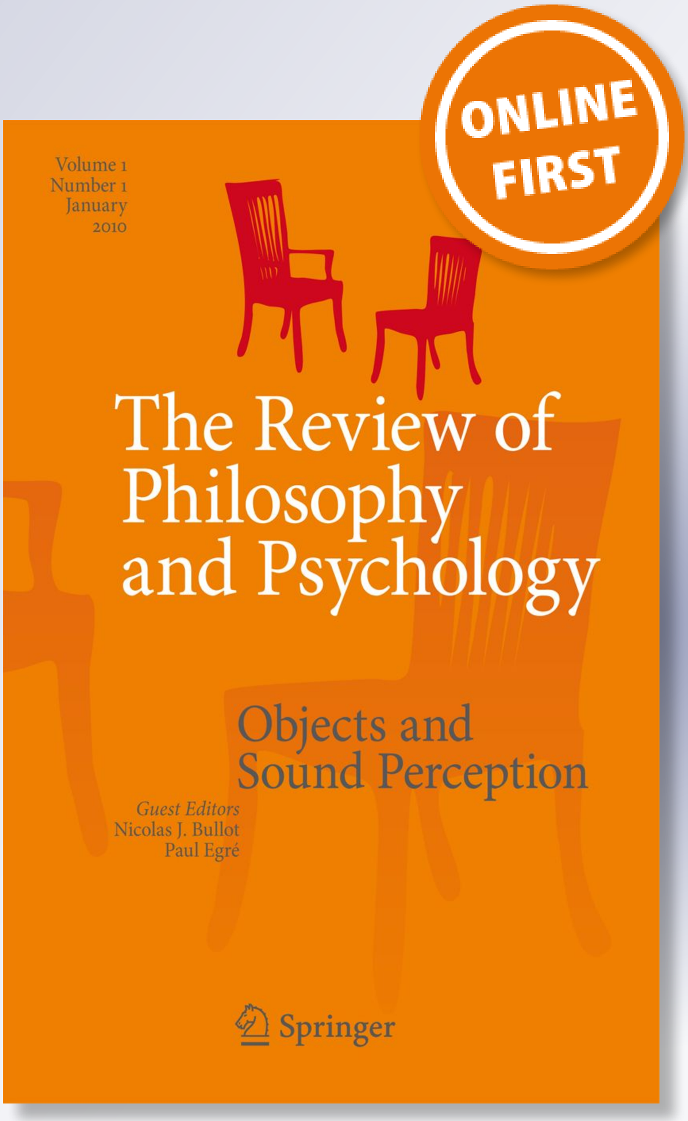
Visual Switching: The Illusion of Instantaneity and Visual Search

Nicoletta Orlandi

Review of Philosophy and Psychology

ISSN 1878-5158

Rev.Phil.Psych.
DOI 10.1007/s13164-012-0098-z



Springer

Your article is protected by copyright and all rights are held exclusively by Springer Science+Business Media B.V.. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your work, please use the accepted author's version for posting to your own website or your institution's repository. You may further deposit the accepted author's version on a funder's repository at a funder's request, provided it is not made publicly available until 12 months after publication.

Springer

Author's personal copy

Rev.Phil.Psych.
DOI 10.1007/s13164-012-0098-z

Visual Switching: The Illusion of Instantaneity and Visual Search

Nicoletta Orlandi

© Springer Science+Business Media B.V. 2012

Abstract This paper questions two *prima facie* plausible claims concerning switching in the presence of ambiguous figures. The first is the claim that reversing is an instantaneous process. The second is the claim that the ability to reverse demonstrates the interpretive, inferential and constructive nature of visual processing. Empirical studies show that optical and cerebral events related to switching protract in time in a way that clashes with its perceived instantaneity. The studies further suggest an alternative theory of reversing: according to such alternative, seeing the same thing in multiple ways is a matter of uncovering what is already present to the senses through visual search.

Our visual world is remarkably stable. We tend to see familiar objects and faces in just one way through time. There are images, however, that defy this stability. Looking at the Necker cube, for example, may cause one to first see the cube oriented in a certain way and then to see it oriented differently. The switch between these two ways of seeing the cube seems to be instantaneous and to be due to the interpretative, constructive and inferential nature of visual processing (Churchland 1989; Fodor 1984; Gregory 1970; Palmer 1999; Rock 1983). This paper questions both of these stances on the phenomenon. It reports on a number of studies that indicate that switching is a process that protracts in time and that involves a number of optical and neurological events. The paper further suggests an alternative account of reversing: according to this alternative, moving from seeing an image in one way to seeing it differently is done by engaging in visual search, where attention is paid to different parts of a figure or object and different visual percepts are processed as a result. Reversing does not involve interpreting or constructing. If there is any sense in which the visual system constructs the percept is that the system slowly finds out what is present to the senses and represents it. Evidence in cognitive and developmental psychology that supports this view is reviewed.

N. Orlandi (✉)
Philosophy and Cognitive Science, Rice University, Houston, TX, USA
e-mail: nico@rice.edu

Published online: 06 July 2012

Author's personal copy

1 The Illusion of Instantaneity

When we move from seeing the Necker cube in Fig. 1a as oriented one way to seeing it oriented differently, we may think that the move is instantaneous.

