Perception and the strongest sensory memory trace of multi-stable displays both form shortly after the stimulus onset

Alexander Pastukhov[1,2,3]

Abstract We investigated the relation between perception and sensory memory of multi-stable structure-from-motion displays. The latter is an implicit visual memory that reflects a recent history of perceptual dominance and influences only the initial perception of multi-stable displays. First, we established the earliest time point when the direction of an illusory rotation can be reversed after the display onset (29–114 ms). Because our display manipulation did not bias perception towards a specific direction of illusory rotation but only signaled the change in motion, this means that the perceptual dominance was established no later than 29–114 ms after the stimulus onset. Second, we used orientation-selectivity of sensory memory to establish which display

orientation produced the strongest memory trace and when this orientation was presented during the preceding prime interval (80–140 ms). Surprisingly, both estimates point towards the time interval immediately after the display onset, indicating that both perception and sensory memory form at approximately the same time. This suggests a tighter integration between perception and sensory memory than previously thought, warrants a reconsideration of its role in visual perception, and indicates that sensory memory could be a unique behavioral correlate of the earlier perceptual inference that can be studied post hoc.

Keywords Multi-stable perception · Sensory memory · Perceptual inference · Implicit visual memory · Structure-from-motion · Perceptual memory · Kinetic-depth effect

✉ Alexander Pastukhov
pastukhov.alexander@gmail.com;
http://www.uni-bamberg.de/allgpsych

[1] Department of General Psychology and Methodology, Otto-Friedrich-Universität Bamberg, Markusplatz 3, 96047 Bamberg, Germany

[2] Center for Behavioral Brain Sciences, Magdeburg, Germany

[3] Cognitive Biology, Otto-von-Guericke Universität, Magdeburg, Germany

Introduction

Multi-stable displays are visual stimuli that are compatible with several distinct perceptual interpretations and, therefore, produce different and alternating perception despite constant sensory inputs. They serve as a highly sensitive psychophysical tool (Leopold & Logothetis, 1999) and, among other things, helped to better characterize numerous implicit visual memories (Leopold, Wilke, Maier, & Logothetis, 2002; Nawrot & Blake, 1991; Pastukhov & Braun, 2008, 2013a, b; Pastukhov, Lissner, Füllekrug, & Braun, 2014). One particularly interesting implicit visual memory they uncovered is sensory memory of multi-stable displays (Adams, 1954; Leopold et al., 2002; Orbach, Ehrlich, & Heath, 1963; Pastukhov & Braun, 2008), a facilitatory memory that can be detected only with multi-stable displays (de Jong, Knapen, & van Ee, 2012; Pastukhov & Braun, 2013a; Sterzer & Rees, 2008). Sensory memory traces

for competing percepts are thought to accumulate independently during periods of perceptual dominance (Brascamp et al., 2008; Leopold et al., 2002; Pastukhov & Braun, 2008), but to influence only the initial state of multi-stable perception at the display onset (de Jong et al., 2012).

Currently, the functional contribution of sensory memory to vision is not well understood. Because it facilitates repeated perception of the same state, it has been suggested to play a role in maintenance of perceptual stability (Knapen, Brascamp, Adams, & Graf, 2009; Pearson & Brascamp, 2008). However, what is already known about properties of sensory memory justifies skepticism about the practical role of this predictive facilitatory effect. First, it is very weak and is detected only with multi-stable displays (de Jong et al., 2012; Pastukhov & Braun, 2013a; Sterzer & Rees, 2008), the nature of which should preferably be known to observers (Oh, 2011; Rock & Mitchener, 1992). Second, multi-stable displays must be presented intermittently and repeatedly, as the initial perceptual stabilization is better explained by an observer-specific bias (Carter & Cavanagh, 2007). Third, during these multiple presentations the multi-stable display should remain mostly unchanged (Chen & He, 2004; Maier, Wilke, Logothetis, & Leopold, 2003; Pastukhov, Füllekrug, & Braun, 2013; Pastukhov et al., 2014) and be presented at the same *retinal* location (Chen & He, 2004; Knapen et al., 2009). Finally, blank periods in-between presentations must be relatively long (>1 s), because shorter interruptions lead to the onset perception being determined primarily by neural persistence or fatigue (Kornmeier, Ehm, Bigalke, & Bach, 2007; Orbach et al., 1963; Pastukhov & Braun, 2013a, b). However, one must note that required blank intervals are several times longer than is typical interruptions produced by either eye blinks or saccades (Baloh, Sills, Kumley, & Honrubia, 1975; (Volkmann, Riggs, & Moore, 1980). Taken together, these constraints are unlikely to be routinely satisfied outside of the laboratory environment.

Therefore, we sought to clarify the role of sensory memory in visual perception and their mutual influence. To this end, we investigated when perceptual dominance and the strongest sensory memory are formed. Because sensory memory influences only the initial perception at the onset of multi-stable displays (de Jong et al., 2012), knowledge about when the perceptual dominance is resolved would also give us an estimate of the only time interval when sensory memory has an effect on it. Conversely, knowledge about when the strongest sensory memory trace is formed and which part of a perceptual dominance history it favors would make it easier to evaluate various hypotheses on its functional role.

We report that for the multi-stable illusory rotation in structure-from-motion, perception and the strongest sensory memory trace both form shortly after the display onset. The fact that both form at approximately the same time indicates that the integration between perception and sensory memory may

be tighter than previously thought. We argue that these findings may require a reinterpretation of the role of sensory memory in visual perception.

General method

Observers

Nine observers participated in experiments. The author participated only in pilot experiments and his data were excluded from the analysis. All the data from pilot experiments were discarded. All participants had normal or corrected-to-normal vision. Observers were naive to the purpose of the experiments and were paid for their participation. Procedures were in accordance with the Declaration of Helsinki and were approved by the medical ethics board of the Otto-von-Guericke Universität, Magdeburg: "Ethik-Komission der Otto-von-Guericke-Universität an der Medizinischen Fakultät".

Apparatus

Stimuli were generated with MATLAB using the Psychophysics Toolbox (Brainard, 1997). Stimuli were displayed on a CRT screen (Iiyama VisionMaster Pro 514, iiyama.com) with a spatial resolution of 1600×1200 pixels and a refresh rate of 100 Hz. The viewing distance was 73 cm, so that each pixel subtended approximately 0.019°. In all experiments, background luminance was kept at 36 cd/m^2. The experimental room was lit dimly (ambient luminance at 80 cd/m^2).

Experiment 1: Estimating when the perception of illusory rotation is formed

In the first experiment, we sought to estimate when the perception of illusory rotation for structure-from-motion (SFM) is formed, as this is the only time interval when sensory memory appears to influence it (de Jong et al., 2012).

Our original approach (data not shown) was to present the SFM display for intervals of various durations (20–320 ms) and ask observers to judge whether they have perceived the 3D illusory rotation rather than a 2D motion. Unfortunately, even highly experienced observers (including the author) found making judgments of "how 3D the motion is" extremely difficult for brief presentations (20–80 ms). In these cases, one can be very confident of seeing motion, but judging whether it is 2D or 3D is very tricky and makes one feel one is guessing rather than knowing.

