Chapter 1

Title goes here

1.1 Introduction

TODO, pull from proposal and proposal-r1.

1.2 Methods

1.2.1 Participants

A total of 58 students from the University of Colorado Boulder participated in the experiment (ages 18-28 years, mean=21; 31 male, 27 female). EEG was recorded from 29 of the participants while they completed the experiment. The remaining 29 participants completed a solely behavioral experiment without EEG recording. All participants were right-handed and reported normal or corrected-to-normal vision. Participants either received course credit or payment of $15 per hour as compensation for their participation. Informed consent was obtained from each participant prior to the experiment in accordance with Institutional Review Board policy at the University of Colorado.

1.2.2 Stimuli

Novel "paper clip" objects similar to those used in previous studies (Bulthoff & Edelman, 1992; Edelman & Bulthoff, 1992; Logothetis, Pauls, Bulthoff, & Poggio, 1994; Logothetis, Pauls, & Poggio, 1995; Sinha & Poggio, 1996) were created using MATLAB. Eight vertices were placed randomly on the surface of a sphere of unit radius and then joined together with line segments. The last and first vertex were also joined to form a closed loop so that line segment terminations were not a salient feature (Balas & Sinha, 2009). Objects were constrained to exclude extremely acute angles between successive segments (less than 20 degrees) and were approximately rotationally balanced (center of mass within 10% of the origin). Objects were were rotated completely about their vertical axis in steps of 12 degrees and rendered to bitmap images under an orthographic projection. A total of 16 objects were created using this procedure, yielding 480 images (30 images per object). Object examples are shown in Figure 1.1.

1.2.3 Procedure

Participants observed an entraining sequence of rotated views of a random object and performed a same-different judgement about a probe stimulus. On each trial, a view was randomly selected as the initial view of the sequence followed by seven additional views spaced 24 degrees apart (Figure 1.2A, blue tick marks). Thus, the eight view entraining sequence spanned 168 degrees of the object. The entraining sequence was either presented in order (i.e., spatially predictable) or randomized. Following the entraining sequence after a 200 ms blank was a probe stimulus consisting of either an unseen view from the entraining object or a novel distractor. Unseen views were randomly sampled from the 12 degree interpolations between views of the entraining sequence (Figure 1.2A, magenta tick marks) and from outside of the span of the entraining sequence in increments of 24 degrees (Figure 1.2A, green tick marks).

Distractors were created from the original target objects by randomly selecting new spherical coordinates for six of the eight vertices and re-rendering them to bitmap images using the same