We seem capable of switching from one percept to the other quickly, as if going from one snapshot to the other of the figure. A number of studies, however, question this impression (Ito et al. 2003; Nakatani and van Leeuwen 2006; Nakatani et al. 2012). The studies are designed to reconstruct the timeline of oculomotor and brain events in the case of switching in the presence of multi-stable images, particularly Necker cube and Attneave triangles (Figs. 1a and 2).

Subjects had to press a button when they perceived that the figure had changed. Button presses (BPs) were used as a reference for when the switch had definitely occurred. It was then found that subjects need an average of 574 ms to produce a response to a change in visual stimulus. This was measured in a stimulus-initiated condition where two non-ambiguous versions of the Necker cube were presented alternatively (see Fig. 1b). The time of 574 ms may therefore also be considered as a reference for when the change has occurred and the subject prepares to signal it. Prior to (and following) 574ms a number of ocular and brain events related to switching take place: reconstructing a single chain of events that occurs every time there is a shift is difficult (more on this below). But, quite apart from these difficulties, preparatory cerebral and ocular events occur as early as 1100 ms prior to the button press (Ito et al. 2003; Nakatani et al. 2012). This means that the visual system prepares the switch approximately 500 ms prior to the initiation of the response on the part of the subject. This is a relatively long time for neural activity and it is surprising given the apparent instantaneity of reversing. For comparison, consider that it takes less that 400ms and generally only about 200ms to attain completion, that is, to perceive an occluded figure or object as complete when presented with only fragments of the figure or object (Sekuler and Palmer 1992).

More in detail, the events measured in the multistability experiments were primarily saccades and blinks for the ocular category, and synchronized activity for the cerebral category. The studies found a progressive reduction in blinks starting 1100ms prior to BPs (Ito et al. 2003). They also found a small increase in saccades around 1000 ms and then a significant peak in saccades 250-150 ms prior to BPs. These saccades presumably followed the switch and occurred in the period devoted to

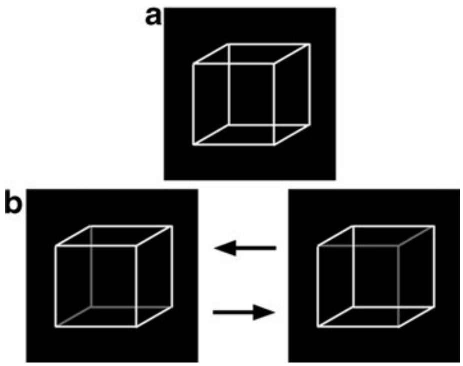


Fig. 1 Ambiguous Necker cube (**a**) and two non-ambiguous versions (**b**). Although not completely unambiguous, the figures in (**b**) are generally perceived in only one way (Nakatani et al. 2012)

Author's personal copy

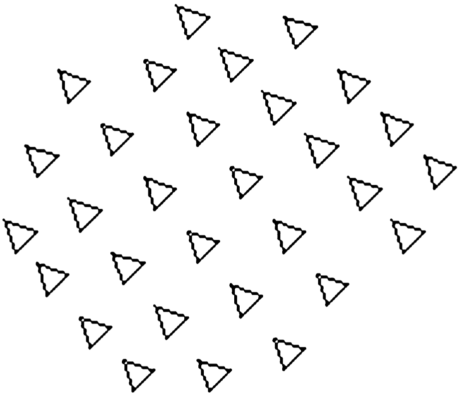


Fig. 2 Ambiguous triangles used in Attneave 1968 and Ito et al. 2003

the preparation of the response (between 574 ms and the button press). Post switch saccades were also found in a similar study using the Necker cube (van Dam and van Ee 2006).

As for cerebral activity measured with EEG, one study found transient periods of synchronized gamma band activity between parietal and frontal areas with a peak around 800-600 ms prior to BPs (Nakatani and van Leeuwen 2006). The activity was localized primarily in the right side of the brain and, in some subjects, accompanied by alpha band activity in occipital areas. A simplified schema of these results is reported in Fig. 3.

As mentioned earlier, reconstructing a single chain of events that occurs every time there is a switch has been difficult for a number of reasons. For one thing, some of the research reported here has not been replicated using the same conditions. Many studies that test ocular and brain events during reversals use binocular rivalry rather than ambiguous figures (Tong et al. 2006; van Eee et al. 2006). In rivalry two different stimuli are presented to each eye: rather than fusing them into a single image, we experience a shift between perceiving one and perceiving the other. Although this phenomenon is similar to perceptual shifts in the presence of ambiguous figures, it is also importantly different. In particular, it seems to be more "stimulus-driven": two stimuli are present and they, unsurprisingly, produce two different percepts. With multi-stable images, by contrast, the same stimulus is perceived in two or more ways. We shouldn't be surprised, then, if different cerebral and optic mechanisms are active in the two conditions.

