



HAL
open science

The continuous Wagon Wheel Illusion is object-based.

Rufin Vanrullen

► **To cite this version:**

Rufin Vanrullen. The continuous Wagon Wheel Illusion is object-based.. *Vision Research*, 2006, 46 (24), pp.4091-5. 10.1016/j.visres.2006.07.030 . hal-00111090

HAL Id: hal-00111090

<https://hal.science/hal-00111090>

Submitted on 10 Nov 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The continuous Wagon Wheel Illusion is object-based

Rufin VanRullen

Centre de Recherche Cerveau et Cognition (UMR 5549)
CNRS - Université Paul Sabatier Toulouse 3
Faculté de Médecine Rangueil, 31062 Toulouse Cedex (France)

email: rufin@klab.caltech.edu
Tel: +33 5 62 17 37 76
Fax: +33 5 62 17 28 09

Summary

The occurrence of perceived reversed motion while observers view a periodic, continuously moving stimulus (the “continuous Wagon Wheel Illusion”) has been taken as evidence that some aspects of motion perception rely on discrete sampling of visual information. The spatial extent of this sampling is currently under debate. When two separate motion stimuli are viewed simultaneously, the illusion of reversed motion rarely occurs for both objects together: this rules out global sampling of the visual field. The same result holds when the objects are superimposed by transparency: this argues against location-based sampling. Here we show that the sampling is in fact object-based: we use a rotating ring stimulus split in two halves. When the two halves move in opposite directions, appearing to belong to separate objects, perceptual reversals occur in either half at a time, but rarely in both. When the two halves physically move in compatible directions, they generally appear to reverse simultaneously: the illusion keeps the perceptual object united. Rather than the local low-level properties of the motion stimulus (which are comparable in both cases), it is thus the high-level organization of the scene that determines the extent of perceived motion reversals. These results imply that the continuous Wagon Wheel Illusion, and any discrete perceptual sampling that may cause it, is restricted to the object of our attention.

Manuscript counts: 2657 words in main text, 2 Figures, 25 references, 1 footnote.

Introduction

In movies or on TV, the wheels of a passing car sometimes appear to rotate in the “wrong” direction. This so-called “Wagon Wheel” effect is simply due to the discrete nature of video frames. Surprisingly, a similar effect can be observed in real life, under continuous conditions of illumination (Schouten, 1967; Purves, Paydarfar, & Andrews, 1996) –although there are important differences between the two phenomena (Pakarian & Yasamy, 2003; Kline, Holcombe, & Eagleman, 2004). This bistable effect, the “continuous Wagon Wheel Illusion” has been taken as evidence that some aspects of motion perception rely on discrete sampling of visual information (Purves et al., 1996; VanRullen, Reddy, & Koch, 2005, 2006; Simpson, Shahani, & Manahilov, 2005; Andrews & Purves, 2005; VanRullen & Koch, 2003). Recently, the spatial extent of this sampling has come under debate (Rojas, Carmona-Fontaine, Lopez-Calderon, & Aboitiz, 2006; Kline, Holcombe, & Eagleman, 2006). When two separate, periodic motion stimuli are viewed simultaneously, the illusion of reversed motion rarely occurs for both objects together: this observation rules out global sampling of the visual field (Kline et al., 2004). The same result holds when two independently moving objects (e.g. a rotating fan and an expanding pattern) are superimposed by transparency: this argues against location-based sampling (Kline et al., 2006). Here we report clear evidence that the illusion is in fact object-based: it affects two separate locations *in alternation* when they are perceived as separate objects, but *simultaneously* when they are perceived as a single object.

Results

To distinguish whether the continuous Wagon Wheel Illusion (c-WWI) is location-based (Kline et al., 2004) or object-based (VanRullen et al., 2005), we designed a motion stimulus that would be perceived as one or two separate objects –depending on a minor manipulation that did not affect (in a statistical sense) its low-level structure. We used a ring stimulus, within which a radial grating was rotated at a temporal frequency of 10Hz, the optimal frequency to generate the c-WWI (VanRullen et al., 2005; Simpson et al., 2005). A vertical gray bar was superimposed on the screen, hiding the midline of the ring stimulus. At the beginning of each trial, we (pseudo-) randomly decided on the rotation direction (clockwise or counterclockwise) for each half of the ring stimulus independently. Therefore, half of the trials had the two half-rings moving in compatible directions (“congruent rotation” trials) and the rest of the trials had the half-rings moving in opposite directions (“incongruent rotation” trials). Although the local low-level properties of the stimulus were statistically comparable in both cases, the Gestalt principles of similarity, good continuation and common fate (Koffka, 1935; Kohler, 1947) resulted in the ring being perceived as a whole, united object in the former case, but not in the latter¹. Thus, location-based sampling of motion information would predict that both trial types should result in comparable amounts of c-WWI; object-based sampling, on the other hand, would predict that congruent rotation trials result in illusory reversals of the entire ring, whereas incongruent trials should yield more independent reversals in each half-ring. (Global sampling, finally, would predict that the whole ring always reverses at once, irrespective of the type of trial.)

