

The Necker cube—an ambiguous figure disambiguated in early visual processing

Jürgen Kornmeier^{a,b}, Michael Bach^{a,*}

^a *Sektion Funktionelle Sehforschung, University Augenklinik Freiburg, Killianstr. 5, 79106 Freiburg, Germany*

^b *Institut für Grenzgebiete der Psychologie (IGPP), Wilhelmstr. 3a, 79098 Freiburg, Germany*

Received 12 August 2004; received in revised form 13 October 2004

Abstract

How can our percept spontaneously change while the observed object stays unchanged? This happens with ambiguous figures, like the Necker cube. Explanations favor either bottom–up factors in early visual processing, or top–down factors near awareness. The EEG has a high temporal resolution, so event related potentials (ERPs) may help to throw light on these alternative explanations. However, the precise point in time of neural correlates of perceptual reversal is difficult to estimate. We developed a paradigm that overcomes this problem and found an early (120ms) occipital ERP signal correlated with endogenous perceptual reversal. Parallels of ambiguous-figure-reversal to binocular-rivalry-reversals are explored.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: Perceptual ambiguity; Ambiguous figures; Necker cube; EEG; ERP; Event related potentials

1. Introduction

Ambiguous figures, e.g. the Necker cube (Necker, 1832) or the drawing “Old/Young Woman” (Boring, 1930) (Fig. 1) are static pictures, which strikingly change their appearance during prolonged viewing: a sudden and unavoidable “mental switch” occurs whenever two or more interpretations of a given picture are equally likely.

Most hypothetical explanations about the neural processes underlying spontaneous perceptual reversals of ambiguous figures fall into two classes, emphasizing either bottom–up, or top–down factors.

Bottom–up approaches assume the perceptual reversal results from ‘passive’ adaptation early in the visual stream (e.g. Köhler, 1940; Toppino & Long, 1987). Support comes from numerous studies: reversal rates were

measured while varying physical features of ambiguous figures (e.g. Ammons & Ammons, 1963; Brigner & Deni, 1990; Lindauer & Lindauer, 1970; Washburn, Mallat, & Naylor, 1931), varying the ambiguous figure’s presentation mode (Orbach, Ehrlich, & Heath, 1963), or while presenting an unambiguous preadapting stimulus preceding the ambiguous one at the same position (e.g. Carlson, 1953; Emerson, 1979; Hochberg, 1950), or presenting the stimulus at different positions in the visual field (e.g. Orbach & Zucker, 1965; Toppino & Long, 1987).

Top–down approaches emphasize attentional or expectational factors, acting in a centrally governed ‘active’ decision process near awareness and consequently later in the visual hierarchy (e.g., Rock, Hall, & Davis, 1994; Vickers, 1972). Support comes from effects of experience, learning and cognitive state of the subjects (Ammons, 1954; Cohen, 1959; Flugel, 1913; Hochberg & Peterson, 1987; Pelton & Solley, 1968; Peterson & Hochberg, 1983; Reisberg & O’Shaughnessy, 1984; Rock, Gopnik, & Hall, 1994; Rock et al., 1994; Spitz

* Corresponding author. Tel.: +49 761 270 4060; fax: +49 761 270 4052.

E-mail address: michael.bach@uni-freiburg.de (M. Bach).

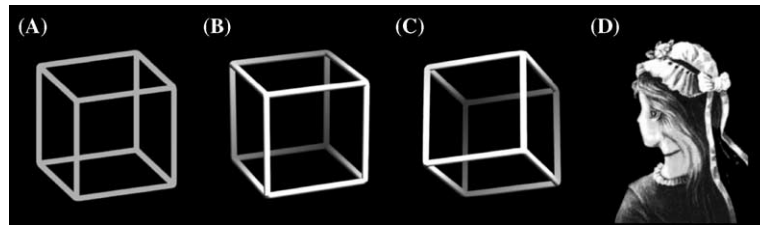


Fig. 1. Ambiguous figures and unambiguous variants. (A) “Necker cube” according to Necker (1832). (B,C) Unambiguous variants with marked depth cues, based on a drawing model incorporating shading, central projection, and aerial perspective. (D) “Old and Young Woman” according to Boring (1930).