Further, and most importantly, there is proven variability in ocular and brain events even during reversals induced by ambiguous figures. This variability is found both inter and intra individually. For example, frequent and infrequent shifters have different patterns of brain activity: frequent shifters display activity in frontal and occipital areas that is absent in infrequent shifters (Nakatani and van Leeuwen 2005). And, even within a single individual, ocular events have been found to vary

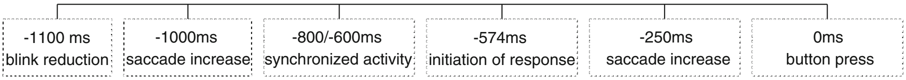


Fig. 3 Timeline of oculomotor and brain events leading to a button press

Author's personal copy

depending on contextual factors. One study found an increase in saccades prior to reversals only in the case of binocular rivalry, and not in the case of switching in the presence of ambiguous figures (van Dam and van Ee 2006). Another study, shows that, in contrast with what has been reported so far, a peak in blinking (rather than a decrease in blinking) occurs around 1000 ms prior to the button press (Nakatani et al. 2011). More research needs to be done to explain these discrepancies: in some cases, context can help explain the differences. For example, the peak in blinking is shown to occur primarily when the subject switches to the preferred, rather than to the non-preferred interpretation of the figure. Each multi-stable image tends to have a preferred percept, that is, a percept that most subjects perceive first and for longer stretches of time. In the Necker-cube case, observers tend to see first and for longer a cube oriented downwards. The peak in blinking occurs mostly when subjects switch to that percept rather than to the alternative one (Nakatani et al. 2011).

This body of evidence suggests that there may not be a single event or sequence of events associated with reversing in all cases and for all individuals: multistable perception may be a radically multiply realizable process, in the sense that different chains of neuro-physiological events constitute it. This demands further experiments and a reconsideration of studies that try to pinpoint a single process as *the* neural or optical correlate of switching (Ellis and Stark 1978; İşoğlu-Alkaç et al 1998; Isoglu-Alkaç et al 2000; Kornmeier and Bach 2004, 2005). But the body of evidence also shows that the chain of events – whatever its components – stretches in time in a way that is at odds with the perceived instantaneity of the process.

In the interest of doing some systematizing, it remains to explain why this time is needed and what happens during it. In other words, it would be good to have a plausible and comprehensive interpretation of the neuro-physiological data that also explained the non-instantaneity of the process. The dominant account of switches in cognitive psychology holds that an interpretation is involved. Palmer says:

> "Ambiguous figures demonstrate the constructive nature of perception because they show that perceivers interpret visual stimulation and that more than one interpretation is sometimes possible. If perception were completely determined by the light stimulating the eye, there would be no ambiguous figures because each pattern of stimulation would map onto a unique percept. This position is obviously incorrect. Something more complex and creative is occurring in vision, going beyond the information strictly given in the light that stimulates our eyes." (Palmer 1999, p. 10)

Understanding switching as the result of an interpretation conforms to the idea that all vision involves a type of inference or construction (Churchland 1989; Gregory 1970; Fodor 1984; Hohwy et al. 2008. Marr 1982, Rock 1983; Palmer 1999; Pylyshyn 1999). The visual system is said to reconstruct the layout of a scene from two-dimensional retinal images that do not uniquely determine the percept. The system is said to move from some initial retinal representations to representations of objects and scenes in stages and by using some stored knowledge of the world (Rock 1983, p. 15).

Richard Gregory, for example, writes:

> "Perception involves a kind of inference from sensory data to object-reality. Further, behavior is not controlled directly by the data, but by the solutions to the perceptual inferences from the data. This is clear from common experience:

Author's personal copy

> if I put a book on a table I do not prod the table first to check that it is solid. I act according to the *inferred* physical object-table – not according to the brown patch in my eye. So perception involves a kind of problem-solving – a kind of intelligence. Helmholtz spoke of perception in terms of 'unconscious inferences'." (Gregory 1970, p. 30)

The idea behind this way of understanding vision is that the visual system moves from some initial representations of sensory data (retinal representations) to representations of objects and scenes in stages and by using some implicit knowledge of the world. In Gregory's example, the system moves from "brown patches" in the eye to inferred representations of objects, by assuming that the causes of brown patches are solid physical entities. Implicit knowledge concerning, in the example, the environmental causes of brown patches, is employed to add information to the insufficient information contained in sensory data. The same style of explanation is applied to the case of multi-stability. Multi-stability is achieved when the same stimulus is interpreted in different ways by using different implicit knowledge of the world.

If being able to see the same thing in different ways demonstrates the constructive nature of visual processing then, presumably, the time taken to switch is employed to interpret the figure in a new way. The reported ocular and neurological events, however varied, should then be understood as components of an interpretive or constructive process. Some of the evidence described here, and some more evidence coming from other areas of psychology, however, suggests an alternative picture: rather than interpreting, switching is an automatic and hardwired process that centrally involves visual search and paying attention to different aspects of an object or figure. The next section expands on this proposal and explains why the evidence supports it.