To sidestep this issue, we estimated when the perception of ambiguous illusory rotation is established by determining the earliest moment at which it is possible to alter it. This simplified the task for observers: an ambiguously rotating sphere

was presented intermittently for at least 320 ms and observers were instructed to report only the final direction of illusory rotation. In contrast to the original task of inferring whether motion was 2D or 3D, reporting on the direction of illusory rotation is straightforward and can be carried out with high confidence by observers. Note that they were explicitly told to report any unclear perception (e.g., two dot planes moving in opposite directions, random motion, etc.), but exercised this option only in ~1 % of all trials.

The display manipulation that is central to the measurement is a sudden inversion of the on-screen motion, schematically illustrated in Fig. 1 (see also Movie 2). It provides a transient signal for the change in motion, irrespective of which illusory rotation is currently dominant (Pastukhov, Vonau, & Braun, 2012; Stonkute, Braun, & Pastukhov, 2012). In contrast to some other methods that trigger perceptual reversals by temporarily biasing the display in favor of the suppressed percept (e.g., using stereoscopic depth for SFM (Pastukhov et al., 2012) or relative contrast for binocular rivalry (Blake, Westendorf, & Fox, 1990), this method does not introduce any state-specific bias, since the on-screen motion before and after the reversal is ambiguous in a balanced way. In other words, instead of prompting a switch towards the other specific percept, it prompts a switch *away from the current* percept, irrespective of what it and alternatives are. The most important consequence of this is that it is effective only *after* the perceptual dominance is established. Prior to that time point, it will only impede the ongoing perceptual inference by introducing a disturbance to the sensory evidence for *both* competing percepts. Accordingly, knowledge about the earliest time-point when the perceptual dominance can be reversed provides us with an upper boundary for the estimate of when it was established.

Method

Stimuli The SFM (Sperling & Dosher, 1994) display consisted of 50 dots (see Movies 1–2). Dots were distributed pseudo-randomly over the surface of an illusory sphere in such a way as to ensure maximal distance between each left-moving dot and the nearest right-moving dot, and vice versa (see Stonkute et al., 2012, for details). The diameter of the sphere was 5° and the rotation rate was 0.2 Hz. The diameter of a single dot was 0.057°, with a luminance of 110 cd/m^2. The dots were semi-transparent, i.e., the luminance of the overlap was a sum of individual luminance levels. This provided no clue on which dot is in front during the overlap to exclude any possible occlusion effects.

Procedure Experimental sessions included eight blocks of 30 trials each. Individual trials consisted of a biasing sequence and a test interval (see Fig. 2a–c). Observers used arrow keys to report the final direction of illusory rotation after each stimulus presentation. Note that observers were not informed about two different types of presentation intervals or about the significance of the planar motion inversion. In addition, within a single block all presentation intervals from all trials were presented in a single succession and observers were not informed about the type of interval (with or without the planar motion inversion) they were currently viewing. They also had the option to report whether the percept was unclear (~1 % of all responses).

During the biasing sequence (see Fig. 2a) observers viewed an intermittently presented, ambiguously rotating SFM sphere (T_{on} = 320 ms, T_{off} = $T_{response}$ + ISI = 1,060 ± 20 ms, where the average response time was $T_{response}$ = 312 ± 19 ms and randomized inter-stimulus interval was ISI = 500–1,000 ms) and reported the dominant direction of illusory rotation *at the end* of each presentation interval. Brief ON intervals precluded spontaneous perceptual reversals within each presentation (~1 % of trials), whereas long OFF intervals ensured a reliable biasing of the initial direction of illusory rotation by sensory memory (Adams, 1954; Leopold et al., 2002; Orbach et al., 1963; Pastukhov & Braun, 2008). The intermittent presentation of the ambiguously rotating sphere was repeated until the observer reported the same direction of illusory rotation on three consecutive trials (labelled as D_{bias}). The average biasing sequence length was 3.3 ± 0.13 trials. The biasing sequence produced the perceptual bias that was strong and allowed us to reliably predict the initial direction of illusory rotation for the test interval that followed immediately (bias reliability was verified using a "no inversion" condition, see *Results* below).

The test interval consisted of a single brief presentation of the ambiguously rotating sphere (Fig. 2b and c), which was
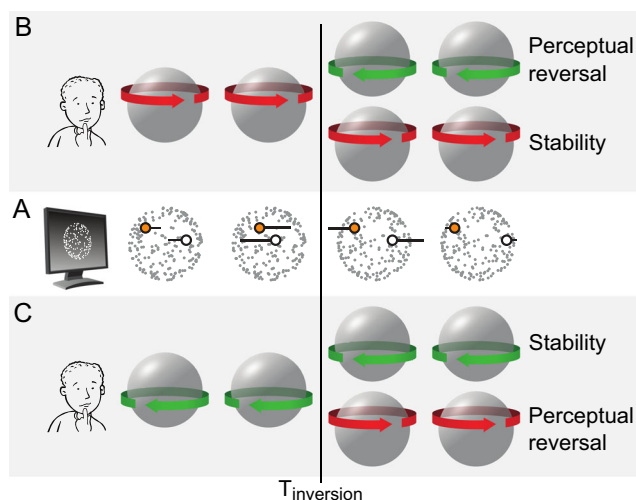

