BIC	Log-Likelihood	χ^2	p Value	Bayes Factor
Observer	3	47	54	-20.4			
+ Aperture area	4	21	31	-6.5	27.80	<.001	241,137 \pm 0.43%
+ Occluder visibility	5	22	35	-6.1	0.88	.35	152,935 \pm 7.12%
+ Aperture area \times Visibility	6	24	39	-6.0	0.18	.67	18,500 \pm 1.06%

The Bayes factor was computed relative to the model with random effects only. χ^2 was computed relative to the preceding simpler model. df : degrees of freedom; AIC: Akaike’s information criterion; BIC: Schwarz’s Bayesian information criterion

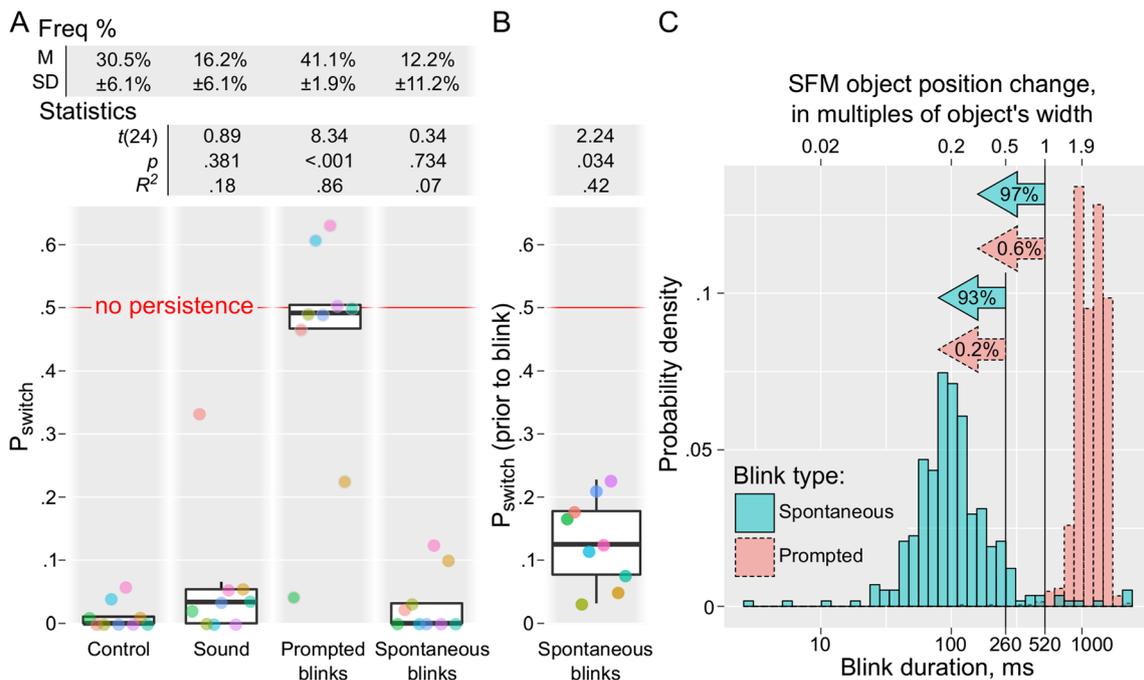


Fig. 4 Experiment 3. (a) Probabilities of a switch being reported between 200 and 800 ms after the event (see the text for details). Circle colors label the individual observers. The Statistics table above the plot depicts the *t* statistics, the corresponding *p* values, and the effect sizes when comparing P_{switch} to the control-condition occluder (linear mixed-effect model with event type as an independent factor and observer identity as a nested random effect). $P_{switch} = .5$, labeled “no persistence,” shows the probability of a switch that corresponds to no persistence. Values below that line indicate persistence, and values above that line indicate

consistent switching. The Freq table summarizes the relative frequencies of the individual trial types across all blocks (please note that ~7% of the trials contained multiple blinks and were excluded from the analysis). (b) Probabilities of the switch being reported within 600 ms prior to the spontaneous blink. The same statistical comparisons were done as in panel A. (c) Distributions of durations for spontaneous and prompted blinks. The arrows show the percentages of spontaneous and prompted blinks that were shorter than the time necessary for an object to travel its half and its full width, respectively

shift of eye gaze toward the anticipated location of the object for the prompted blinks: $0.83^\circ \pm 0.65^\circ$ (mean and *SEM*), $\chi^2(1) = 13.8$, $p = .0002$, $R^2 = .087$ (for a linear mixed model of the gaze shift during the blink, with object motion direction as a fixed factor and observer identity as a random factor). No systematic shift was observed for spontaneous blinks: $-0.05^\circ \pm 0.13^\circ$ ($M \pm SEM$), $\chi^2(1) = 0.5$, $p = .48$, $R^2 = .074$ (for a linear mixed model of the gaze shift during the blink, with object motion direction as a fixed factor and observer identity and block condition as random factors).