As for other bistable effects, we evaluated the strength of the c-WWI as the percentage of the time spent reporting reversed motion (Kline et al., 2004; VanRullen et al., 2005) (Figure 1a). As predicted by the object-based account, there were significantly more joint reversals of both half-rings in the congruent than in the incongruent rotation trials (paired t-test, $t(6)=2.92, p=0.026$), and more reversals of a single half-ring in the incongruent than in the congruent rotation trials ($t(6)=6.43, p<0.001$ for the left half-ring, $t(6)=5.72, p<0.002$ for the right half-ring). Note that, as pointed out by Kline, Holcombe and Eagleman (Kline et al., 2004), the fact that many reversals were restricted to a single half-ring (in particular, during the incongruent trials) rules out global sampling of the visual field.

We also compared the amount of joint reversals in each condition with the amount that would be predicted if each half-ring reversed independently (which was calculated by multiplying the total proportion of reversals obtained for each half-ring). The actual amount of joint reversals was significantly higher than this predicted number in the congruent trials ($t(6)=3.47, p<0.02$), and lower than this prediction for the incongruent trials (although the effect did not reach significance, $t(6)=1.65, p>0.05$). Thus, the perceived unity of the two half-rings in the congruent condition resulted in a strong tendency of the c-WWI to encompass both half-rings. In other words, the c-WWI affected the entire ring when it was perceived as a single entity, and each half independently when they were perceived as separate. This is evidence that the c-WWI effect is object-based.

Figure 1 here

Since motion adaptation is likely to play a role in the c-WWI (e.g. by tipping off the perceptual balance in favor of the illusory motion direction), we verified that the strong qualitative change in perceptual appearance between congruent and incongruent trials was not due to a difference in the absolute amount of motion adaptation. To this end, we measured the duration of the “static” and “flicker” motion aftereffects (MAE) induced by our ring stimulus, in both conditions (Mather, Verstraten, & Anstis, 1998). Both measures characterize the amount of neuronal adaptation to a motion stimulus, but the flicker MAE is generally thought to provide a more sensitive measure and/or to tap into a higher-order motion system (Ledgeway, 1994; Nishida & Sato, 1995; Culham, Verstraten, Ashida, & Cavanagh, 2000). For both types of MAE, the obtained durations did not differ sensibly between congruent and incongruent rotation trials. The average (\pm standard error) static MAE duration was 4.48s (± 1.13 s) for congruent trials and 4.77s (± 1.20 s) for incongruent trials ($t(6)=0.18, p>0.05$). The average flicker MAE duration was 5.08s (± 1.09 s) for congruent trials and 4.62s (± 1.05 s) for incongruent trials ($t(6)=0.30, p>0.05$). Thus, the absolute level of motion adaptation does not depend on whether the ring is perceived as a united object or not, and is unlikely to explain our finding that the c-WWI is object-based.

Our results so far indicate that the perceived direction of motion in one hemifield is heavily influenced by the rotation direction in the other hemifield. We wondered whether this effect could be explained by the heavy neural circuitry involved in spatially linking visual attributes that straddle the vertical meridian and thus project in distinct brain hemispheres (Gazzaniga, 1995). To explore this possibility, we repeated the entire experiment, but with the entire stimulus rotated by 90° (Figure 2a). Now interactions between the two half-rings could take place within each cortical hemisphere rather than only across hemispheres. Results were, however, essentially the same as before. More joint reversals of the two halves occurred in the congruent trials than in the incongruent trials ($t(4)=3.27$, $p<0.05$), whereas more individual reversals of a single half-ring occurred in the incongruent than in the congruent trials ($t(4)=8.50$, $p=0.001$ for the top half-ring, $t(4)=4.50$, $p<0.02$ for the bottom half-ring). During congruent (respectively, incongruent) rotation there were also more (respectively, less) joint reversals than would be predicted if the two halves reversed independently ($t(4)=3.54$, $p<0.05$ for congruent rotation, $t(4)=2.95$, $p<0.05$ for incongruent rotations). Finally, neither the static nor the flicker MAEs (Figure 2b) differed significantly in magnitude between congruent and incongruent trials ($t(4)=0.13$, $p>0.05$ and $t(4)=0.73$, $p>0.05$ respectively). Overall, there was thus no evidence that inter-hemispheric communication is an important factor in our findings of object-based influences on the c-WWI effect.