Fig. 1 Induced reversals of an illusory rotation, schematic procedure. (**a**) Static snapshots of physical displays with moving dots (schematic, frontal view, x-, y-plane). The on-screen velocity and its changes are indicated for two example dots. (**b**, **c**) Two alternative perceptual interpretations of the initial on-screen motion, with front surface of a sphere moving initially either to the right (**b**) or to the left (**c**). Following the inversion of the on-screen motion at time $T_{inversion}$, illusory rotation may reverse (outer outcomes) or remain stable (inner outcomes). The on-screen motion is fully ambiguous both before and after the inversion and the inversion itself is dominance-agnostic, i.e., its effectiveness does not depend on the initially dominant direction of illusory rotation
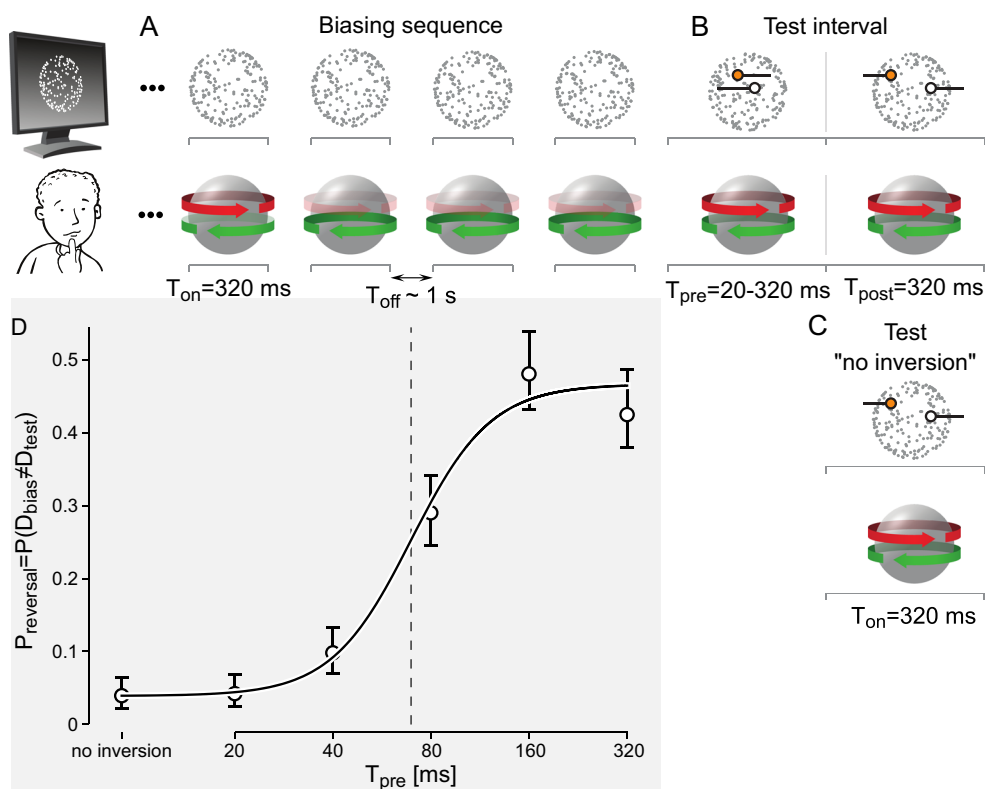
Fig. 2 Experiment 1. **a-c**) Schematic procedure for a single trial. **a**) The biasing sequence. An ambiguously rotating sphere was presented intermittently. Observers reported the *final* direction of illusory rotation. Presentation was repeated until the observer reported the same direction of illusory rotation (labelled as D_{bias}) on three consecutive presentations. **b**) During the test interval, the planar motion of all flow elements was inversed at the end of interval T_{pre} and the presentation continued for another 320 ms (T_{post}) Observers reported the final direction of illusory rotation (D_{test}). Illusory rotation either remained stable ($D_{test} = D_{bias}$) or reversed ($D_{test} \neq D_{bias}$). See (Stonkute et al., 2012) and text for details. **c**) During the control test interval ("no inversion") the planar motion remained unperturbed during the entire presentation interval. Observers reported the final direction of illusory rotation (D_{test}) D) Perceptual destabilization ($P_{reversal}$) as a function of the pre-inversion interval (T_{pre}) of on-screen motion. The dashed line marks the 50% threshold: 71.4 ms

presented either with or without the planar motion inversion. The latter case, labeled as "no inversion" in Fig. 2d, was identical to the display presented during the preceding biasing sequence. In the former case, the ambiguously rotating sphere was presented for a variable amount of time ($T_{pre} \in [20, 40, 80, 160, 320]$ ms) before the on-screen motion was inversed and the presentation continued unperturbed for another 320 ms (i.e., total presentation duration $T_{on} = T_{pre} + T_{post} \in [340, 360, 400, 480, 600]$ ms, see Fig. 2b and Movie 2). The purpose of the on-screen motion inversion was to induce a reversal in the direction of illusory rotation (see above; Pastukhov et al., 2012; Stonkute et al., 2012). As during the biasing sequence, observers reported the dominant direction of illusory rotation *at the end* of the presentation interval (D_{test}). The long presentation time ensured that observers reliably perceived illusory rotation at the end of the interval (~1 % of trials were reported as having an "unclear" perception).

Statistical analysis The proportion of the *test* trials on which the inversion of the planar motion did induce the reversal of the illusory rotation was estimates as

$$P_{reversal} = P(D_{bias} \neq D_{test}). \tag{1}$$

Low values of $P_{reversal}$ mean that the on-screen motion inversion was ineffective in triggering a perceptual reversal, whereas values of $P_{reversal} \gg 0$ correspond to the fraction of trials in which the inversion did trigger a reversal. Please note that the probability of reversal is complementary to the probability of survival used in Experiment 2, such that $P_{reversal} = 1-P_{survival}$.

Group averages were fitted with a logistic function using the Palamedes toolbox (Prins & Kingdom, 2009). The error bars represent the 95 % confidence interval (CI) around the mean binomial proportion based on the total number of valid trials.

Results and discussion

The probability of a perceptual reversal ($P_{reversal} = P(D_{bias} \neq D_{test})$) as a function of the time at which planar motion is inverted is plotted in Fig. 2d. "No inversion" and $T_{pre} = 320$ ms conditions served as control conditions. For the "no inversion" condition, there was *no* planar motion inversion, thus the reported direction of illusory rotation should have reflected primarily the influence of sensory memory created

by a preceding biasing sequence (i.e., we expected that $D_{test} = D_{bias}$ and, therefore, $P_{reversal} \approx 0$). Confirming this prediction, changes in illusory rotation were extremely rare: $P_{reversal}$ (no inversion) = 0.04 [0.02 to 0.06] (mean and 95 % CI based on the binomial distribution). Conversely, the longest T_{pre} interval (T_{pre} = 320 ms) should give enough time for the perception of SFM rotation to be resolved, and therefore revealed the maximum destabilization induced by the inversion of the on-screen motion: $P_{reversal}(T_{pre}$ = 320 ms) = 0.43 [0.38 to 0.49] (for individual observers maximal destabilization ranged from $P_{reversal}$ = 0.3 to $P_{reversal}$ = 0.85). To characterize the dependence of the probability of reversal on the duration of interval T_{pre}, the group average was fitted with a logistic function (solid line in Fig. 2d). The 50% threshold was α = 71.4 ms and the 5%-95% range corresponded to [29–114] ms.

As we have argued above, this estimate corresponds to the upper boundary for the when the perception of illusory rotation itself was established and is in a good agreement with visual response latencies for the dorsal visual stream both in monkeys (Schmolesky et al., 1998) and humans (Foxe & Simpson, 2002). However, this could also be an estimate of the lower boundary, i.e. the moment of the earliest significant difference between two competing representations then motion-inversion becomes effective. This would explain why latencies reported here are much shorter than those obtained for random-dot stimuli (Watamaniuk & Sekuler, 1992). Note, however, that Watamaniuk and Sekuler (1992) used refresh rate that was five times lower (20 Hz vs. 100 Hz used in here) and, therefore, their threshold of 9.3 frames could correspond to 93 ms for the 100 Hz presentation, making two estimates very similar. Crucially, both interpretations indicate that perception is formed approximately at the same time as the strongest sensory memory trace (see Experiment 2).

For other multi-stable displays, first-percept disambiguation time was estimated to be ~220–260 ms for Necker cube (Kornmeier & Bach, 2006, 2012) and >150 ms binocular rivalry (Wolfe, 1983). This difference could stem both from different experimental methods used in these studies and from shorter visual response latencies in the dorsal visual stream (Foxe & Simpson, 2002; Schmolesky et al., 1998).