The results of Experiment 3 are summarized in Fig. 4a. We found that only prompted blinks significantly destabilized the perception of rotation. This effect did not depend on the duration of eye closure, $\chi^2(1) = .08$, $p = .78$. Thus, in contrast to the earlier report on stationary bistable displays (Leopold et al., 2002), we found that for the moving bistable SFM object, long, prompted blinks can be as destabilizing as the complete occlusion used in Experiments 1 and 2.

In contrast to the prompted blinks and complete occlusion in Experiments 1 and 2, spontaneous blinks produced very little destabilization (see Fig. 4a). Although this may indicate that spontaneous (but not prompted) blinks recruit memory mechanisms that maintain activity in the visual cortex (Hyo, Chung, Song, & Park, 2005), it must be noted that the

spontaneous blinks were very brief (102 ms [49.215 ms], geometric mean and confidence interval based on geometric standard deviation; see Fig. 4c). This means that, in contrast to the prompted blinks and the complete-occlusion events in Experiments 1 and 2, the SFM object reappeared at the location that overlapped with the object’s location before the blink. For 93% of all spontaneous blinks, the SFM display moved no more than a half-width during the blink, whereas it moved at least an entire width for virtually all prompted blinks (see Fig. 4c). Thus, it is very likely that the perception of rotation was stabilized by the lingering activity (i.e., neural persistence, discussed in the introduction) of the recently active and spatially overlapping neural populations (Pastukhov & Braun, 2013). Additional factors for the stability following spontaneous blinks are the generally low levels of perceptual adaptation that destabilize perception for stationary bistable displays (Blake et al., 2003; Pastukhov & Braun, 2011) and perceptual stabilization that occurs shortly before the blinks themselves (see Fig. 4b and Van Dam & Van Ee, 2005).

Experiment 1–3: Location-specific perceptual bias

Since our experimental procedure relied on the moving SFM object, this provided us with an opportunity to

examine location-specific biases for the perception of rotation. To this end, we divided the trajectory into 20 sub-intervals and computed the probability of observers reporting the upward direction of rotation for each interval:

$$P_{up} = \frac{N_{up}}{N_{up} + N_{down}}, \tag{2}$$

where N_{up} and N_{down} are the numbers of trials for which the reported percepts were, correspondingly, upward or downward in rotation. Trials with unclear perception were excluded from the analysis.

As can be seen in Fig. 5, the location-specific analysis showed divergent patterns of results, since the strength, direction, and location specificity of the bias varied greatly among the observers. Some of these location-specific biases were dynamic (e.g., observer SDA95m in Fig. 5), but some were remarkably stable (observer UKS89m; please note that Exps. 1 and 2 occurred on the same day, whereas Exp. 3 was conducted two months later). Our results match those found for other multistable displays (Carter & Cavanagh, 2007; Wexler, Duyck, & Mamassian, 2015) and provide further evidence for representation-specific influences on multistable perception.

Discussion

The main aim of the study was to investigate whether the physical plausibility of an occlusion/disappearance episode would facilitate the persistence of a temporarily occluded moving bistable display. To this end, we used a bistable rotating structure-from-motion object that repeatedly traversed the screen and either passed behind an occluding object (a cardboard strip in Exp. 1, a computer-generated image in Exp. 2) or was rendered temporarily invisible due to participants closing their eyes (Exp. 3). We observed that bistable rotation persisted only when the successive locations of the object along the trajectory were overlapping, as was the case for the occluder with apertures in Experiment 2 and for brief, spontaneous blinks in Experiment 3, or if the object was large enough to allow for the grouping of dots on both sides of the occluding strip (Exp. 1). In contrast, whenever the successive locations were nonoverlapping and spatially distant, we found no persistence (i.e., both directions of bistable rotation were equally likely to become dominant following the interruption). This was the case for the complete-occlusion conditions in Experiments 1 and 2 and the long, prompted blinks in Experiment 3. In short, we found no predictive remapping of a representation or of the features of a moving object onto an anticipated but nonoverlapping spatial location.