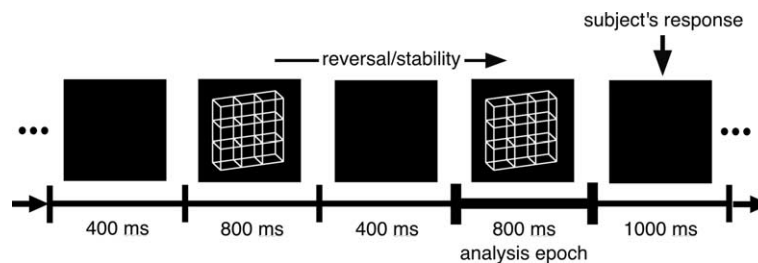


Fig. 2. Experimental paradigm. The “Necker lattices” were presented successively in an onset/offset mode. Subjects compared the front-back orientation of any given stimulus with the preceding one. In different conditions they indicated either “perceptual reversal” or “perceptual stability” via button press during the off-periods. All traces depicted in Fig. 3 were acquired during the 800ms analysis epoch.

& Lipman, 1962; Strüber & Stadler, 1999; Vickers, 1972). A detailed review about experimental results and potential explanations of the phenomenon can be found in Long and Toppino (2004).

Event related potentials (ERPs) as correlates of perceptual processes may help to shed new light on the ongoing debate, due to their high temporal resolution. However, because of their low signal-to-noise ratio, they are only visible after averaging many EEG epochs, time-locked to the relevant events. With ambiguous figures no external time reference for the instance of perceptual reversal is available. Previous ERP studies of ambiguous figures thus relied on backward-averaging with respect to the manual response of the subject. They found P300-like responses (Basar-Eroglu, Strüber, Stadler, & Kruse, 1993) as correlates of the perceptual reversal. Since the P300 is well known as a “cognitive component” (Verleger, 1997), these late positive components were interpreted as evidence for the top-down approach. Unfortunately, marked temporal jitter is introduced when the point in time of an endogenous event—like the perceptual reversal of ambiguous figures—is estimated via manual response. This jitter will obscure brief ERP components such as the typically sharp early peaks, while having little effect on broad later components.

We have developed an experimental paradigm (Kornmeier & Bach, 2004) which allows to time-lock on endogenous events, and thus to identify early ERP com-

ponents specific for Necker cube reversals, as follows: we presented Necker cubes intermittently with a temporal regime (800 ms on, 400 ms off) that optimally induced perceptual reversals at stimulus onset (Fig. 2). Using stimulus onset as time reference for averaging we uncovered an early reversal-related negative ERP component (“Reversal Negativity”, “RN”, Kornmeier & Bach, 2004). We also recognized an earlier positive component at the occipital electrode. Since this ERP signal was very weak, and we were concerned about significance inflation by multiple testing, we ignored it at that time. With some ramifications and more electrodes we therefore repeated the experiment with new subjects. The additional electrodes (O1 and O2) picked up this earlier component. We here report—in addition to a full reproduction of the Reversal Negativity—an even earlier, reversal-associated ERP component, the Reversal Positivity at 120 ms.

2. Methods

2.1. Subjects

Twenty-one subjects (aged 20–31 years) with normal or corrected-to-normal visual acuity participated in the experiment. All subjects gave informed written consent to participate and were naive as to the specific experimental question. The studies were performed in accor-

dance with the ethical standards laid down in the Declaration of Helsinki (World Medical Association, 2000).