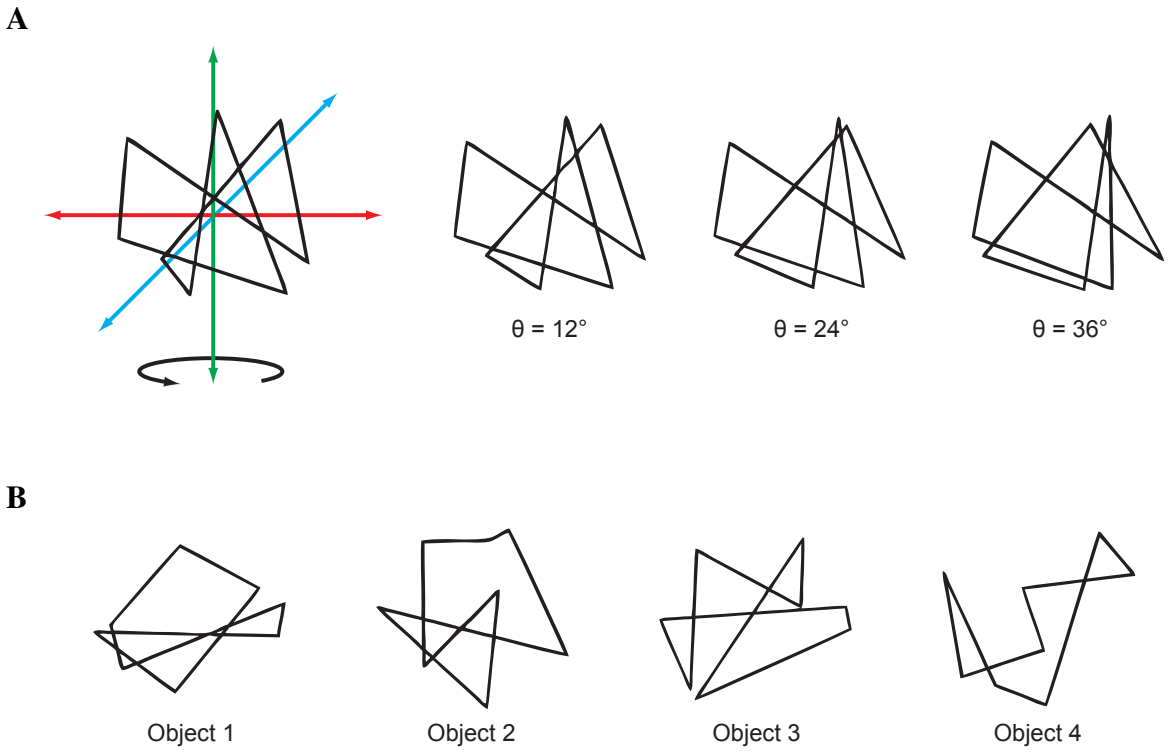

Figure 1.1: Novel "paper clip" objects

A: Objects were composed of eight three-dimensional vertices joined together with line segments. To render the objects to bitmap images, each object was rotated completely about its vertical axis in steps of 12 degrees and reduced to an orthographic projection. **B:** Four of the 16 objects used in the experiment.

method as the original target objects (12 degree steps about the vertical axis). Distractors conformed to the same constraints as the original target objects (no extremely acute angles, approximately rotationally balanced). Participants were instructed to respond "same" if they believed the probe depicted the same object as the entraining sequence or "different" if it depicted a distractor object. Participants received feedback after each trial according to whether their response was correct or incorrect.

During the entraining sequence, object views were presented for 50 ms at either 10 Hz (i.e., temporally predictable) or at a variable rate by manipulating the interstimulus interval (ISI) between subsequent views. Temporally predictable ISIs were 50 ms, totaling 350 ms across the

entraining sequence. Variable ISIs were selected by randomly generating seven ISIs that also summed to 350 ms (Figure 1.2B). ISIs were in the range of 16.67 ms (minimum) to 216.67 ms (maximum) in increments of 16.67 ms. Temporal unpredictability was maximized by generating 400 such ISI sequences, calculating the summed squared error (SSE) across subsequent ISIs in a sequence, and selecting the 100 sequences with the highest SSE for use during the experiment.