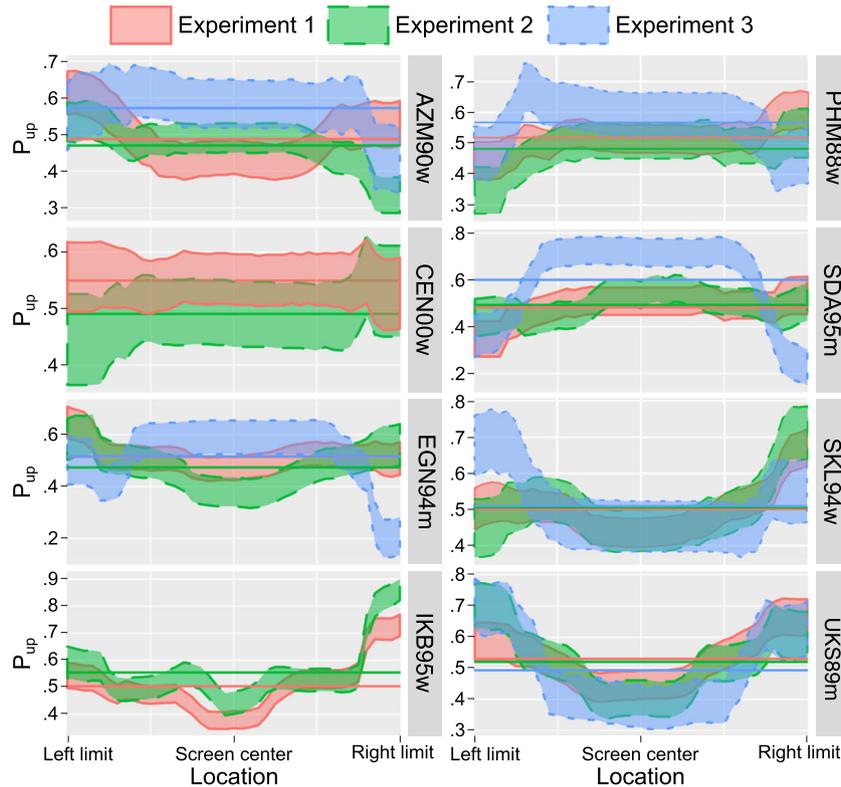


Fig. 5 Location-specific bias: Probability of an observer reporting the upward direction of rotation (P_{up}) as a function of location. The stripes denote a 95% binomial confidence intervals. The solid lines indicate probabilities of reporting the upward direction of rotation

Our results fit well with prior work on multistable displays moving through adjacent neural populations. In our case, as well as in Blake et al. (2003), perceptual dominance was passed on to a different neural representation that mapped an overlapping neighboring spatial location. Other cases included the adjacent axis of rotation for a “wobbling” SFM globe (Blake et al., 2003), object orientation for bistable kinetic depth (Pastukhov & Braun, 2013), grating orientation for a binocular rivalry display (Denison, Piazza, & Silver, 2011), and relative dot locations for ambiguous-motion quartets (Maloney, Martello, Sahm, & Spillmann, 2005). In all these cases, the dominant percept was “passed on” to the next neural representation, and the perception remained stable. However, this was not the case whenever the gap between two spatial locations (Blake et al., 2003) or between two orientations (Pastukhov & Braun, 2013) was too large. Our present study confirms these findings, showing both persistence (for an occluding object with apertures in Exp. 2 or for brief, spontaneous blinks in Exp. 3) and the lack of it (for the solid occluding objects in Exps. 1 and 2 and for long, prompted blinks in Exp. 3). However, the present results extend them by showing that this gap is not bridged by a predictive remapping of neural representations, even if the occluding event is physically plausible, ecologically valid, predictable (as in Exps. 1 and 2), and internally generated (as in Exp. 3). Thus, the lack of persistence most likely reflects an absence of such predictive neural mechanisms rather than our inability to tap into them.