2.2. Stimuli

Stimuli were perceptually ambiguous “Necker Lattices” (Fig. 2) with a viewing angle of $7.5^\circ \times 7.5^\circ$ and a luminance of 20 cd/m^2 on a background of 0.01 cd/m^2 . In all experimental trials, successive stimuli were jittered in virtual 3D-space over $\pm 12^\circ$ of both elevation and azimuth angle, resulting in altogether seven ambiguous variants; this avoided afterimages and trivial local cues. A small cross in the centre of the screen served as fixation target.

2.3. Paradigm

The stimuli were presented in an onset/offset mode with 800 ms presentation time followed by a blank screen for 400 ms. Subjects viewed the stimuli monocularly. They compared the front-back orientation of any given stimulus with that of the preceding one (“observation sequence”). In different experimental conditions (A and B) they indicated either “reversal” or “stability” by button press after stimulus offset (Fig. 2); this allowed to factor out the task. The experiment was balanced in an ABBA scheme and comprised an equal number of “reversal” and “same” trials. The instruction explained that the criterion to press the button was not met, if the reversal did not occur at stimulus onset or if more than one reversal occurred during the presentation period [both situations occurred very rarely]. The instruction also included a statement that perceived reversal and stability were equally important. After any button press the current blank screen interval was extended to 1000 ms and the observation sequence was restarted, so that stimuli never were compared across intervening manual reactions. This minimized interference from motor tasks between successive stimuli of interest.

Subjects were instructed to gaze at the fixation cross, not to provoke perceptual reversals, and only to respond if they were certain about their percept. If they responded by accident already during the presentation, the computer program emitted a warning tone and the current EEG epoch was excluded from further analysis. EEG was recorded from 11 gold-cup scalp electrodes at O1, O2, Oz, P3, P4, Pz, C3, C4, Cz, Fz and Fpz (American Encephalographic Society, 1994), referenced to averaged ears. Vertical EOG electrodes detected eye blinks.

2.4. Data analysis

The recordings were automatically checked for artifacts from eye movements and amplitude excursions

exceeding $\pm 100 \mu\text{V}$. They were averaged (around 120 trials for each condition) and digitally filtered with a low-pass of 25 Hz.

Peak amplitudes of the ERP components were measured relative to baseline, which was defined as the average from 40 ms before to 60 ms after stimulus onset. The first prominent ERP component in the difference traces was identified by repeated significance in a “running *t*-test” (testing each sample point against zero, Guthrie, 1990). The component was not unequivocally identifiable in all subjects. As a solution, we first determined the unique peak position in the grand mean across all subjects. Then, within a $\pm 10 \text{ ms}$ window around this position, in each subject the largest excursion was selected.

3. Results

Fig. 3A illustrates group average data for these two conditions (perceptual reversal and non-reversal = stability) arranged according to the scalp locations of the electrodes. Clear ERPs to the onset of the stimulus are present at all locations. The earliest visual ERP is a positivity 80 ms (P80) after stimulus onset and most prominent at the occipital pole (O1, Oz, O2; Fig. 3A). Small but highly significant differences between reversal and stability are more easily seen in the difference traces (reversal minus stability, Fig. 3B). The earliest significant ERP difference associated with perceptual reversal, the “Reversal Positivity”, peaks around 120 ms after stimulus onset, and is most prominent at occipital electrode positions ($p < 0.001$) and also present but weaker, at parietal locations. It is followed by a negative deflection, the “Reversal Negativity” at 250 ms ($p < 0.001$), most prominent at occipital and parietal electrode positions, and also visible but much weaker at central locations. Subsequent positivities maximal at frontopolar (410 ms) and parietal (470 ms) locations can also be observed.