2 Reversing as Finding Out

The last section presented a seemingly plausible way to explain visual multi-stability. The idea is to suppose that vision is an interpretive, constructive and inferential process, and to take switching in the presence of ambiguous figures to be a demonstration of this fact. Two or more percepts are the product of two or more interpretations of the same image where the visual system adds information to the one contained in retinal projections. In this sense, the approach regards vision as interpretive, a process similar to higher-order intellectual activities, such as reasoning, in being mediated by representations and informed by implicit knowledge.

Among those accepting this type of view, there is disagreement concerning how the interpretation is carried out. Virtually everybody believes that the inference is performed unconsciously (von Helmholtz 1962). Some suppose that it is performed subpersonally by dedicated and encapsulated visual modules (Fodor 1983; Pylyshyn 1999). Others, in contrast, suppose that past experience, background concepts and expectations can influence the outcome of the interpretation (Churchland 1989; Gregory 1970). Philosophers in the Kantian tradition may find this latter position attractive because it seems to conform to Kant's adage that perception without

Author's personal copy

concepts is blind. (Kant 1997 [1781]). And in psychology, this type of idea accords well with the view that top-down operations influence both stable and multi-stable vision. Irvin Rock's proposed "knowledge hypothesis", for example, holds that knowledge of the ambiguity and of the terms of the ambiguity is a necessary condition for switching (Girgus et al. 1977; Rock and Mitchener 1992; Rock et al. 1994).

Part of the reason for accepting the interpretive position, in both its modular and non-modular versions, is that there is a lack of plausible alternatives. The impression is that there is nothing else that could explain our ability to see the same thing in different ways. Since the figure in front of us stays the same, what changes must be something about us. The suggestion is that what changes is how we (or our visual modules) interpret the figure to be.

Reflection on the evidence, however, suggests a non-interpretive alternative. The alternative has it that the visual system is out to find what is already present to the senses, rather than adding information through an interpretation or a construction. Nothing particularly creative is involved in seeing the world in different ways: shifts are prompted by curiosity, rather than by interpreting. If talk of a construction is necessary at all, it should be understood as the gradual process of rendering what is present in the scene by careful search of it, rather than as a process where the visual system adds information to what is already available to it. More or less explicit precursors of this view can be found in both philosophy and psychology (Wollheim 1980; Gibson 1979; Noë 2004; Pylyshyn 1999). And the idea that visual search, together with other noise-inducing oculomotor processes, play an important role in multistability has already been used to explain binocular rivalry (Lee et al. 2007; van Brascamp et al. 2006).

A full defense of this alternative – and a full explanation of why it is a genuine alternative, together with its implications for the philosophy of mind and the computational theory of mind – is outside the scope of this paper (see Orlandi 2011b, c). Here, we can give a preliminary picture of the alternative, and, most importantly, focus on why the evidence supports it.

The basic idea is to view the visual system as situated in a structured environment that shaped, and continues to shape, the system to work in a certain way. Like other biological systems that are hardwired to operate as they do, the visual system is "naturally constrained" to process information the way that it does, not needing to encode implicit knowledge of the world (Pylyshyn 1999, p. 354).

Being "naturally constrained" means roughly being built to work in accordance with principles without having any implicit knowledge of the principles. Many biological and artifactual systems work by being so hardwired. Thinking of the visual system in this way explains how the system produces percepts without "unconscious inference." (Pylyshyn *ibid*). A naturally constrained visual apparatus produces representations of solid objects upon detection of brown patches at the eye (Gregory's example above) *not* by having any implicit knowledge of the environmental causes of brown patches, but by being hardwired to produce representations of solid objects in response to certain retinal projections. The fact that brown patches at the eye are, in our environment, typically caused by solid objects shaped the visual system to be so hardwired. The system is simply built to produce certain representational outputs given informational inputs.

The further proposal for understanding visual reversals in this framework is to think of the system as integrated in a larger system (paradigmatically a person) that is capable of performing visual search. This means that the larger system is capable of

Author's personal copy

directing, holding and shifting attention to different parts of a scene, figure or object: shifts in attention determine differences in how the stimulus is processed, and, derivatively, in what percepts are produced.

There is a growing literature concerning what attention consists in, and the different kinds of attentional mechanisms that humans (and non-humans) are endowed with (Mole 2010; Pylyshyn 2009). One thing that most theorists seem to agree about is that the allocation of attention has a selective function: attention selects, and, according to some, enhances (Posner and Peterson 1990), the part of the stimulus that needs to be processed, and it filters out unwanted information, that is, it ignores conditions that are not to be processed, because they are not relevant to the production of behavior (Desimone and Duncan 1995). The primary reason why attention has this function is because of limited processing capacities on the part of biological organisms.

This characterization appears to be true, for example, of focal attention, the ability used in tasks of object search and identification in cluttered environments. The direction of focal attention typically, but not invariably, coincides with the point of fixation, that is, the point where we foveate (Posner and Peterson 1990). Focal attention, whether understood as a type of spotlight or as a type of cognitive unison (Mole 2010), has a selective function: when the task is identifying an object among others, it selects the stimulus to be processed and it ignores stimuli that are irrelevant.