The observed lack of persistence indicates that, although the location of an object may be tracked throughout an occlusion episode, as in the tunnel effect (Burke, 1952; Flombaum, Kundey, Santos, & Scholl, 2004; Flombaum & Scholl, 2006; Scholl & Pylyshyn, 1999; Tougas & Bregman, 1990), its properties may be predictively remapped only during active viewing (e.g., when making a saccade while viewing an object; Melcher, 2007), but not during passive viewing (fixating while viewing a moving object, as in the present study). The likely source of this difference could be a far lower confidence in the future object’s location in the latter case. In contrast to an endogenously generated saccade, a moving object may abruptly alter its velocity, rendering extrapolation based on its prior motion erroneous. Moreover, a moving object that is worth tracking, such as a predator or prey, is likely to exhibit deliberately random behavior and, therefore, motion. For example, animal studies consistently show that predictable behavior can be a profound disadvantage once your strategy has been found out (Lee, Conroy, McGreevy, & Barraclough, 2004; Lee, McGreevy, & Barraclough, 2005). Conversely, assuming regular, predictable motion for a potentially unpredictably behaving agent would mean that your predictions might be wrong more often than not. In this case, making no predictions and relying on the immediate visual evidence alone

may prove to be a better and safer strategy, at least at the level of perceptual representation.

Finally, we would note the slightly different natures of the destabilization observed for stationary versus moving multistable displays. In the former case, destabilization is strongest for blank durations of approximately half a second (Klink et al., 2008; Kornmeier & Bach, 2004; Orbach, Ehrlich, & Heath, 1963; Pastukhov & Braun, 2013). Here, perception was pushed away from the previously dominant state by the habituation of its neural representation (Noest, van Ee, Nijs, & van Wezel, 2007; Wolfe, 1984). For shorter interruptions, the destabilizing effect of adaptation is partially mitigated by neural persistence (Pastukhov & Braun, 2013), whereas longer interruptions allow for recovery from perceptual adaptation and reveal a weak facilitating effect of the sensory memory of multistable displays (Adams, 1954; Leopold et al., 2002; Orbach, Ehrlich, & Vainstein, 1963). In our case, the movement of the bistable SFM object minimizes the buildup of adaptation (Blake et al., 2003); thus, destabilization reveals a lack of persistence for the recently dominant percept at a new location. In this case, the perceptual decision is dominated by location-specific biases and location-specific memories (Knapen, Brascamp, Adams, & Graf, 2009). Perhaps the latter processes were unable to bridge the gap between the distant locations used in this study.

Conclusions

We report that a complete interruption of sensory evidence of an object’s continued existence strongly and significantly destabilized bistable rotation, as long as the object reappeared at a nonoverlapping location, thus engaging previously unaccessed neural representations.

References

- Adams, P. A. (1954). The effect of past experience on the perspective reversal of a tridimensional figure. *American Journal of Psychology*, *67*, 708–710. <https://doi.org/10.2307/1418496>
- Amano, K., Wandell, B. A., & Dumoulin, S. O. (2009). Visual field maps, population receptive field sizes, and visual field coverage in the human MT+ complex. *Journal of Neurophysiology*, *102*, 2704–2718. <https://doi.org/10.1152/jn.00102.2009>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using (lme4). *Journal of Statistical Software*, *67*, 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Blake, R., Sobel, K. V., & Gilroy, L. A. (2003). Visual motion retards alternations between conflicting perceptual interpretations. *Neuron*, *39*, 869–878. <https://doi.org/10.1016/S0896-627300495-1>
- Bolstad, W. M., & Curran, J. M. (2016). *Introduction to Bayesian statistics*. Hoboken: Wiley.