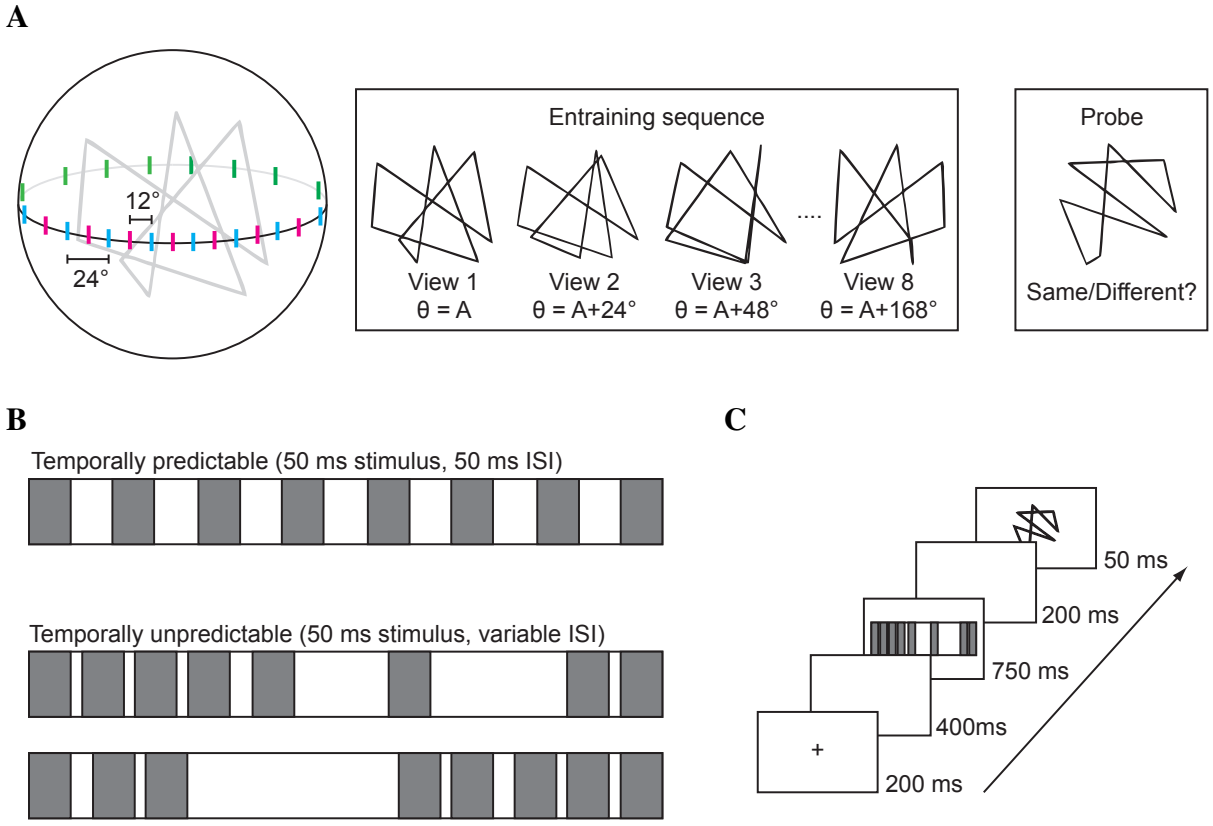

Figure 1.2: Experimental procedure

A: Experimental trials contained an entraining sequence composed of eight views of a single object, followed by a probe stimulus. Entraining views were spaced 24 degrees apart (blue tick marks). The probe depicted an unseen view from the 12 degree interpolations between views of the entraining sequence (magenta tick marks) or from outside the span of the entraining sequence in increments of 24 degrees (green tick marks). **B:** Entraining views were either presented at 10 Hz with a 50 ms on time and 50 ms off time or in a temporally unpredictable manner with a 50 ms on time (gray segments) and variable off time (white segments). In both cases, the duration of the total entraining sequence was held constant at 750 ms. **C:** Order and timing of events within a single trial.

The experiment was displayed on an LCD monitor with a native resolution of 1280x1024 operating at 60 Hz using the Psychophysics Toolbox Version 3 (Brainard, 1997; Pelli, 1997). Stimuli were presented on an isoluminant 50% gray background and subtended approximately 5 degrees of visual angle. Trials began with a fixation cross (200 ms) followed by a blank (400 ms), the entraining sequence (750 ms total), a second blank (200 ms), and ended with the probe stimulus (50 ms) (Figure 1.2C). Participants were required to respond within 2000 ms. Trials were separated by a variable intertrial interval of 2000-2400 ms. The experiment contained 500 trials with an additional 20 practice trials that contained a longer blank (1000 ms) between the entraining sequence and the probe to familiarize participants with the order of events during trials. Participants completed the 20 practice trials (which were discarded from analysis) prior to performing the 500 experimental trials.

1.2.4 EEG recording and preprocessing

The EEG was recorded using an Electrical Geodesics, Inc. (EGI) system composed of a 128 channel net (HCGSN 130) amplified through 200 MΩ amplifiers (Net Amps 200). The signal was sampled at 250 Hz with impedances for each electrode were adjusted to less than 40 kΩ before and during the recording. Stimulus and response trigger onsets were measured via the Psychophysics Toolbox using a high precision realtime clock that was synchronized within 2.5 ms of the EEG system's clock before every trial during the experiment.

EEG data were preprocessed using the FieldTrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011). Raw data were first band-pass filtered between 1 Hz and 100 Hz with a 59-61 Hz band-stop and then epoched into 2350 ms segments that spanned the start of the pre-trial blank to 1000 ms after the probe stimulus. Individual segments were visually inspected and rejected if found to contain muscle artifacts or atypical noise. Bad channels were also identified and temporarily removed from the data before performing ICA decomposition (Delorme & Makeig, 2004) to remove of ocular artifacts. Components related to ocular artifacts were identified based on their topographical distribution across electrodes. The data were reconstructed without the ocular com-

ponents and any bad channels were replaced using spherical spline interpolation (Perrin, Pernier, Bertrand, & Echallier, 1989). The resulting segments were re-referenced to the average reference.