If we think of observers as capable of directing focal attention not only to whole objects and figures, but also to features of said objects and figures, we can begin to see how visual multistability occurs. By focusing attention on some features rather than others, the visual system gradually processes the stimulus given the features that are detected initially, and delivers percepts that are determined both by what is detected initially, and by how the system is naturally constrained to process information. A different percept may be produced if one focuses on different initial features.

Looking at the empirical evidence we can see why this is a plausible way of understanding visual multistability. Consider, first, the stimulation of frontal and parietal areas during reversals (Nakatani and van Leeuwen 2006). These areas are commonly associated with the control of selective attention (Britz et al. 2009; Kleinschmidt et al. 1998; Leopold and Logothetis 1999; Nakatani and van Leeuwen 2006; Sterzer and Kleinschmidt 2007; Tong et al. 2006). Derivatively, the studies suggest that attentional shifts are involved in switching, and this is confirmed by other studies. For example, it was found that, in adults, the part of a reversible figure that the observer focuses on determines which percept the viewer experiences (Chastain and Burnham 1975). And patients impaired in attentional control – in particular, patients with unilateral frontal brain damage – have greater difficulty in shifting than normal control subjects (Ricci and Blundo 1990).

The experiments on ocular activity reported above further support this picture. Saccades prior to, and during, reversals are positively associated with the process. They increase in the period leading to a shift (Ito et al. 2003), and in conditions where fixation instructions restrict them perceptual switching rate is considerably reduced (Glen 1940; Toppino 2003). Since saccades are closely associated with shifts in spatial attention (Slotnick and Yantis 2005) this again implicates attentional mechanisms in the switching process.

Blinks, on the other hand, have been found to reduce prior to reversing (Ito et al. 2003). They increase, primarily, when the subject shifts to the preferred percept

Author's personal copy

(Nakatani et al. 2011). This is consistent with the present proposal. Blinking is associated with a reduction in attentiveness: it increases when the visual stimulus does not require cognitive engagement, and it otherwise tends to decrease. The found diminution in blinking can then be explained as the need to deploy attentional resources for the processing of the non-preferred percept.

The present hypothesis is also able to account for the results reported by Rock (Girgus et al. 1977; Rock and Mitchener 1992; Rock et al. 1994) and mentioned above. Knowledge of the ambiguity and of the terms of the ambiguity positively correlates with visual reversals, and appears to be necessary for them, because when observers know that a figure is ambiguous they are more likely to engage in visual search. Further, by knowing the terms of the ambiguity observers are more likely to guide attention to those features of the figure that are suggestive of one of the known interpretations. Rock's results are then unsurprising if we think of multistability as induced by attentional changes.

Research in other areas of psychology further supports the view that shifts in attention guide reversing: bilingual children at 6 years of age are more likely to experience reversals than their monolingual peers (Bialystok and Shapero 2005). This result can be attributed to the fact that bilingual children develop control over selective attention earlier than monolingual children because they have to control two active language systems (Bialystok 2001; Bialystok and Martin 2004). Further, children younger than 5 are usually unable to reverse ambiguous figures spontaneously. One proposed explanation of this fact is that children are unable to reverse because they lack a theory of mind (Gopnik and Rosati 2001) that is, they lack, among other things, the capacity to interpret, and they don't understand that the same thing can be interpreted in different ways. But this proposal runs at odds with research on autistic children (Ropar et al. 2003): autistic children, while famously impaired in theory of mind tasks, appear to have no trouble seeing ambiguous figures in different ways. A more plausible way to explain children's inability to reverse is that prefrontal cortex, the brain area that controls attention, develops relatively late in humans reaching maturation only during adolescence (Diamond 2002). This is further suggested by a fact that has already been reported: frequent and infrequent switchers differ in brain activation precisely in these areas (Nakatani and van Leeuwen 2006).

This body of evidence indicates that the capacity to direct and hold attention to certain features of a scene is positively correlated with reversals. More work needs to be done to establish the exact role of attention in shifting, in particular to establish if attentional shifts are a necessary condition for reversing. Some studies suggest that they are not, but such studies did not test subjects on paradigmatic ambiguous figures, such as the Necker cube (Ilg et al. 2008; Pastukhov and Braun 2007).

Other research suggests that attention plays a more modest role in binocular rivalry than in multistability in the presence of ambiguous images (Meng and Tong 2004; van Eee et al. 2006; van Ee 2005; van Ee 2009). This, however, is relatively unsurprising. In rivalry two stimuli are presented to each eye, and although this set-up gives rise to shifts, the shifts are importantly "stimulus-driven": they don't seem to require any active visual search on the part of the subject. By contrast, we should expect attention to play a bigger role in perceptual multi-stability where the *same* stimulus is perceived in two or more ways. And this seems to be corroborated by the empirical results.

Author's personal copy

Further research is also needed to understand the relation between attentional shifts and neurophysiological processes such as adaptation. Adaptation of neural pathways to one of the two percepts has been used to explain both binocular rivalry and perceptual multistability (see Long and Toppino 2004 for a review of the literature). Combining adaptation to attentional shifts is in fact a promising way to provide the explanation of multistability that this article recommends: adaptation is a hardwired process that doesn't seem to be properly (and non-arbitrarily) describable as an inference, construction or interpretation.