Experiment 2: Estimating the time interval when the strongest sensory memory trace is formed

The purpose of the second experiment was to estimate the time when the strongest sensory memory trace for illusory rotation is formed. Prior research indicates that sensory memory traces build-up throughout the corresponding perceptual dominance period (Brascamp et al., 2008; Leopold et al., 2002; Pastukhov & Braun, 2008). However, their long decay (estimated to be on the scale of dozens of seconds (Brascamp

et al., 2008; Leopold et al., 2002; Pastukhov & Braun, 2008)) makes it hard establish how the perceptual evidence is weighted over time. Yet, it is the weighting that could be most informative about sensory memory. For example, if, as suggested (Knapen et al., 2009), sensory memory helps to stabilize perception, one would expect it to favor the most recent perceptual state, with earlier experience contributing relatively little (Fig. 3, top row). Alternatively, if it is used to anticipate the most likely future perception (making it similar to the repetition priming (Pearson & Brascamp, 2008)), we would expect a larger part of the perceptual history to be taken into account. Still, one would expect higher weights to be given to more recent perceptual configurations, as, typically, it is the most recent not the most distant history that allows for better predictions (Fig. 3, second row). Of course, many other weighting strategies are possible, including, for example, a perfect integrator or extrapolation (Fig. 3, two lower rows).

However, prior studies cannot provide the answer as perceptual configurations of each state of a multi-stable display remained constant throughout each dominance period. In this case, the knowledge that sensory memory favors a particular direction of illusory rotation of an ambiguous sphere (e.g., front surface rotating to the right) is of a little use. As illustrated in Fig. 3a, the visual system samples the same object configuration at every time point, so that all weighting variants produce the same prediction.

To overcome this issue, we relied on shape specificity of sensory memory for illusory rotation (Pastukhov et al., 2013). Sensory memory for a particular direction of rotation appears to contain some information about the shape of an ambiguously rotating object, in a sense that its influence is strongest if the same object is viewed again but weakens progressively as the consequently presented object becomes more and more different from the object that generated this sensory memory. (Note that both objects in question rotate around the same axis and so are ambiguous in an equally balanced way with respect to the same two opposite directions of illusory rotation.)

Here, instead of different shapes we used an ambiguously rotating *band* (see Movie 3), as our pilot experiments indicated that sensory memory is also orientation-specific. As the band rotates, it changes its visual configuration, producing disparate predictions for different weighting strategies (Fig. 3b). Accordingly, we used an intermittently presented band to determine which onset orientation (relative to orientations used in the previous trial) would maximize influence of sensory memory. Next, we used our knowledge about when the band at this particular orientation was presented during the previous trial to estimate how the sensory evidence was weighted over time and when the strongest sensory memory trace was generated. We have employed this procedure previously but using very short interruptions (100–200 ms) to establish orientation-specificity of neural persistence (Pastukhov & Braun, 2013b), it favors the most recent perceptual state.
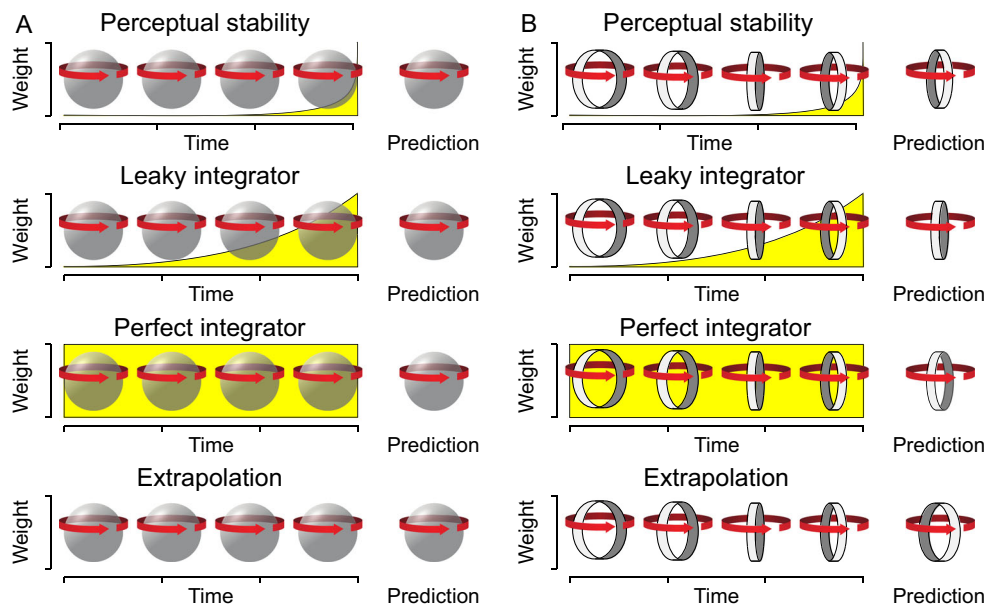
Fig. 3 Predictions for various strategies for weighting sensory evidence over time. (**a**) Rotationally symmetric objects, such as a sphere, result in the same visual configuration at any time point through a perceptual dominance period. Accordingly, all weighting strategies produce the same prediction. (**b**) In contrast, rotationally asymmetric objects, such as a band (lighter gray denotes the front surface), produce different appearances at different time-points despite constant illusory rotation and, correspondingly, different predictions for different weighting strategies

Method

Stimuli The structure-from-motion (SFM) band stimulus consisted of 500 dots distributed randomly over the surface of a band (see Movie 3). It had a height of 5.7°, with individual dots having a size of 0.057° and a luminance of 110 cd/m^2.

Procedure The band stimulus was presented intermittently with observers reporting on the *initial* direction of illusory rotation. A single block consisted of 90 On- and Off-intervals (see Table 1 for further details on four experimental conditions used in the experiment). Within each pair of consecutive presented intervals, the first On-interval is denoted as the prime and the second is denoted as the probe (see Fig. 4a).

The orientation of the band was systematically manipulated. The term "orientation of the band" refers to the orientation of the model of the 3D band, which orthographic projection observers viewed on the screen. For each On-interval, the onset orientation of the band was computed as

$$\Theta^{on}_{probe} = \Theta^{off}_{prime} + \Delta\Theta, \qquad (2)$$

where Θ^{on}_{probe} is the initial orientation at the onset of the probe interval, Θ^{off}_{prime} is the final orientation at the offset of the prime interval, and $\Delta\Theta \in [0°, \pm 22.5°, \pm 45°, \pm 67.5°, \pm 90°]$ (see Movie 4). In other words, the onset orientation for a probe interval could correspond to one of the orientations of the band presented during the prime interval ($\Delta\Theta < 0°$), to the orientation of the band immediately before the interruption ($\Delta\Theta = 0°$, this orientation produced the strongest effect for neural persistence (Pastukhov & Braun, 2013b)), or to the orientation the band would have had if it continued rotating during the Off-interval ($\Delta\Theta > 0°$, i.e., an extrapolated orientation of the band). The relation between the presentation time and the orientation of the band during the prime interval is illustrated in Fig. 4b. The correspondence between onset orientations in the probe interval and the range of orientations presented during the prime interval is shown in Fig. 4c. Note that, due to the symmetry of the band stimulus, its orientation is equivalent for $\Delta\Theta$ and $\Delta\Theta + 180°$.