4. Discussion

We recorded the ERP in two visually identical conditions which only differed in the percept, as reported by the subject with subsequent button presses. The difference between the ERPs in these two conditions contains various structures, beginning with an early (120 ms) positive ERP component (Reversal Positivity). This represents a new finding. We tentatively denote this early positive ERP component to the endogenous perceptual reversal of an ambiguous Necker lattice as “Reversal Positivity”. It is most prominent at occipital electrodes and present, but weaker at parietal positions. Subsequent to the Reversal Positivity, an occipital/parietal Reversal Negativity at 250 ms can be seen. This is a

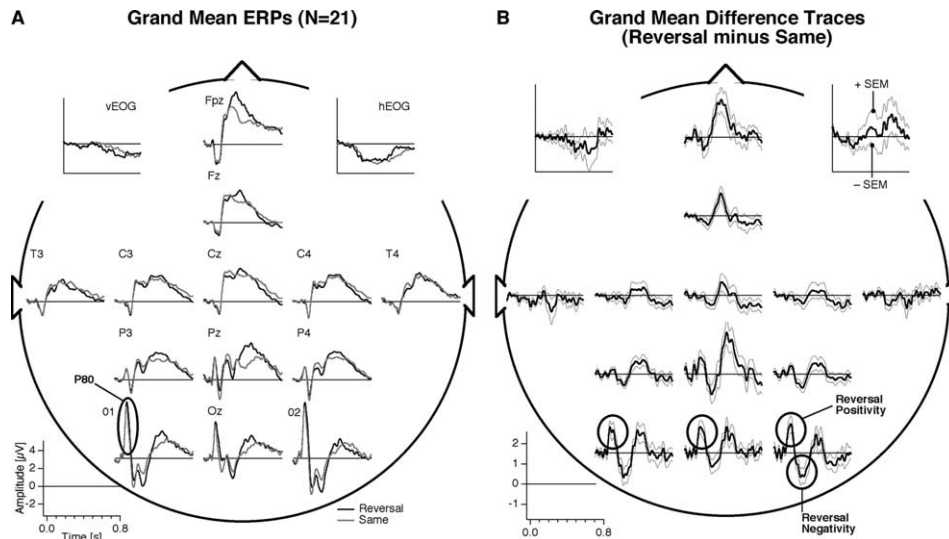


Fig. 3. ERP correlates of perceptual reversal. (A) Grand mean ERPs across all subjects. The grey traces indicate trials with perceptual reversal, the black traces represent trials with stable perception. (B) Difference traces (bold lines) \pm SEM (thin lines) [‘Reversal’-ERPs minus ‘Stability’-ERPs from (A)]. The averaged traces are geographically arranged on a symbolized scalp, including horizontal (top left) and vertical (top right) EOG traces. In (A) the electrode positions are denoted on top of each graph. Each trace represents the grand mean of 21 subjects (except electrodes C3/C4 with $n = 16$, and T3/T4 with $n = 10$).

replication of a previously reported finding (Kornmeier & Bach, 2004).

Let us first discuss possible shortcomings of the underlying assumptions. The paradigm aimed to evoke endogenous reversals using discontinuous presentation. But are such reversals an adequate model for reversals during continuous presentation? As argued by Kornmeier and Bach (2004), findings from Orbach et al. (1963) suggest they are: they presented Necker cubes discontinuously and measured reversal rates as a function of the on- and off-times. A smooth monotonous change was found at the transition from interrupted to continuous presentation.

The paradigm aimed to temporally entrain endogenous reversals using discontinuous presentation. But is the stimulus onset, used as time reference for averaging EEG data, a reliable marker for the reversal instance? Any temporal jitter of the endogenous process convolves the ERP components with the jitter distribution function. When this is wide, e.g. in case of reaction times, sharp ERP components are obliterated, as demonstrated in Kornmeier and Bach (2004). In the present situation, the mere existence of the relatively sharp and symmetric peak, the Reversal Positivity at 120ms, argues that a sizable number of endogenous reversals were temporally entrained to stimulus onset within a time window of less than ± 30 ms. We estimated this time window as follows: let us denote the early occipital ERP components, the ones before subtraction (Fig. 3A), as ‘‘low-level ERPs’’ (effectively, these are VEPs). Any such low-level ERPs are by necessity fully synchronized to the stimulus onset. The earliest low-level ERP here is the occipital positivity at 80ms latency (‘‘P80’’) and has a peak width of ± 20 ms. In comparison, the earliest ‘‘high level’’, reversal-associated structure, the Reversal Positivity, has a peak width of ± 35 ms (Fig. 3B); this must result from a convolution of the jitter producing the P80-width with the jitter of the process generating the Reversal Positivity. Assuming a Gaussian peak shape, the unknown jitter width is equal to $\sqrt{(35^2 - 20^2)} \approx 29$ ms (this formula is known from statistics: additivity of variances). Giving round numbers, we conclude that the endogenous reversals were indeed entrained within ± 30 ms.