1.2.5 Event-related averaging

Event-related averaging was performed separately for the entraining sequence and the subsequent probe. For the entraining sequence, data were aligned to the onset of entraining views 2 through 8 and averaged from the period beginning 50 ms before each entrainer and ending 50 ms after. Baseline correction was performed using the first 50 ms of this period. For the probe, data were aligned to the probe onset and averaged from the period beginning 200 ms before the probe and ending 400 ms after. This allowed detection of predictability effects during the blank period elicited by the entraining sequence as well as probe-evoked predictability effects.

In both cases, waveforms were further averaged over pools of seven electrodes centered over locations from the 10-10 system that are commonly associated with perceptual processing (e.g., Doherty, Rao, Mesulam, & Nobre, 2005; Rohenkohl & Nobre, 2011). These pools included one midline site (OZ) and three sites over each hemisphere (O1/O2, PO3/PO4, and PO7/PO8) (Figure 1.3).

1.2.6 Time-frequency analysis

$$ITC(f,t) = \left| \frac{1}{N} \sum_{n=1}^{N} \frac{F_n(f,t)}{|F_n(f,t)|} \right|$$

1.3 Results

1.3.1 Behavioral measures of spatial and temporal predictability

Five subjects were excluded from behavioral analyses for accuracy 2.7σ (or further) below mean accuracy across subjects. The remaining 53 subjects were submitted to a 2x2 ANOVA with spatial and temporal predictability as within-subjects factors. Experiment type (EEG or behavioral

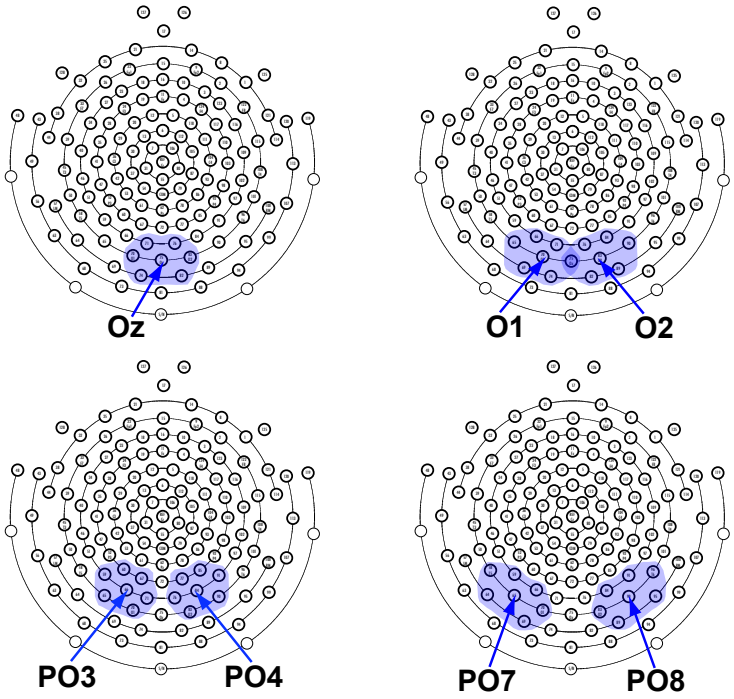

Figure 1.3: Electrode poolings

Pools of seven electrodes centered over locations from the 10-10 system used for all event-related averaging. These pools included one midline site (OZ) and three sites over each hemisphere (O1/O2, PO3/PO4, and PO7/PO8).

only) was included as an additional between-subjects factor to ensure that it did not interact with any of the within-subjects factors. Accuracy and reaction times were collected during the experiment and were used to compute *d'*, a measure of sensitivity that takes into account response bias, and inverse efficiency, a measure that combines accuracy and reaction times (Townshend & Ashby, 2005). These behavioral measures are plotted in Figure 1.4.