The influence of perception during the prime interval on the onset perception of the probe interval was quantified as the probability of survival ($P_{survival}$), which is the probability that the same direction of illusory rotation reported in the prime interval was also reported at the onset of the probe interval (Pastukhov et al., 2013, 2014). Note that the probability of survival is complementary to the probability of reversal used in Experiment 1, such that $P_{survival} = 1 - P_{reversal}$. Higher values of $P_{survival}$ (close to 1.0) indicate a strong influence of prime (i.e., a closer match between the strongest sensory memory trace and the orientation of the band in the probe display), whereas lower values indicate a weak influence (i.e., a mismatch between the strongest sensory memory trace and the orientation of the probe display).

Observers reported the direction of illusory rotation using a keyboard. They were instructed to withhold responses when their perception was mixed (e.g., two half-rings rotating independently) or unclear. Trials with no responses or with

Table 1 Summary of experimental conditions. See Methods for details

Label	1 × 1@72	1 × 2@72	0.5 × 1@72	1 × 1@36
T_{on} [s]	1	1	0.5	1
T_{off} [s]	1	2	1	1
Rotation speed V_θ [°/s]	72	72	72	36

multiple responses were discarded (2 % and 0 % of total trials, respectively).

Statistical analysis Group means for each condition were fitted with a von Mises (circular Gaussian) tuning function using a maximum likelihood approach:

$$F(\Theta) = m + R \cdot exp(k \cdot cos[\Theta + \Theta_{best}] \cdot 2), \qquad (3)$$

where θ_{best} is the orientation of the probe display that produces the highest values of $P_{survival}$. The 95 % CIs were estimated using a parametric bootstrap procedure with 2,000 iterations per condition.

Results and discussion

Results for the four experimental conditions are presented in Fig. 5 and are summarized in Table 2. For all conditions, the influence of sensory memory was strongest when the onset orientation for the probe interval was similar (although not identical) to the onset orientation for the prime. Group averages were fitted with a von Mises (circular Gaussian) tuning function, depicted as red curves in Fig. 5 (see Method for details). The key free parameter θ_{best}, which is the onset orientation of the probe display that produced the highest values of $P_{survival}$, is marked by a radial red line and a red stripe (denoting, respectively, the mean and 95 % CI using a parametric bootstrapping).

The knowledge on which orientation of the band corresponded to the strongest sensory memory trace (θ_{best}) allowed us to calculate when this orientation was presented during the preceding prime interval. For all four conditions the strongest sensory memory trace corresponds to the time interval approximately 60–140 ms after the stimulus onset. Note that this range overlaps with the previously estimated time range of when the perception of illusory rotation is formed (see Experiment 1, the corresponding ranges of orientations
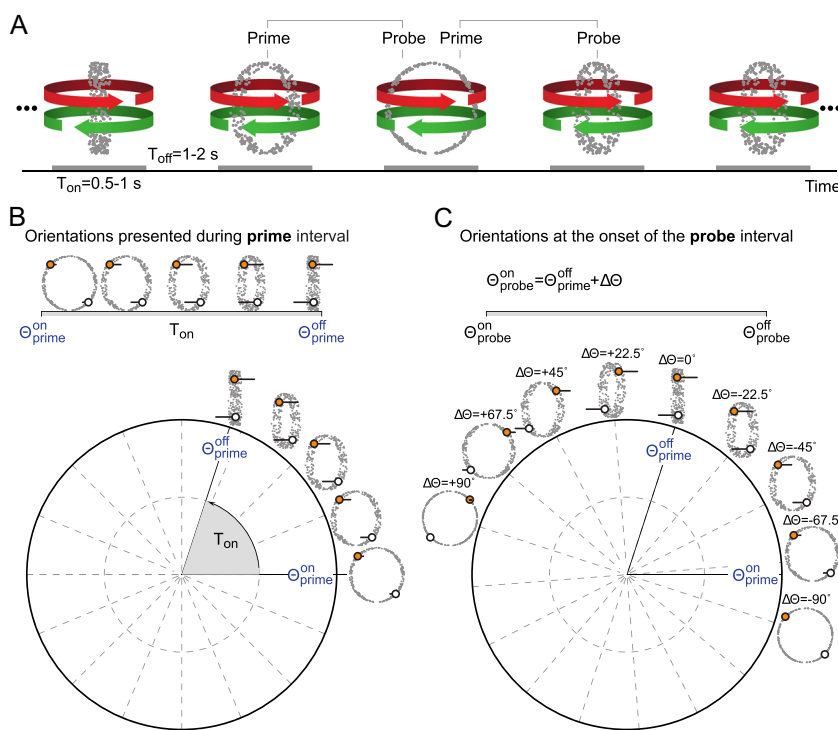

Fig. 4 Experiment 2, schematic procedure: (**a**) An ambiguously rotating band shape was presented intermittently (see Method for details). Subjects reported the perceived direction of illusory rotation after each On-interval. Within each pair of consecutive On-intervals, the first is denoted as the prime and the second as the probe. (**b**) Prime. An ambiguously rotating band continuously changes its orientation throughout the T_{on} interval. The changes in the orientation can also be plotted in a polar coordinate system with Θ_{prime}^{on} aligned to 0° for all trials and time advancing counterclockwise. (**c**) The onset orientation of the band was systematically varied (see Method). The polar plot illustrates all possible onset orientations for the probe trial relative to orientations presented during the preceding prime. Note that, due to the symmetry of the band stimulus, its orientation for $\Delta\Theta + 180°$ is equivalent to its orientation for $\Delta\Theta$. The polar plots illustrate conventions for only two conditions ([T_{on} = 1s, T_{off} = 1 s, $V_\theta \in 72°$/s] and [T_{on} = 1s, T_{off} = 2 s, $V_\theta \in 72°$/s]). Two other conditions are plotted in Fig. 5 in an analogous fashion
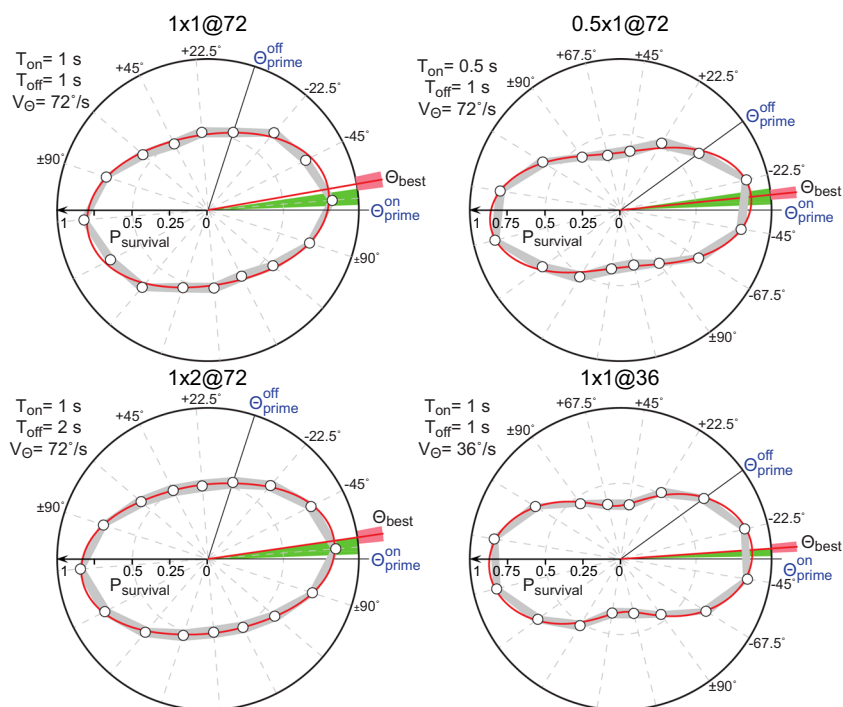
Fig. 5 Experiment 2, results: Probability of survival (P$_{survival}$, plotted radially) as a function of the onset orientation for the probe trial interval relative to the range of orientations presented during the prime (see Fig. 4 for further details on plotting conventions). Circles and gray stripes show, respectively, group means and 95 % CIs based on the binomial distribution. Red curves show the best model fit (see text). Red radial lines and stripes show, respectively, fitted optimal probe-onset orientations (Θ_{best}) and 95 % CIs (parametric bootstrapping, 2,000 trials). Green sectors mark estimated ranges of orientations at the time the perception of illusory rotation was formed during the prime trial (see Experiment 1). For all four conditions, there is a good correspondence between the time when the perception of illusory rotation (green sectors) and the strongest sensory memory trace of illusory rotation (red stripes) were formed