The Reversal Positivity at 120ms can be considered as ‘‘early’’ when compared to the earliest low-level peak (P80, Fig. 3A), to other visual ERPs (Regan, 1989), and to the neural timing of V1 cells in awake non-human primates (Nowak & Bullier, 1997). This early occurrence of the Reversal Positivity combined with its relatively sharp peak and its occipital scalp distribution suggests that the mechanisms of endogenous perceptual reversals are situated early in visual processing.

Later ERP components are also prominent in the difference traces (Fig. 3B). Around 240ms, a ‘‘Reversal Negativity’’ at occipital and parietal derivations can be observed. Further, nearly a quarter of a second after the Reversal Positivity, we see a marked frontopolar positivity at 350ms, followed by a parietal positivity at 410ms. The positive components probably correspond to the P300-like positivity to ambiguous figures as reported by Basar-Eroglu et al. (1993) which was maximal at 250ms before the subject’s reaction. In the present experiments, the occipital Reversal Positivity peaks 300ms before the parietal positivity, so in Basar-Eroglu et al.’s traces a Reversal Positivity should also appear

at a similar relative position; presumably at 250 ms + 300 ms = 550 ms before the manual response. Why is the Reversal Positivity not seen in their data near this point in time? Basar-Eroglu et al. (1993) averaged backwards from the reaction time. It appears likely that convolution of the reaction-time distribution with the ERP structures has obliterated any sharp early peaks leaving later (and broader) components less affected, as discussed in Kornmeier and Bach (2004).

The EEG timing information thus suggests that disambiguation is initiated early in the visual processing chain. Other findings, like the influence of stimulus features, or presentation mode (Orbach et al., 1963) on reversal rates also point to low-level factors. Since this cumulation of evidence seems to tip the balance in favor of the bottom-up approach, how can results about high-level influences on reversal rates, e.g. the subject's attitude towards the stimulus (Horlitz & O'Leary, 1993), be explained?

The situation may be comparable to that of binocular rivalry. In this well-known phenomenon dissimilar images are presented to corresponding regions of the two eyes and compete for perceptual dominance. Binocular rivalry and perceptual reversals of ambiguous figures have several features in common, some examples are: the durations of dominance-phases for a given ambiguous image seem to be gamma-distributed (Borsellino, De Marco, Allazetta, Rinesi, & Bartolini, 1972; Lehky, 1995); there is a high inter-subject variability in reversal rates (Aafjes, Hueting, & Visser, 1966; Borsellino et al., 1972); reversal rates can be sizably influenced by physical properties of the stimulus (e.g. Ammons & Ammons, 1963; Walker, 1978) and through voluntary control by the subjects (e.g. Horlitz & O'Leary, 1993; Lack, 1974). Finally, Kaernbach, Schroger, Jacobsen, and Roeber (1999) reported in an EEG study of binocular rivalry similar results to those presented here. It was recently suggested that binocular rivalry can be induced at several processing steps, depending on the experimental setting (Blake & Logothetis, 2002; Tong, 2001). Applying this to the Necker phenomenon, perceptual reversal could also take place at several sequential processing steps. In the present case, where the subjects had a passive attitude whilst fixating the ambiguous stimulus, bottom-up processing supposedly dominated the phenomenon. Other experimental conditions (e.g. changing the subject's instruction) might then affect only later reversal-related ERP structures, but not the early "Reversal Positivity".