Figure 2 here

Discussion

What is an object? Can we be confident that the reports of our observers reflect perceived “objecthood” rather than merely a high-level response bias? As pointed out by David Marr (Marr, 1982), objects are an elusive concept in vision science:

“Is a nose an object? Is a head one? Is it still one if it is attached to a body? What about a man on horseback? These questions show that the difficulties in trying to formulate what should be recovered as a region from an image are so great as to amount almost to philosophical problems. There is really no answer to them --all these things can be an object if you want to think of them that way, or they can be part of a larger object.” (p. 270)

Indeed, decades of investigation in the Gestalt tradition (Koffka, 1935; Kohler, 1947) have only started to reveal what features tend to group together, and under what conditions, to produce perceived “objects”. In a more pragmatic way –and even though this merely displaces our definition problem- the notion of object is intimately linked to the concept of attention, as suggested in the famous quote from William James (James, 1890):

“Every one knows what attention is. It is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects [...]” (p 403).

Without committing to a precise definition which would inevitably be too restrictive, we can thus simply consider that “everyone knows what an object is”, and that whatever our observers individuated as “objects” during the experiments should probably be regarded as such.

With this in mind, our results indicate that the continuous Wagon Wheel Illusion is a high-level effect, whose spatial extent is entirely determined by the global perceptual organization of the scene into objects. This implies that the illusion cannot be explained simply in terms of the properties of individual motion detectors in early cortical areas. Together with a previous report of a strong attentional influence on the c-WWI (VanRullen et al., 2005), these findings imply that the illusion may reflect the periodic operation of attention. According to our calculations (VanRullen et al., 2005), and to our recent experimental confirmation using EEG (VanRullen et al., 2006), the “rate” of this periodic attentional capture of motion information would lie around 13Hz.

Both our account and the opposing view of this illusion championed by Kline and colleagues (Kline et al., 2004, 2006), agree on the bistable nature of the effect: a (relatively weak) signal supporting illusory motion must be rivalling with the veridical motion signal to induce the occasional perceptual reversals. One major point of disagreement between the two camps is on the origin of these signals supporting the non-veridical direction of motion: attention-based (and object-based) discrete sampling in our view, motion adaptation (and/or Reichardt motion detector aliasing) according to Kline and colleagues (Kline et al., 2004, 2006). While the present experiment does not in itself provide definitive evidence for our view, it restrains the range of adaptation-based models that could account for the illusion. In particular, our data could not be explained by the specific adaptation of high-level, rotation-sensitive neurons (Sakata et al., 1994), since we measured no stronger motion aftereffect in the congruent rotation trials than in the incongruent ones (Figures 1b and 2b). On the other hand, as mentioned above, adaptation of low-level motion detectors (such as Reichardt-based detectors) cannot easily account for our high-level object-based effects. Future work will hopefully disambiguate the relative contributions of discrete sampling vs. motion adaptation to the continuous Wagon Wheel Illusion.

During incongruent trials, the total amount of time spent by our observers in a non-veridical percept (i.e. with at least one half-ring reported in the illusory direction) was about twice as high as during the congruent trials. The congruent ring stimuli are similar to the conventional “wheel” stimuli often used in studies of the c-WWI (Purves et al., 1996; VanRullen et al., 2005; Simpson et al., 2005; VanRullen et al., 2006) (except for an empty disk in the center). Why would the incongruent ring increase the strength of the illusion compared to the more conventional, congruent one? This can be understood by relating the bistable motion percept of the c-WWI to a dynamic system (Poston & Stewart, 1978; Richards, Wilson, & Sommer, 1994). A congruent wheel or ring stimulus would constitute a stable point for motion perception: it takes considerable energy (i.e. long adaptation durations) to venture away from this point. An incongruent ring such as ours, on the other hand, can be considered a highly unstable point, because it cannot be bound into a single “object” representation: less energy (i.e. shorter adaptation durations) may be required to swing from this point to a more stable motion percept –which is achieved when a single half-ring reverses, resulting in a perceptually uniform object. Our procedure of introducing

incongruent motion between the two halves of the stimulus may thus be a promising way to increase the strength of the c-WWI in future experimental studies.