marked by green sectors in Fig. 5). These results indicate that the strongest sensory memory trace of illusory rotation was formed at approximately the same time or shortly after the perception of illusory rotation was formed itself.

Although results presented above indicate that the *strongest* sensory memory trace is formed shortly after the display onset, prior work demonstrated that sensory memory is continuously accumulating throughout the entire dominance period (Brascamp et al., 2008; Leopold et al., 2002; Pastukhov & Braun, 2008). Consistent with that, we found that the range of orientations presented during a single trial (72° or 36°) had a significant effect on the shape of the curve. There is a noticeable dip in the probability of survival for the 36° range (conditions 0.5 × 1@72 and 1 × 1@36), but not for the 72° range (conditions 1 × 1@72 and 1 × 2@72, compare left and rights plots in Fig. 5). This indicates all orientations presented during the trial produced a sensory memory trace of variable strength. To quantify the modulation of P$_{survival}$, $\Delta P = \max(P_{survival}) - \min(P_{survival})$ was computed for each bootstrapped fit. Statistical analysis showed a greater difference between 1 × 1@72 and 0.5 × 1@72 conditions (p < 0.001, ranksum test), and between 1 × 1@72 and 1 × 1@36 conditions (p < 0.001, rank sum test), than between 1 × 1@72 and 1 × 2@72 conditions (p = 0.003, rank sum test).

General discussion

We investigated the interaction between the perception and sensory memory of multi-stable displays by identifying when the strongest sensory memory trace is formed (by perception) and when it, in turn, the multi-stable perception is formed and is affected by sensory memory. Surprisingly, we found that both are formed approximately at the same time, shortly after the

Table 2 Summary of experimental results. 95 % CIs were estimated using a parametric bootstrapping procedure (2,000 iterations). See Method for details

Condition	1 × 1@72	1 × 2@72	0.5 × 1@72	1 × 1@36
Optimal orientation of the band for the probe trial $\Theta_{best} - \Theta_{prime}^{on}$ [°]				
Mean	10.0°	8.2°	5.8°	3.9°
95 % CI	7.4°–12.6°	5.7°–10.7°	3.9°–7.7°	2.3°–5.6°
Time corresponding to $\Theta_{best} - \Theta_{prime}^{on}$ orientation [ms]				
Mean	138 ms	113 ms	81 ms	109 ms
95 % CI	103–175 ms	80–148 ms	54–107 ms	65–157 ms
Modulation of sensory memory bias strength $\Delta P = \max(P_{survival}) - \min(P_{survival})$				
Mean	0.31	0.35	0.49	0.52
95 % CI	0.28–0.34	0.32–0.39	0.46–0.52	0.49–0.55

stimulus onset. The perception of illusory rotation in structure-from-motion (SFM) display was estimated to be fully established within 29–114 ms after the stimulus onset (see Experiment 1). The orientation specificity of sensory memory indicates that the strongest trace was formed about 80–140 ms after the stimulus onset (see Experiment 2). Combined with the results of de Jong et al. (2012), who demonstrated that sensory memory influences multi-stable perception only at the stimulus onset, the data presented here suggest a tighter integration between perception and sensory memory than previously thought.

The current results may call for a reinterpretation of sensory memory and its role in visual perception. It was hypothesized previously to be a predictive memory that stabilizes the current visual perception and maintains it during interruptions in the stimulation (Knapen et al., 2009) or, analogously to repetition priming (Kristjánsson & Campana, 2010), creates the expectation that a previous perceptual experience will be repeated (Pearson & Brascamp, 2008). Both hypotheses assume that the observed facilitation during the following probe interval is the primary effect of sensory memory. In other words, both assume that the key functional contribution of sensory memory to vision is in the anticipation (prediction) of future perception.

However, as discussed in the Introduction, the known properties of sensory memory indicate that this predictive facilitatory effect can influence perception only under very special conditions, which are very rarely satisfied during typical viewing. The results of Experiment 2 also pose a problem for current hypotheses about the functional role of sensory memory. Specifically, any hypothesis that focuses on its facilitatory effect would have to explain the advantage of having a predictive

memory whose strongest trace favors the earliest perceptual state (i.e., the most distant in time) over the more recent ones. Note that these hypotheses are also virtually impossible to falsify, as any information about the past is better than no information and, therefore, one could always argue that in this case the visual system *might* rely on it under special circumstances.