We conclude: perceptual reversals can be initiated during the first visual processing steps as early as 120 ms. High-level processing can modulate the perceptual process at later stages, as others have shown (see above). The beginning of visual awareness is estimated to occur in the time range of the P300 (Desmedt & Tomberg, 1995; Posner, 1978; Verleger, 1997), thus disam-

biguation of ambiguous figures seems to occur already 200–300 ms before perceptual awareness is established.

Acknowledgment

Support by the Deutsche Forschungsgemeinschaft (BA 977-11) is gratefully acknowledged.

References

- Aafjes, M., Hueting, J. E., & Visser, P. (1966). Individual and interindividual differences in binocular retinal rivalry in man. *Psychophysiology*, 3(1), 18–22.
- American Encephalographic Society (1994). Guideline thirteen: Guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, 11, 111–113.
- Ammons, C. H., & Ammons, R. B. (1963). Perspective reversal as affected by physical characteristics of Necker Cube drawings. *Proceedings of the Montana Academy of Sciences*, 23, 287–302.
- Ammons, R. B. (1954). Experimental factors in visual form perception: I. Review and formulation of problems. *Journal of General Psychology*, 67, 708–710.
- Basar-Eroglu, C., Strüber, D., Stadler, M., & Kruse, E. (1993). Multistable visual perception induces a slow positive EEG wave. *International Journal of Neuroscience*, 73, 139–151.
- Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience*, 3(1), 13–21.
- Boring, E. G. (1930). A new ambiguous figure. *American Journal of Psychology*, 42, 444–445.
- Borsellino, A., De Marco, A., Allazetta, A., Rinesi, S., & Bartolini, B. (1972). Reversal time distribution in the perception of visual ambiguous stimuli. *Kybernetik*, 10(3), 139–144.
- Brigner, W. L., & Deni, J. R. (1990). Depth reversals with an equiluminant Necker cube. *Perceptual and Motor Skills*, 70(3 Pt 2), 1088.
- Carlson, V. R. (1953). Satiation in a reversible perspective figure. *Journal of Experimental Psychology—General*, 45, 442–448.
- Cohen, L. (1959). Rate of apparent change of a Necker cube as a function of prior stimulation. *American Journal of Psychology*, 72, 327–344.
- Desmedt, J. E., & Tomberg, C. (1995). Consciousness. *Electroencephalography and Clinical Neurophysiology Supplement*, 44, 227–234.
- Emerson, P. (1979). Necker cube: Duration of preexposure of an unambiguous form. *Bulletin of the Psychonomic Society*, 14(6), 397–400.
- Flugel, J. C. (1913). The influence of attention in illusions of reversible perspective. *Atlantic Monthly*, 131, 357–397.
- Guthrie, D. (1990). Intergroup and intrasubject principal component analysis of event-related potentials. *Psychophysiology*, 27(1), 111–119.
- Hochberg, J., & Peterson, M. (1987). Piecemeal organization and cognitive components in object perception. *Journal of Experimental Psychology—General*, 116, 370–380.
- Hochberg, J. E. (1950). Figure-ground reversal as a function of visual satiation. *Journal of Experimental Psychology—General*, 40, 682–686.
- Horlitz, K. L., & O'Leary, A. (1993). Satiation or availability? Effects of attention, memory, and imagery on the perception of ambiguous figures. *Perception & Psychophysics*, 53(6), 668–681.
- Kaernbach, C., Schroger, E., Jacobsen, T., & Roeber, U. (1999). Effects of consciousness on human brain waves following binocular rivalry. *Neuroreport*, 10(4), 713–716.