Equally as interesting is the question of whether the model of multistability presented here is capable of explaining reversals in the non-visual senses. Multistable perception has been induced in both the auditory (Pressnitzer and Hupé 2006) and the tactile (Carter et al. 2008) modalities. While Pressnitzer et al. hypothesize that mechanisms of neural competition may regulate auditory multi-stability, more research needs to be done to test whether mechanisms of attentional control are active during episodes of auditory and tactile reversal.

Now, one may object in two ways to the kind of position introduced in this article. First, one may notice that the view that reversing is driven by attentional shifts is not incompatible with the idea that switching requires an interpretation. Performing visual search may just be what is needed, together with interpreting, in order to reverse. But further reflection reveals that an interpretation may not be needed after all.[1] The idea that switching requires interpreting is made plausible by the tendency to view the visual system as statically and internally reconstructing the visual layout of a scene from retinal stimulation that is, by itself, insufficient to determine the percept (Gregory 1970; Hohwy et al. 2008; Palmer 1999; Rock 1983). In the case of multistable images, the stimulus is blatantly insufficient to determine the percept because *one* image gives rise to two distinct perceptions. How could we get two things out of just one?

But this idea overlooks both the complex structure of the world, and the dynamic features of seeing. The visual system is located in a moving head and body, and its parts dynamically interact with a structured environment through eye and head movements. If we take this into consideration, then we can introduce the idea that interpreting the stimulus is not necessary because the stimulus itself can be explored in all its complexity to gather information about its features. Things in the world, including ambiguous images, look a certain way, and, in some cases, in more than one way. The visual system can be thought of as simply finding out how they look like without having to add any interpretation to it. This requires understanding perceivers as essentially located in the environment and capable of dynamically interacting with it. Instead of supposing that we internally reconstruct a model of the world from poor resources, we may instead think that we represent what's in the world by exploration of its richness.

Secondly, one may notice that visual search is itself a cognitive process that may require some interpretive skills.[2] There is some truth to this claim: shifting focal attention is a cognitive activity at least in the sense that, sometimes, it constitutes an intentional action that may be accompanied and preceded by thought. But conceding

---

[1] A full argument for this position is provided elsewhere: Orlandi 2011b, c.

[2] See, Orlandi 2011a for a similar worry.

Author's personal copy

this point does not amount to conceding that shifts in attention always, or even frequently, require cognitive and/or interpretive abilities. Attentional mechanisms are found in cognitively unsophisticated creatures, such as pigeons, that appear capable of identifying objects by sight and, more importantly, appear capable of reversing (Vettel et al. 2000). This indicates that the capacity to direct, hold and shift attention does not presuppose interpretive faculties.

Moreover, focal attention is often data-driven, that is, steered by intrinsic biases of the perceptual system (Remington 1980; Blaser et al. 1999). Local inhomogeneities, new objects, and objects that are larger, brighter, and faster moving are among the things that capture one's attention automatically, and independently of the subject's intentions (Desimone and Duncan 1995).

We reach similar conclusions when we look at so-called indexing mechanisms, that is, mechanisms that allow us to track objects prior to the initiation of visual processing. Their operation is overwhelmingly stimulus-driven: something may just "grab" a visual index (Pylyshyn 2009, p. 3).

If this is true, then we can understand multi-stability as initiated by, sometimes voluntary, shifts in attention, and as being carried out by a system that is not properly described as inferential. When we see, we simply find out how the world is, rather than constructing it to be a certain way. If talk of 'construction' is appropriate at all it should be taken to denote the gradual and dynamic rendering of what is in the world through visual curiosity. The relatively long time taken to switch, then, is not employed interpreting, but gradually searching for what is present in the scene. The optical and neurological events leading to a switch can then be interpreted as enabling visual search, and their variability as the way in which search is realized by the visual system in different contextual conditions.

Outlining the advantages of this view would take us too far astray: what is sufficient, for present purposes, is that we have a viable alternative to the common idea that reversing indicates the interpretive and constructive nature of perceiving. Interpreting is no longer the only option for explaining visual shifting.

Acknowledgements Some of the research cited in this paper was conducted by Cees van Leeuwen and Hironori Nakatani at the Riken Laboratory for Perceptual Dynamics in Tokyo, Japan. Thanks to both for giving me the opportunity to work with them.

References

Attneave, F. 1968. Triangles as ambiguous figures. *The American Journal of Psychology* 81: 447–453.

Bialystok, E. 2001. *Bilingualism in development: Language, literacy, and cognition*. Cambridge: Cambridge University Press.

Bialystok, E., and M. Martin. 2004. Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science* 7: 325–339.

Bialystok, E., and D. Shapero. 2005. Ambiguous benefits: The effect of bilingualism on reversing ambiguous figures. *Developmental Science* 8(6): 595–604.

Blaser, E., G. Sperling, and Z.L. Lu. 1999. Measuring the amplification of_attention. *Proceedings of the National Academy of Science, USA* 96: 11681–11686.

van Brascamp, J.W., R. Ee, A.J. Noest, R.H. Jacobs, and A.V. van den Berg. 2006. The time course of binocular rivalry reveals a fundamental role of noise. *Journal of Vision* 6: 1244–1256.