Methods

Seven subjects, including the author, participated in the main experiment (Figure 1), and 5 of these (including the author) also performed the control experiment (Figure 2). All subjects provided informed consent prior to the experiments, which were run according to local ethical guidelines. The subjects were seated in the dark, approximately 60cm away from a computer screen with a refresh rate of 100Hz. Given the limited temporal frequency of our stimuli, this refresh rate was sufficiently fast to avoid any spurious temporal aliasing that may have contaminated our results (Burr, Ross, & Morrone, 1986; VanRullen et al., 2005).

Our stimulus was a bisected ring (7° diameter and 1° width), within which a radial grating (spatial frequency 1 cycle per degree) was rotated at a temporal frequency of 10Hz, the optimal frequency to generate the c-WWI (VanRullen et al., 2005; Simpson et al., 2005). A vertical gray bar (width 1.5°) was superimposed on the screen, hiding the midline of the ring stimulus. For the control experiment, a horizontal bar was hiding the horizontal midline. Observers fixated a small dot in the middle of the bar. At the beginning of each trial, we (pseudo-) randomly decided on the rotation direction (clockwise or counterclockwise) for each half of the ring stimulus independently. Therefore, half of the trials had the two half-rings moving in compatible directions (“congruent rotation” trials) and the rest of the trials had the half-rings moving in opposite directions (“incongruent rotation” trials). The two main stimulus conditions can be viewed online at <http://www.klab.caltech.edu/~rufin/ringmovies/>

To measure the occurrence of illusory motion reversals during each 40s trial (16 trials per condition, randomly interleaved), we instructed the subjects to press the left arrow key whenever the left half-ring appeared to reverse, and to keep it pressed for the entire duration of the reversal; similarly, they were told to press the right arrow key when the right half-ring reversed; no key was pressed when no illusion was present, and both keys were pressed when the two half-rings reversed as one. In the control experiment, the up and down arrows replaced the left and right arrow keys. As for other bistable effects, we evaluated the strength of the c-WWI as the percentage of the time spent reporting reversed motion (Kline et al., 2004; VanRullen et al., 2005). Even though this “classic” procedure of subjective self-paced report can be sometimes prone to response biases (Mamassian & Goutcher, 2005), it was deemed more suitable for our particular stimulus conditions in which two bistable objects (i.e., the two half-rings) had to be monitored simultaneously. Note that in another study we have proposed a more objective way of assessing the strength of the c-WWI, based on unbalanced counterphase gratings (VanRullen et al., 2005).

To evaluate the contribution of neuronal adaptation, we measured the duration of the static and flicker motion aftereffects (MAE) induced by our ring stimulus, in both conditions. Immediately after each 40s trial, one of the two half-rings disappeared (and was replaced by the gray background), while the other half-ring remained but stopped its rotation; the observer was

instructed to wait until any impression of (reversed) motion was extinguished in this remaining half-ring, and then press a key to signal the end of the aftereffect. In half of the trials (randomly determined), the remaining half-ring was static, while in the other half it flickered steadily at 10Hz. The resulting durations represented a measure of the strength of the “static” and “flicker” MAE, respectively (Mather et al., 1998).

Acknowledgements

The author wishes to thank Christof Koch, Patrick Cavanagh and Leila Reddy for constructive discussions and/or comments on the manuscript.

¹ In a separate control experiment, we verified that observers do tend to perceive congruent and incongruent rotating rings as containing 1 vs. 2 separate “objects”, respectively. We presented each stimulus type (in counterbalanced order across observers) for an unlimited duration to 10 observers (5 of which had participated in at least one of our experiments) and asked them “Do you see one moving object, two moving objects, or is it too ambiguous to tell?”. All observers reported seeing one moving object for congruent rotation and two objects during incongruent rotation. Readers can experience these stimuli online at: <http://www.klab.caltech.edu/~rufin/ringmovies/>