- Köhler, W. (1940). *Dynamics in psychology*. New York: Liverlight.
- Kornmeier, J., & Bach, M. (2004). Early neural activity in Necker-cube reversal: Evidence for low-level processing of a gestalt phenomenon. *Psychophysiology*, *41*, 1–8.
- Lack, L. C. (1974). Selective attention and the control of binocular rivalry. *Perception & Psychophysics*, *15*, 193–200.
- Lehky, S. R. (1995). Binocular rivalry is not chaotic. *Proceedings of the Royal Society of London Series B—Biological Sciences*, *259*(1354), 71–76.
- Lindauer, M. S., & Lindauer, J. G. (1970). Brightness differences and the perception of figure-ground. *Journal of Experimental Psychology*, *84*(2), 291–295.
- Long, G. M., & Toppino, T. C. (2004). Enduring interest in perceptual ambiguity: Alternating views of reversible figures. *Psychological Bulletin*, *130*(5), 748–768.
- Necker, L. A. (1832). Observations on some remarkable Optical Phenomena seen in Switzerland; and on an Optical Phenomenon which occurs on viewing a Figure of a Crystal or geometrical Solid. *The London and Edinburgh Philosophical Magazine and Journal of Science*, *1*(5), 329–337.
- Nowak, L. G., & Bullier, J. (1997). The timing of information transfer in the visual system. In J. H. Kaas, K. Rockland, & A. Peters (Eds.), *Cerebral cortex* (pp. 205–241). New York: Plenum.
- Orbach, J., Ehrlich, D., & Heath, H. (1963). Reversibility of the Necker cube: I. An examination of the concept of “satiation of orientation”. *Perceptual and Motor Skills*, *17*, 439–458.
- Orbach, J., & Zucker, E. (1965). Reversibility of the Necker cube: VI. Effects of interpolating a non-reversing cube. *Perceptual and Motor Skills*, *20*, 470–472.
- Pelton, L. H., & Solley, C. M. (1968). Acceleration of reversals of a Necker cube. *American Journal of Psychology*, *81*(4), 585–588.
- Peterson, M. A., & Hochberg, J. (1983). Opposed-set measurement procedure: A quantitative analysis of the role of local cues and intention in form perception. *Journal of Experimental Psychology*, *9*, 183–193.
- Posner, M. I. (1978). *Chronometric explorations of mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Regan, D. (1989). *Human brain electrophysiology. Evoked potentials and evoked magnetic fields in science and medicine*. New York: Elsevier.
- Reisberg, D., & O’Shaughnessy, M. (1984). Diverting subjects’ concentration slows figural reversals. *Perception*, *13*(4), 461–468.
- Rock, I., Gopnik, A., & Hall, S. (1994). Do young children reverse ambiguous figures? *Perception*, *23*(6), 635–644.
- Rock, I., Hall, S., & Davis, J. (1994). Why do ambiguous figures reverse? *Acta Psychologica (Amst)*, *87*(1), 33–59.
- Spitz, H. H., & Lipman, R. S. (1962). Some factors affecting Necker cube reversal rate. *Perceptual and Motor Skills*, *15*, 611–625.
- Strüber, D., & Stadler, M. (1999). Differences in top-down influences on the reversal rate of different categories of reversible figures. *Perception*, *28*, 1185–1196.
- Tong, F. (2001). Competing theories of binocular rivalry: A possible resolution. *Brain and Mind*, *2*, 55–83.
- Toppino, T. C., & Long, G. M. (1987). Selective adaptation with reversible figures: Don’t change that channel. *Perception & Psychophysics*, *42*(1), 37–48.
- Verleger, R. (1997). On the utility of P3 latency as an index of mental chronometry. *Psychophysiology*, *34*(2), 131–156.
- Vickers, D. (1972). A cyclic decision model of perceptual alternation. *Perception*, *1*(1), 31–48.
- Walker, P. (1978). Binocular rivalry: Central or peripheral selective processes? *Psychological Bulletin*, *85*, 376–389.
- Washburn, M. F., Mallat, H., & Naylor, A. (1931). The influence of the size of an outline cube on the fluctuations of its perspective. *American Journal of Psychology*, *43*, 484–489.
- World Medical Association (2000). Declaration of Helsinki: Ethical principles for medical research involving human subjects. *JAMA*, *284*(23), 3043–3045.