Author's personal copy

Britz, J., et al. 2009. Right parietal brain activity precedes perceptual alternation of bistable stimuli. *Cerebral Cortex* 19: 55–65.

Carter, O.Konkle, T. Wang, Q. Hayward, and C. Moore. 2008. Tactile rivalry demonstrated with ambiguous apparent motion quartet. *Current Biology* 18(14): 1050–4.

Chastain, G., and C.A. Burnham. 1975. The first glimpse determines the perception of an ambiguous figure. *Perception & Psychophysics* 17(3): 221–224.

Churchland, P. 1989. Perceptual plasticity and theoretical Neutrality: A reply to Jerry Fodor. In his *A Neurocomputational Perspective: the Nature of Mind and the Structure of Science*. Cambridge: MIT.

Desimone, R., and J. Duncan. 1995. Neural mechanism of selective attention. *Annual Review of Neuroscience* 18: 198–222.

Diamond, A. 2002. Normal development of pre-frontal cortex from birth to young adulthood: cognitive functions, anatomy, and biochemistry. In *Principles of frontal lobe functioning*, ed. D. Stuss and R. Knight, 466–503. New York: Oxford University Press.

Ellis, S.R., and L. Stark. 1978. Eye movements during the viewing of Necker cubes. *Perception* 7: 575–581.

Fodor, J.A. 1983. *Modularity of mind*. Cambridge: MIT.

Fodor, J.A. 1984. Observation reconsidered. *Philosophy of Science* 51: 23–43.

Gregory, R.L. 1970. *The intelligent eye*. New York: McGraw-Hill.

Gibson, J.J. 1979. *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Girgus, J., I. Rock, and R. Egatz. 1977. The effect of knowledge of reversibility on the reversibility of ambiguous figures. *Perception & Psychophysics* 22: 550–556.

Glen, J.S. 1940. Ocular movements in reversibility of perspective. *The Journal of General Psychology* 23: 243–281.

Gopnik, A., and A. Rosati. 2001. Duck or rabbit? Reversing ambiguous figures and understanding ambiguous representations. *Developmental Science* 4(2): 175–183.

Hohwy, J., A. Roepstorff, and K. Friston. 2008. Predictive coding explains binocular rivalry: An epistemological review. *Cognitiion* 108(3): 687–701.

Ilg, R., A.M. Wohlschlager, S. Burazanis, A. Woller, S. Nunnemann, and M. Muhlau. 2008. Neural correlates of spontaneous percept switches in ambiguous stimuli: An event-related functional magnetic resonance imaging study. *The European Journal of Neuroscience* 28(11): 2325–2332.

İşoğlu-Alkaç, Ü., C. Başar-Eroğlu, A. Ademoğlu, T. Demiralp, M. Miener, and M. Stadler. 1998. Analysis of the electroencephalographic activity during the Necker cube reversals by means of the wavelet transform. *Biological Cybernetics* 79: 437–442.

Isoglu-Alkaç, Ü., C. Başar-Eroglu, A. Ademoglu, T. Demiralp, M. Miener, and M. Stadler. 2000. Alpha activity decreases during the perception of Necker cube reversals: An application of wavelet transform. *Biological Cybernetics* 82: 313–320.

Ito, J., A.R. Nikolaev, M. Luman, M.F. Aukes, C. Nakatani, and C. van Leeuwen. 2003. Perceptual switching, eye movements, and the bus paradox. *Perception* 32: 681–698.

Kant, I. 1997 [1781/1787]. *Critique of Pure Reason* (trans. P. Guyer and A. Wood). Cambridge and New York: Cambridge University Press.

Kleinschmidt, A., C. Buchel, S. Zeki, and R.S.J. Frackowiak. 1998. Human brain activity during spontaneously reversing perception of ambiguous figures. *Proceedings of the Royal Society of London - Series B: Biological Sciences* 265: 2427–2433.

Kornmeier, J., and M. Bach. 2004. Early neural activity in Necker-cube reversal: Evidence for low-level processing of a gestalt phenomenon. *Psychophysiology* 41: 1–8.

Kornmeier, J., and M. Bach. 2005. The Necker cube – an ambiguous figure disambiguated in early visual processing. *Vision Research* 45: 955–960.

Long, G.M., and T.C. Toppino. 2004. Enduring interest in perceptual ambiguity: Alternating views of reversible figures. *Psychological Bulletin* 130(5): 748–768.

Lee, Blake, and Heeger. 2007. Hierarchy of cortical responses underlying binocular rivalry. *Nature NeuroScience* 10(2007): 1048–1052.

Leopold, D.A., and N.K. Logothetis. 1999. Multistable phenomena: Changing views in perception. *Trends in Cognitive Science* 3(7): 254–264.

Marr, D. 1982. *Vision: A computational investigation into the human representation and processing of visual information*. New York: W.H. Freeman and Company.

Meng, M., and F. Tong. 2004. Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision* 4: 539–551.