Subjects that completed the full EEG experiment were on average less accurate ($F(1, 51) = 4.80$, $p = 0.033$) but responded more quickly ($F(1, 51) = 10.05$, $p = 0.003$) than subjects that completed the solely behavioral experiment. These differences reflect a speed-accuracy tradeoff, likely due to differences in instructions given to subjects by experimenters or motivational differences between subject groups. Importantly, experiment type did not interact with any within-subjects factors (all p's > 0.05) indicating that the behavioral measures of interest were not dependent on

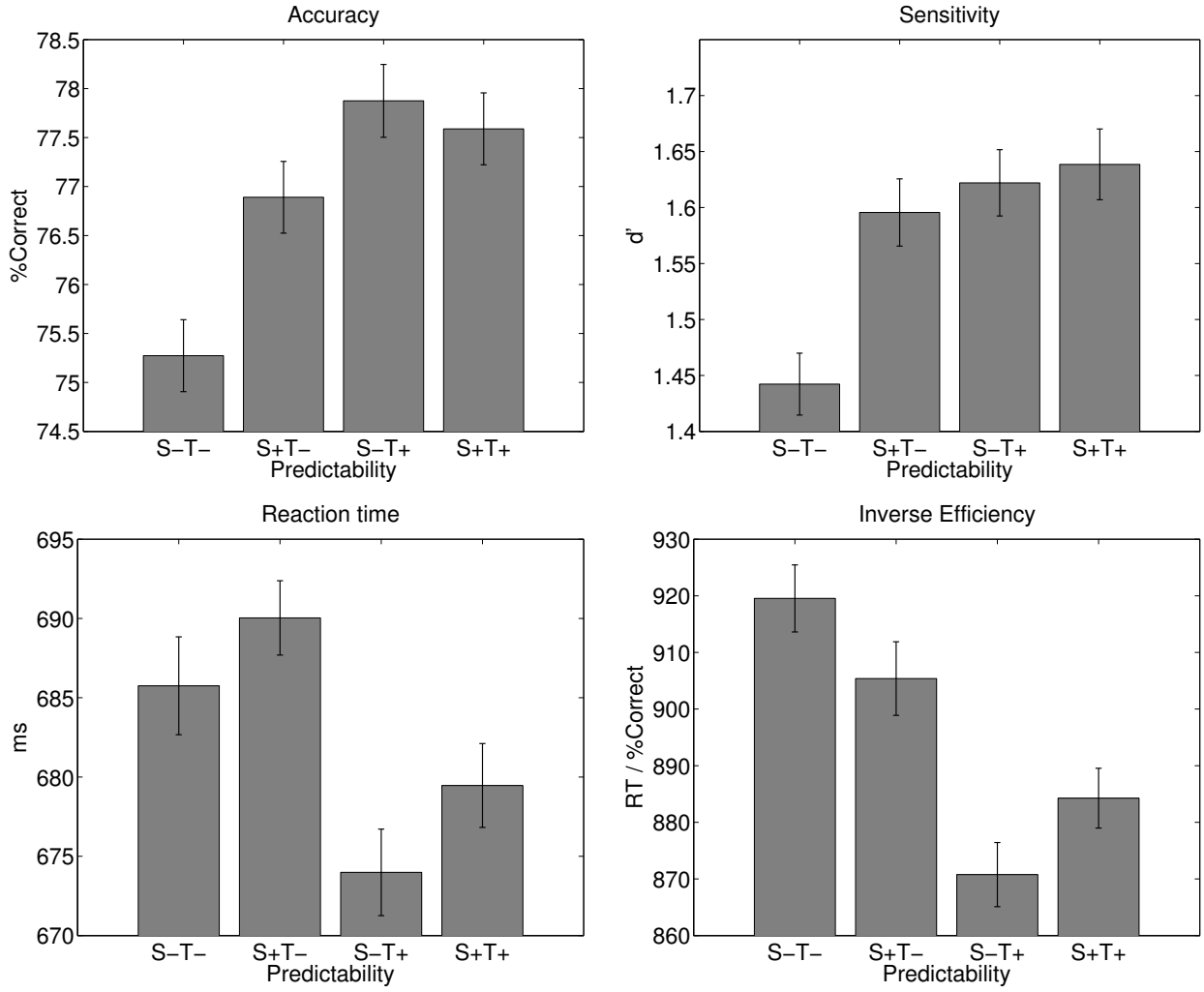

Figure 1.4: Behavioral measures of spatial and temporal predictability

Accuracy, *d'* (sensitivity), reaction time, and inverse efficiency (reaction time divided by percent correct) as a function of entrainment condition. S-/+ refers to spatially unpredictable and predictable, T-/+ to temporally unpredictable and predictable. Error bars depict within-subjects error using the method described in Cousineau (2005) adapted for standard error.

which type of experiment subjects completed.