References

- Andrews, T., & Purves, D. (2005). The wagon-wheel illusion in continuous light. *Trends Cogn Sci*, 9(6), 261-263.
- Burr, D. C., Ross, J., & Morrone, M. C. (1986). Smooth and sampled motion. *Vision Res*, 26(4), 643-652.
- Culham, J. C., Verstraten, F. A., Ashida, H., & Cavanagh, P. (2000). Independent aftereffects of attention and motion. *Neuron*, 28(2), 607-615.
- Gazzaniga, M. S. (1995). Principles of human brain organization derived from split-brain studies. *Neuron*, 14(2), 217-228.
- James, W. (1890). *The principles of Psychology* (Vol. I). New York: Holt.
- Kline, K., Holcombe, A. O., & Eagleman, D. M. (2004). Illusory motion reversal is caused by rivalry, not by perceptual snapshots of the visual field. *Vision Res*, 44(23), 2653-2658.
- Kline, K., Holcombe, A. O., & Eagleman, D. M. (2006). Illusory motion reversal does not imply discrete processing: Reply to Rojas et al. *Vision Res*, 46(6-7), 1158-1159.
- Koffka, K. (1935). *Principles of gestalt psychology*. NY: Harcourt.
- Kohler, W. (1947). *Gestalt psychology: An introduction to new concepts in modern psychology*. NY: Liverright.
- Ledgeway, T. (1994). Adaptation to second-order motion results in a motion aftereffect for directionally-ambiguous test stimuli. *Vision Res*, 34(21), 2879-2889.
- Mamassian, P., & Goutcher, R. (2005). Temporal dynamics in bistable perception. *J Vis*, 5(4), 361-375.
- Marr, D. (1982). *Vision: a computational investigation into the human representation and processing of visual information*. San Francisco: W.H. Freeman.
- Mather, G., Verstraten, F. A., & Anstis, S. (1998). *The motion aftereffect: a modern perspective*. Cambridge, MA: MIT Press.
- Nishida, S., & Sato, T. (1995). Motion aftereffect with flickering test patterns reveals higher stages of motion processing. *Vision Res*, 35(4), 477-490.
- Pakarian, P., & Yasamy, M. T. (2003). Wagon-wheel illusion under steady illumination: real or illusory? *Perception*, 32(11), 1307-1310.
- Poston, T., & Stewart, I. (1978). Nonlinear modeling of multistable perception. *Behav Sci*, 23(5), 318-334.
- Purves, D., Paydarfar, J. A., & Andrews, T. J. (1996). The wagon wheel illusion in movies and reality. *Proc Natl Acad Sci U S A*, 93(8), 3693-3697.
- Richards, W., Wilson, H. R., & Sommer, M. A. (1994). Chaos in percepts? *Biol Cybern*, 70(4), 345-349.
- Rojas, D., Carmona-Fontaine, C., Lopez-Calderon, J., & Aboitiz, F. (2006). Do discreteness and rivalry coexist in illusory motion reversals? *Vision Res*, 46(6-7), 1155-1157.
- Sakata, H., Shibutani, H., Ito, Y., Tsurugai, K., Mine, S., & Kusunoki, M. (1994). Functional properties of rotation-sensitive neurons in the posterior parietal association cortex of the monkey. *Exp Brain Res*, 101(2), 183-202.
- Schouten, J. F. (1967). Subjective stroboscopy and a model of visual movement detectors. In I. Wathen-Dunn (Ed.), *Models for the perception of speech and visual form* (pp. 44-45). Cambridge, MA.: MIT Press.
- Simpson, W. A., Shahani, U., & Manahilov, V. (2005). Illusory percepts of moving patterns due to discrete temporal sampling. *Neurosci Lett*, 375(1), 23-27.
- VanRullen, R., & Koch, C. (2003). Is perception discrete or continuous? *Trends Cogn Sci*, 7(5), 207-213.
- VanRullen, R., Reddy, L., & Koch, C. (2005). Attention-driven discrete sampling of motion perception. *Proc Natl Acad Sci U S A*, 102(14), 5291-5296.
- VanRullen, R., Reddy, L., & Koch, C. (2006). The continuous wagon wheel illusion is associated with changes in electroencephalogram power at approximately 13 Hz. *J Neurosci*, 26(2), 502-507.