Perhaps it is possible that the facilitation of future perception is not the primary purpose of sensory memory. The fact that perception and the strongest sensory memory trace both form and influence each other shortly after the display onset leads us to suggest an alternative hypothesis. We propose that the primary functional contribution of sensory memory *mechanisms* could be to assist in the construction of perception at the stimulus onset. These memory mechanisms, which are responsible for the later perceptual effect dubbed sensory memory, could serve as buffers that accumulate sensory evidence for competing percepts over time. The percept that is quickest in accumulating evidence in its favor wins the "race to the boundary" and dominates the perception, whereas the competing percept is suppressed. Such an accumulation of evidence for conflicting hypotheses (albeit for longer time scales) has been reported as part of the perceptual decision making process (race models (Gold & Shadlen, 2007)). This is also consistent with slower perceptual inference for full ambiguous displays as compared to their unambiguous versions (Kornmeier & Bach, 2004). The latter are strongly biased towards one particular perceptual interpretation, which accelerates evidence accumulation in their favor. The later facilitatory effect on the perception of the *following (probe)* display could be caused by a residual activation of, or changes in, the synaptic connections of these memory buffers,
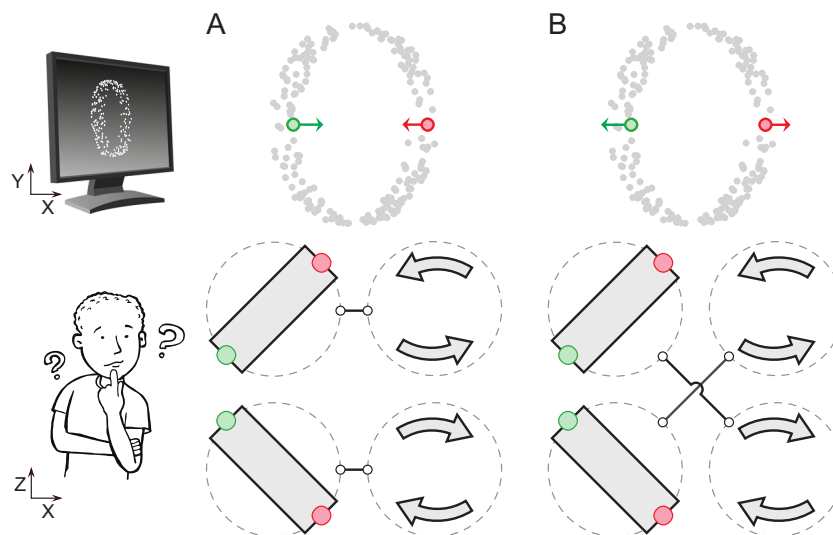

Fig. 6 Possible and internally consistent combinations of perceptual states of illusory rotation and illusory depth in structure-from-motion displays. (a) A given 2D planar motion is consistent with two alternative states of illusory depth (the orientation of the bar) and two directions of illusory rotation (indicated by the arrows). However, only two of the four combinations give an internally consistent perception

(valid combinations are connected by lines). (b) When sensory input changes (note the reversed direction of planar motion for marked dots in the top row), initially consistent combinations become incompatible, whereas the two other combinations (inconsistent in a) produce an internally consistent perception

providing a small head start in this "race to the boundary" for one of the competing states. Whether this residual effect has a functional role of its own is open for debate. However, the fact that it can be detected only with a very sensitive visual display and even then only under very specific conditions suggests that it might have little practical importance (see Introduction).

Thus, it is possible that the results of sensory memory experiments do not tell us about how the visual system anticipates future perceptions. Instead, it might inform us about how perception is constructed in the visual system. In this case, one can view sensory memory as a sort of "fossil" of the initial perception, as it is remarkably robust to intervening and divergent stimuli and perceptions. This would make it a unique behavioral correlate of the earlier perceptual inference that can be studied post hoc. For example, the specificity of sensory memory may reflect constraints that are imposed on the inference of complex objects. Such objects are represented by a distributed network of coordinated, yet independent, neural representations (Grill-Spector, 2003; Orban, 2011; Pastukhov & Braun, 2013b; Pastukhov et al., 2014). Some object properties and, corresponding representations, may not constrain each other and, therefore, can be constructed independently of each other. But at least some representations of the object's properties must be constructed in a coordinated way. In the case of SFM, co-dependence of representations of object's shape, illusory rotation and illusory depth (see Fig. 6) could explain the shape selectivity of sensory memory (Pastukhov et al., 2013, 2014). Conversely, the lack of interdependence between two representations, e.g., between representations of the object's color and illusory rotation, could be manifested in a lack of specificity for sensory memory (Chen & He, 2004; Maier et al., 2003; Pastukhov et al., 2014). This makes sensory memory a potentially useful tool for the characterization of dynamic networks of sensory representations that emerge when a perception is constructed. Knowledge about these co-dependent representation networks would provide a powerful guide for both modeling and imaging studies, particularly those that investigate functional connectivity (Deco et al., 2013).

Above, we focused on the role of sensory memory during the initial perceptual inference at the onset of the visual stimulation. However, for multi-stable displays this is not the only time at which perception must be disambiguated. During continued viewing, perception is constantly destabilized and needs to be reconstructed. These periods of destabilization manifest themselves as very brief transition phases, which may be too brief for observers to report (Huguet, Rinzel, & Hupe, 2014; Kim, Grabowecky, & Suzuki, 2006; Naber, Frässle, & Einhäuser, 2011). It is likely that the same perceptual and memory mechanisms assist perceptual inference during these destabilizations and during perceptual reversals (Knapen, Brascamp, Pearson, van Ee, & Blake, 2011). This is consistent with weaker but detectable sensory memory traces for orientations presented later in a trial in Experiment

2, as well as with prior work that shows a positive dependence between the duration of a dominance interval and the strength of sensory memory (Brascamp et al., 2008; Leopold et al., 2002; Pastukhov & Braun, 2008).

To conclude, we demonstrated that perception and the strongest sensory memory trace of illusory rotation both form shortly after the stimulus onset. This suggests a tighter integration between perception and sensory memory than previously thought and warrants a reconsideration of the role of sensory memory in visual perception.

References

Adams, P. A. (1954). The Effect of past Experience on the Perspective Reversal of a Tridimensional Figure. *The American Journal of Psychology, 67*(4), 708. doi:10.2307/1418496

Baloh, R. W., Sills, A. W., Kumley, W. E., & Honrubia, V. (1975). Quantitative measurement of saccade amplitude, duration, and velocity. *Neurology, 25*(11), 1065–1065. doi:10.1212/WNL.25.11.1065

Blake, R., Westendorf, D., & Fox, R. (1990). Temporal perturbations of binocular rivalry. *Perception & Psychophysics, 48*(6), 593–602. doi:10.3758/BF03211605

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433–436. doi:10.1163/156856897X00357

Brascamp, J. W., Knapen, T. H. J., Kanai, R., Noest, A. J., van Ee, R., & van den Berg, A. V. (2008). Multi-timescale perceptual history resolves visual ambiguity. *PLoS One, 3*(1), e1497. doi:10.1371/journal.pone.0001497

Carter, O., & Cavanagh, P. (2007). Onset rivalry: Brief presentation isolates an early independent phase of perceptual competition. *PLoS One, 2*(4), e343. doi:10.1371/journal.pone.0000343

Chen, X., & He, S. (2004). Local Factors Determine the Stabilization of Monocular Ambiguous and Binocular Rivalry Stimuli. *Current Biology, 14*(11), 1013–1017. doi:10.1016/j.cub.2004.05.042

De Jong, M. C., Knapen, T. H. J., & van Ee, R. (2012). Opposite influence of perceptual memory on initial and prolonged perception of sensory ambiguity. *PLoS One, 7*(1), e30595. doi:10.1371/journal.pone.0030595

Deco, G., Ponce-Alvarez, A., Mantini, D., Romani, G. L., Hagmann, P., & Corbetta, M. (2013). Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience, 33*(27), 11239–52. doi:10.1523/JNEUROSCI.1091-13.2013

Foxe, J. J., & Simpson, G. V. (2002). Flow of activation from V1 to frontal cortex in humans. A framework for defining "early" visual processing. *Experimental Brain Research, 142*(1), 139–50. doi:10.1007/s00221-001-0906-7