- Burke, L. (1952). On the tunnel effect. *Quarterly Journal of Experimental Psychology*, 4, 121–138. <https://doi.org/10.1080/17470215208416611>
- Carter, O., & Cavanagh, P. (2007). Onset rivalry: Brief presentation isolates an early independent phase of perceptual competition. *PLoS ONE*, 2, e343. <https://doi.org/10.1371/journal.pone.0000343>
- Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics*, 27, 183–228. <https://doi.org/10.3758/BF03204258>
- Denison, R. N., Piazza, E. A., & Silver, M. A. (2011). Predictive context influences perceptual selection during binocular rivalry. *Frontiers in Human Neuroscience*, 5, 166:1–11. <https://doi.org/10.3389/fnhum.2011.00166>
- Flombaum, J. I., Kundery, S. M., Santos, L. R., & Scholl, B. J. (2004). Dynamic object individuation in rhesus macaques: A study of the tunnel effect. *Psychological Science*, 15, 795–800. <https://doi.org/10.1111/j.0956-7976.2004.00758.x>
- Flombaum, J. I., & Scholl, B. J. (2006). A temporal same-object advantage in the tunnel effect: Facilitated change detection for persisting objects. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 840–853. <https://doi.org/10.1037/0096-1523.32.4.840>
- Hupé, J.-M., Lamirel, C., & Lorenceau, J. (2009). Pupil dynamics during bistable motion perception. *Journal of Vision*, 9(7), 10. <https://doi.org/10.1167/9.7.10>
- Hyo, W. Y., Chung, J. Y., Song, M. S., & Park, H. W. (2005). Neural correlates of eye blinking: Improved by simultaneous fMRI and EOG measurement. *Neuroscience Letters*, 381, 26–30. <https://doi.org/10.1016/j.neulet.2005.01.077>
- Kawachi, Y., & Gyoba, J. (2006). A new response-time measure of object persistence in the tunnel effect. *Acta Psychologica*, 123, 73–90. <https://doi.org/10.1016/j.actpsy.2006.04.003>
- Klink, P. C., van Ee, R., Nijs, M. M., Brouwer, G. J., Noest, A. J., & van Wezel, R. J. A. (2008). Early interactions between neuronal adaptation and voluntary control determine perceptual choices in bistable vision. *Journal of Vision*, 8(5), 16.1–18. <https://doi.org/10.1167/8.5.16>
- Knäpen, T. H. J., Brascamp, J. W., Adams, W. J., & Graf, E. W. (2009). The spatial scale of perceptual memory in ambiguous figure perception. *Journal of Vision*, 9(13), 16.1–12. <https://doi.org/10.1167/9.13.16>
- Kornmeier, J., & Bach, M. (2004). Early neural activity in Necker-cube reversal: Evidence for low-level processing of a gestalt phenomenon. *Psychophysiology*, 41, 1–8. <https://doi.org/10.1046/j.1469-8986.2003.00126.x>
- Kuznetsova, A., Bruun Brockhoff, P., & Haubo Bojesen Christensen, R. (2016). *lmerTest: Tests in linear mixed effects models*. Retrieved from <https://cran.r-project.org/package=lmerTest>
- Lee, D., Conroy, M. L., McGreevy, B. P., & Barraclough, D. J. (2004). Reinforcement learning and decision making in monkeys during a competitive game. *Cognitive Brain Research*, 22, 45–58. <https://doi.org/10.1016/j.cogbrainres.2004.07.007>
- Lee, D., McGreevy, B. P., & Barraclough, D. J. (2005). Learning and decision making in monkeys during a rock–paper–scissors game. *Brain Research. Cognitive Brain Research*, 25, 416–30. <https://doi.org/10.1016/j.cogbrainres.2005.07.003>
- Leopold, D. A., Wilke, M., Maier, A., & Logothetis, N. K. (2002). Stable perception of visually ambiguous patterns. *Nature Neuroscience*, 5, 605–9. <https://doi.org/10.1038/nn851>
- Maier, A., Wilke, M., Logothetis, N. K., & Leopold, D. A. (2003). Perception of temporally interleaved ambiguous patterns. *Current Biology*, 13, 1076–1085. <https://doi.org/10.1016/S0960-982200414-7>
- Maloney, L. T., Martello, M. F. D., Sahn, C., & Spillmann, L. (2005). Past trials influence perception of ambiguous motion quartets through pattern completion. *Proceedings of the National Academy of Sciences*, 102, 3164–3169. <https://doi.org/10.1073/pnas.0407157102>
- McDermott, J., & Adelson, E. H. (2004). The geometry of the occluding contour and its effect on motion interpretation. *Journal of Vision*, 4, 944–954. <https://doi.org/10.1167/4.10.9>
- Melcher, D. (2007). Predictive remapping of visual features precedes saccadic eye movements. *Nature Neuroscience*, 10, 903–907. <https://doi.org/10.1038/nm1917>