Mole, C. 2010. *Attention is cognitive unison*. Oxford University Press.

Nakatani, H., and C. van Leeuwen. 2005. Individual differences in perceptual switching rates; the role of occipital alpha and frontal theta band activity. *Biological Cybernetics* 93: 343–354.

Author's personal copy

Nakatani, H., and C. van Leeuwen. 2006. Transient synchrony of distant brain areas and perceptual switching in ambiguous figures. *Biological Cybernetics* 94: 445–457.

Nakatani, et al. 2012. Reversing as a dynamic process: variability of ocular and brain events in perceptual switching. *Journal of Consciousness Studies* 19(5-6): 117–140(24).

Nakatani et al. 2011. Precisely timed oculomotor and parietal EEG activity in perceptual switching. *Cognitive Neurodynamics* 5: 399–409.

Noë, A. 2004. *Action in perception*. Cambridge: MIT.

Orlandi, N. 2011a. The innocent eye: Seeing-as without concepts. *American Philosophical Quarterly* 48: 1.

Orlandi, N. 2011b. Embedded seeing-as: Multi-stable visual perception without interpretation. *Philosophical Psychology*. doi:10.1080/09515089.2011.579425.

Orlandi, N. 2011c. Embedded seeing: Vision in the natural world. *Nous*. doi:10.1111/14680068.2011.00845.

Palmer, S.E. 1999. *Vision science: Photons to phenomenology*. Cambridge: MIT.

Pastukhov, A., and J. Braun. 2007. Perceptual reversals need no prompting by attention. *Journal of Vision* 7 (10): 17.

Posner, M.I., and S. Peterson. 1990. The attention system of the human brain. *Annual Review of Neuroscience* 13: 25–42.

Pressnitzer, D., and J.M. Hupé. 2006. Temporal Dynamis of auditory and visual bistability reveal common principles of perceptual organization. *Current Biology* 11: 1351–1357.

Pylyshyn, Z. 1999. Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *The Behavioral and Brain Sciences* 22: 341–423.

Pylyshyn, Z. 2009. Perception, representation and the world: The FINST that binds. Available online at: uccs.rutgers.edu/faculty/pylyshyn/Dedrick&TrickFinal.pdf

Remington, R.W. 1980. Attention and saccadic eye-movements. *Journal of Experimental Psychology. Human Perception and Performance* 6: 726–744.

Ricci, C., and C. Blundo. 1990. Perception of ambiguous figures after focal brain lesion. *Neuropsychologia* 28: 1163–1173.

Rock, I. 1983. *The logic of perception*. Cambridge: MIT.

Rock, I.A., A.A. Gopnik, and S. Hall. 1994. Do young children reverse ambiguous figures? *Perception* 23: 635–644.

Rock, I., and K. Mitchener. 1992. Further evidence of failure of reversal of ambiguous figures by uninformed subjects. *Perception* 21: 39–45.

Ropar, D., P. Mitchell, and K. Ackroyd. 2003. Do children with autism find it difficult to offer alternative interpretations to ambiguous figures? *British Journal of Developmental Psychology* 21: 387–395.

Sekuler, A.B., and S.E. Palmer. 1992. Perception of partly occluded objects: A microgenetic analysis. *Journal of Experimental Psychology. General* 121(1): 95–111.

Slotnick, S.D., and S. Yantis. 2005. Common neural substrates for the control and effects of visual attention and perceptual bistability. *Cognitive Brain Research* 24: 97–108.

Sterzer, P., and A. Kleinschmidt. 2007. A neural basis for inference in perceptual ambiguity. *Proceedings of the National Academy of Sciences of the United States of America* 104: 323–328.

Tong, F., M. Meng, and R. Blake. 2006. Neural basis of binocular rivalry. *Trends in Cognitive Science* 10 (11): 502–511.

Toppino, T.C. 2003. Reversible-figure perception: Mechanisms of intentional control. *Perception & Psychophysics* 65: 1285–1295.

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

van Eee, R., A.J. Noest, J.W. Brascamp, and A.V. van den Berg. 2006. Attentional control over either of the two competing percepts of ambiguous stimuli revealed by a two-parameter analysis: Means do not make the difference. *Vision Research* 46: 3129–3141.

van Ee, R. 2005. Dynamics of perceptual bi-stability for stereoscopic slant rivalry and a comparison with grating, house-face, ad Necker cube rivalry. *Vision Research* 45: 29–40.

van Ee, R. 2009. Stochastic variations in sensory awareness are driven by noisy neuronal adaptation: Evidence from serial correlations in perceptual bistability. *Journal of the Optical Society of America. A* 26: 2612–2622.

Vettel, G., J.D. Haynes, and S. Pfaff. 2000. Evidence for multistability in the visual perception of pigeons. *Vision Research* 40(16): 2177–2186.

von Helmholtz, H. 1962[1867]. *Treatise on physiological optics*. Vol. 3. Translated from the German by J.P.C. Southall. New York: Dover.

Wollheim, R. 1980. Seeing-as, seeing in and pictorial representation. *Arts and its objects: With six supplementary essays*. Cambridge University Press.