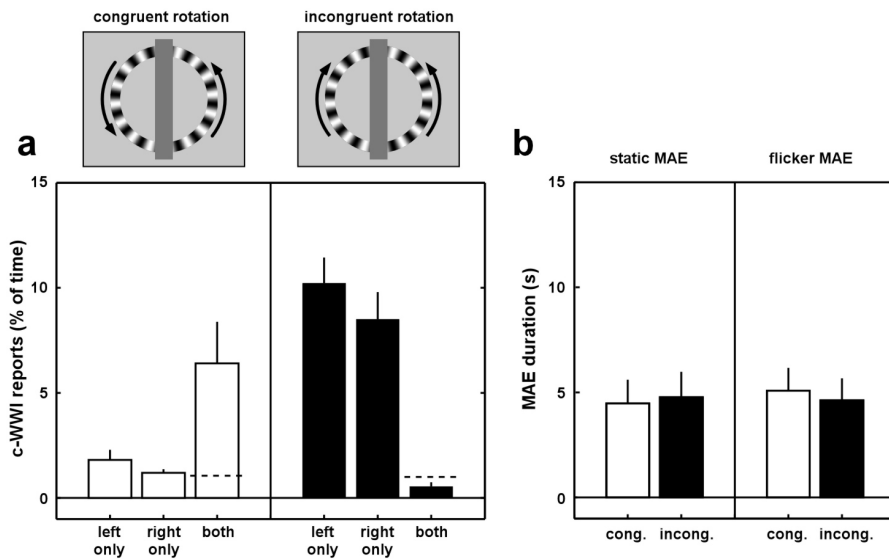


Figure 1. a. The ring stimulus was viewed for 40s at a time, during which subjects ($n=7$) continuously reported which of the two halves momentarily appeared to reverse its motion direction, i.e. to undergo a c-WWI. When rotation in the two half-rings was congruent (i.e. both clockwise or both counterclockwise), most perceptual reversals encompassed both half-rings (open bars). The amount of joint reversals was higher in this case ($p<0.02$) than would be predicted if the two halves reversed independently (dashed line). On the other hand, when the two half-rings rotated incongruently, most reversals were restricted to one or the other half (black bars). The amount of joint reversals in this case was significantly lower ($p<0.05$) than in the previous case. **b.** We verified that the congruency of the ring stimulus did not directly influence neuronal adaptation to the motion stimulus. At the end of each 40s trial, one of the two halves disappeared and the subjects estimated the duration of the motion aftereffect (MAE) for the half-ring that remained on the screen, which could be either static (static MAE, left), or flickering steadily at 10Hz (flicker MAE, right). In both cases, the duration of the MAE (reflecting the amount of adaptation) did not depend on whether the ring was congruent or not ($p>0.05$). Error bars represent s.e.m.

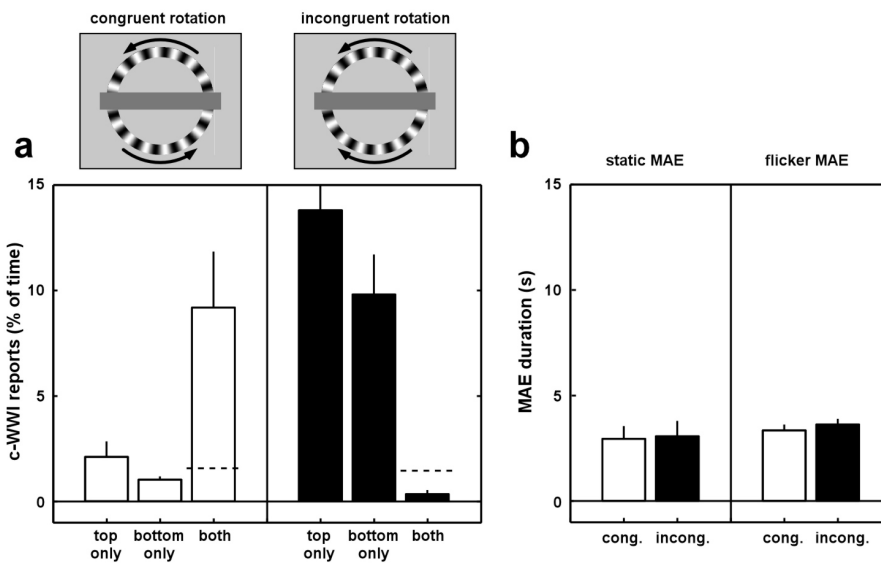


Figure 2. a. When the entire display was rotated by 90°, the motion information from the two half-rings could interact not only across but also within cortical hemispheres. Results (n=5) were similar to the previous case: during congruent rotation there was an increased tendency to perceive illusory motion reversals of both half-rings together (open bars); during incongruent rotation, on the other hand, most reversals were restricted to a single half (black bars). **b.** The amount of neuronal adaptation, measured by the static or the flicker MAE, did not differ significantly between the congruent and incongruent rotation trials ($p > 0.05$). Error bars represent s.e.m.