- Morey, R. D., & Rouder, J. N. (2015). *BayesFactor: Computation of Bayes factors for common designs*. Retrieved from <https://cran.r-project.org/package=BayesFactor>
- Mossbridge, J. A., Ortega, L., Grabowecky, M., & Suzuki, S. (2013). Rapid volitional control of apparent motion during percept generation. *Attention, Perception, & Psychophysics*, 75, 1486–1495. <https://doi.org/10.3758/s13414-013-0504-3>
- Noest, A. J., van Ee, R., Nijs, M. M., & van Wezel, R. J. A. (2007). Percept-choice sequences driven by interrupted ambiguous stimuli: A low-level neural model. *Journal of Vision*, 7(8), 10. <https://doi.org/10.1167/7.8.10>
- O’Regan, J. K., Deubel, H., Clark, J. J., & Rensink, R. A. (2000). Picture changes during blinks: Looking without seeing and seeing without looking. *Visual Cognition*, 7, 191–211. <https://doi.org/10.1080/135062800394766>
- Orbach, J., Ehrlich, D., & Heath, H. A. (1963). Reversibility of the Necker cube: I. An examination of the concept of “satiation of orientation.” *Perceptual and Motor Skills*, 17, 439–458. <https://doi.org/10.2466/pms.1963.17.2.439>
- Orbach, J., Ehrlich, D., & Vainstein, E. (1963). Reversibility of the Necker cube: III. Effects of interpolation on reversal rate of the cube presented repetitively. *Perceptual and Motor Skills*, 17, 571–582. <https://doi.org/10.2466/pms.1963.17.2.571>
- Orban, G. A. (2011). The extraction of 3-D shape in the visual system of human and nonhuman primates. *Annual Review of Neuroscience*, 34, 361–388. <https://doi.org/10.1146/annurev-neuro-061010-113819>
- Pastukhov, A. (2017). *edfImport: Matlab interface to Eyelink EDF files*. Open Science Framework. Retrieved January 2017 from <https://osf.io/xfumn>. <https://doi.org/10.17605/OSF.IO/FXUMN>
- Pastukhov, A., & Braun, J. (2011). Cumulative history quantifies the role of neural adaptation in multistable perception. *Journal of Vision*, 11(10), 12. <https://doi.org/10.1167/11.10.12>
- Pastukhov, A., & Braun, J. (2013). Structure-from-motion: Dissociating perception, neural persistence, and sensory memory of illusory depth and illusory rotation. *Attention, Perception, & Psychophysics*, 75, 322–340. <https://doi.org/10.3758/s13414-012-0390-0>
- Pastukhov, A., Vonau, V., & Braun, J. (2012). Believable change: Bistable reversals are governed by physical plausibility. *Journal of Vision*, 12(1), 17. <https://doi.org/10.1167/12.1.17>
- Pearson, J., & Brascamp, J. W. (2008). Sensory memory for ambiguous vision. *Trends in Cognitive Sciences*, 12, 334–41. <https://doi.org/10.1016/j.tics.2008.05.006>
- R Core Team. (2016). *R: A Language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.r-project.org/>
- Scholl, B. J., & Pylyshyn, Z. W. (1999). Tracking multiple items through occlusion: Clues to visual objecthood. *Cognitive Psychology*, 38, 259–290. <https://doi.org/10.1006/cogp.1998.0698>
- Song, C., & Yao, H. (2009). Duality in binocular rivalry: Distinct sensitivity of percept sequence and percept duration to imbalance between monocular stimuli. *PLoS ONE*, 4, e6912. <https://doi.org/10.1371/journal.pone.0006912>
- Tougas, Y., & Bregman, A. S. (1990). Auditory streaming and the continuity illusion. *Perception & Psychophysics*, 47, 121–126. <https://doi.org/10.3758/BF03205976>
- Van Dam, L. C. J., & Van Ee, R. (2005). The role of (micro)saccades and blinks in perceptual bi-stability from slant rivalry. *Vision Research*, 45, 2417–2435. <https://doi.org/10.1016/j.visres.2005.03.013>

- Volkman, F., Riggs, L., & Moore, R. (1980). Eyeblinks and visual suppression. *Science*, *207*, 900–902. <https://doi.org/10.1126/science.7355270>
- Wexler, M., Duyck, M., & Mamassian, P. (2015). Persistent states in vision break universality and time invariance. *Proceedings of the National Academy of Sciences*, *112*, 14990–14995. <https://doi.org/10.1073/pnas.1508847112>
- Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. New York: Springer. Retrieved from ggplot2.org
- Wolfe, J. M. (1984). Reversing ocular dominance and suppression in a single flash. *Vision Research*, *24*, 471–478. [https://doi.org/10.1016/0042-6989\(84\)90044-0](https://doi.org/10.1016/0042-6989(84)90044-0)