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience, 30*, 535–74. doi:10.1146/annurev.neuro.29.051605.113038

Grill-Spector, K. (2003). The neural basis of object perception. *Current Opinion in Neurobiology, 13*(2), 159–66. doi:10.1016/S0959-4388(03)00040-0

Huguet, G., Rinzel, J., & Hupe, J. (2014). Noise and adaptation in multistable perception: Noise drives when to switch, adaptation determines percept choice. *Journal of Vision, 14*(3), 19–19. doi:10.1167/14.3.19

Kim, Y.-J., Grabowecky, M., & Suzuki, S. (2006). Stochastic resonance in binocular rivalry. *Vision Research, 46*(3), 392–406. doi:10.1016/j.visres.2005.08.009

Knapen, 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. doi:10.1167/9.13.16

Knapen, T. H. J., Brascamp, J. W., Pearson, J., van Ee, R., & Blake, R. (2011). The Role of Frontal and Parietal Brain Areas in Bistable Perception. *Journal of Neuroscience, 31*(28), 10293–10301. doi:10.1523/JNEUROSCI.1727-11.2011

Kornmeier, J., & Bach, M. (2004). Early neural activity in Necker-cube reversal: Evidence for low-level processing of a gestalt phenomenon. *Psychophysiology, 41*(1), 1–8. doi:10.1046/j.1469-8986.2003.00126.x

Kornmeier, J., & Bach, M. (2006). Bistable perception – along the processing chain from ambiguous visual input to a stable percept. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology, 62*(2), 345–9. doi:10.1016/j.ijpsycho.2006.04.007

Kornmeier, J., & Bach, M. (2012). Ambiguous Figures – What Happens in the Brain When Perception Changes But Not the Stimulus. *Frontiers in Human Neuroscience, 6*(March), 1–23. doi:10.3389/fnhum.2012.00051

Kornmeier, J., Ehm, W., Bigalke, H., & Bach, M. (2007). Discontinuous presentation of ambiguous figures: How interstimulus-interval durations affect reversal dynamics and ERPs. *Psychophysiology, 44*(4), 552–60. doi:10.1111/j.1469-8986.2007.00525.x

Kristjánsson, Á., & Campana, G. (2010). Where perception meets memory: A review of repetition priming in visual search tasks. *Attention, Perception, & Psychophysics, 72*(1), 5–18. doi:10.3758/APP.72.1.5

Leopold, D. A., & Logothetis, N. K. N. (1999). Multistable phenomena: Changing views in perception. *Trends in Cognitive Sciences, 3*(7), 254–264. doi:10.1016/S1364-6613(99)01332-7

Leopold, D. A., Wilke, M., Maier, A., & Logothetis, N. K. (2002). Stable perception of visually ambiguous patterns. *Nature Neuroscience, 5*(6), 605–9. doi:10.1038/nn851

Maier, A., Wilke, M., Logothetis, N. K., & Leopold, D. A. (2003). Perception of Temporally Interleaved Ambiguous Patterns. *Current Biology, 13*(13), 1076–1085. doi:10.1016/S0960-9822(03)00414-7

Naber, M., Frässle, S., & Einhäuser, W. (2011). Perceptual Rivalry: Reflexes Reveal the Gradual Nature of Visual Awareness. *PLoS ONE, 6*(6), e20910. doi:10.1371/journal.pone.0020910

Nawrot, M., & Blake, R. (1991). The interplay between stereopsis and structure from motion. *Perception & Psychophysics, 49*(3), 230–44. doi:10.3758/BF03214308

Oh, S. (2011). The eyeglass reversal. *Attention, Perception, & Psychophysics, 73*(5), 1336–43. doi:10.3758/s13414-011-0116-8

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,* 439–58. doi:10.2466/pms.1963.17.2.439

Orban, G. A. (2011). The extraction of 3D shape in the visual system of human and nonhuman primates. *Annual Review of Neuroscience, 34,* 361–88. doi:10.1146/annurev-neuro-061010-113819

Pastukhov, A., & Braun, J. (2008). A short-term memory of multi-stable perception. *Journal of Vision, 8*(13), 7.1–14. doi:10.1167/8.13.7

Pastukhov, A., & Braun, J. (2013a). Disparate time-courses of adaptation and facilitation in multi-stable perception. *Learning & Perception, 5*(s2), 101–118. doi:10.1556/LP.5.2013.Suppl2.7

Pastukhov, A., & Braun, J. (2013b). Structure-from-motion: Dissociating perception, neural persistence, and sensory memory of illusory depth and illusory rotation. *Attention, Perception, & Psychophysics, 75*(2), 322–40. doi: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. doi:10.1167/12.1.17

Pastukhov, A., Füllekrug, J., & Braun, J. (2013). Sensory memory of structure-from-motion is shape-specific. *Attention, Perception, & Psychophysics, 75*(6), 1215–29. doi:10.3758/s13414-013-0471-8

Pastukhov, A., Lissner, A., Füllekrug, J., & Braun, J. (2014). Sensory memory of illusory depth in structure-from-motion. *Attention, Perception, & Psychophysics, 76*(1), 123–32. doi:10.3758/s13414-013-0557-3

Pearson, J., & Brascamp, J. W. (2008). Sensory memory for ambiguous vision. *Trends in Cognitive Sciences, 12*(9), 334–41. doi:10.1016/j.tics.2008.05.006

Prins, N., & Kingdom, F. A. A. (2009). *Palamedes: Matlab routines for analyzing psychophysical data.* Http://www.palamedestoolbox.org

Rock, I., & Mitchener, K. (1992). Further evidence of failure of reversal of ambiguous figures by uninformed subjects. *Perception, 21*(1), 39–45. doi:10.1068/p210039

Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D., & Leventhal, A. G. (1998). Signal timing across the macaque visual system. *Journal of Neurophysiology, 79*(6), 3272–8.

Sperling, G., & Dosher, B. A. (1994). Depth from motion. In T. V. Papathomas, A. G. Charles Chubb, & E. Kowler (Eds.), *Early Vision and Beyond* (pp. 133–142). Cambridge: MIT Press.

Sterzer, P., & Rees, G. (2008). A neural basis for percept stabilization in binocular rivalry. *Journal of Cognitive Neuroscience, 20*(3), 389–99. doi:10.1162/jocn.2008.20039

Stonkute, S., Braun, J., & Pastukhov, A. (2012). The role of attention in ambiguous reversals of structure-from-motion. *PLoS One, 7*(5), e37734. doi:10.1371/journal.pone.0037734

Volkmann, F., Riggs, L., & Moore, R. (1980). Eyeblinks and visual suppression. *Science, 207*(4433), 900–902. doi:10.1126/science.7355270

Watamaniuk, S. N. J., & Sekuler, R. (1992). Temporal and spatial integration in dynamic random-dot stimuli. *Vision Research, 32*(12), 2341–2347. doi:10.1016/0042-6989(92)90097-3

Wolfe, J. M. (1983). Influence of spatial frequency, luminance, and duration on binocular rivalry and abnormal fusion of briefly presented dichoptic stimuli. *Perception, 12*(4), 447–456. doi:10.1068/p120447