Overall, subjects were more accurate when the entraining sequence was temporally predictable ($F(1, 51) = 17.84$, $p < 0.001$). A similar effect for spatial predictability failed to reach significance ($F(1, 51) = 1.85$, $p = 0.18$). The interaction between spatial and temporal predictability, however, was significant ($F(1, 51) = 6.13$, $p = 0.017$). LeabraTI as well as previous investigations

of predictability (e.g., Doherty et al., 2005) suggest that spatial and temporal predictability should have an additive effect on behavioral outcomes. However, the combined spatial and temporal predictability condition here (denoted S+T+ in Figure 1.4) does not exhibit this pattern. Although not significantly different from spatial predictability alone ($t(52) = 1.29$, $p = 0.204$) or from temporal predictability alone ($t(52) = 0.45$, $p = 0.652$), this result merits further investigation.

When responses are transformed into *d'*, there is a significant effect of both spatial ($F(1, 51) = 4.71$, $p = 0.035$) and temporal predictability ($F(1, 51) = 11.99$, $p < 0.001$). This result suggests that response bias can partially explain why spatial predictability failed to reach significance for raw accuracy. Other potential reasons for this result are discussed later. The interaction between spatial and temporal predictability remained significant for *d'* ($F(1, 51) = 4.49$, $p = 0.039$). The interaction is additive, but is driven primarily by the strong effect of introducing spatial or temporal predictability over complete unpredictability (S-T- versus S+T-, $t(52) = 3.19$, $p = 0.002$; S-T- versus S+T+, $t(52) = 4.26$, $p < 0.001$) opposed to any synergistic effect of combined spatial and temporal predictability (S+T+ versus S+T-, $t(52) = 0.90$; S+T+ versus S-T+, $t(52) = 0.31$; both p's > 0.05).

Reaction times were significantly faster when the entraining sequence was temporally predictable ($F(1, 51) = 12.38$, $p < 0.001$). A similar effect for spatial predictability failed to reach significance ($F(1, 51) = 1.96$, $p = 0.168$) nor did the interaction term ($F(1, 51) = 0.05$, $p = 0.83$).

Inverse efficiency, which considers reaction time as a function of accuracy (defined as reaction time divided by percent correct) can be thought of as the amount of energy consumed by the system to produce a behavioral outcome (Townshend & Ashby, 1983). It is often used to remove non-monotonicities present in accuracy or reaction times alone, although that effect is not observed here. Nevertheless, it provides another lens under which to interpret the results, and thus it is considered here. Inverse efficiency was significantly lower when the entraining sequence was temporally predictable ($F(1, 51) = 23.31$, $p < 0.001$), but not when it was spatially predictability ($F(1, 51) = 0.002$, $p = 0.963$). Inverse efficiency is characterized by a significant cross-over interaction ($F(1, 51) = 5.85$, $p = 0.019$). Spatial predictability of the entraining sequence produces lowers inverse efficiency over complete unpredictability. Inverse efficiency is lowest on average when

stimuli are temporally predictable, but the addition of spatial predictability causes an increase in inverse efficiency. These results suggest... **TODO**

1.3.2 Time course of spatial and temporal predictability

1.3.3 Entrainment of alpha oscillations

1.3.4 Alpha phase at probe onset

1.4 Discussion

References

Balas, B. J., & Sinha, P. (2009). The role of sequence order in determining view canonicality for novel wire-frame objects. Attention, Perception & Psychophysics, 71(4), 712–723.

Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. The Annals of Statistics, 29(4), 1165–1188.

Brainard, D. (1997). The Psychophysics Toolbox. Spatial Vision, 10(4), 433–436.

Bulthoff, H. H., & Edelman, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. Proceedings of the National Academy of Sciences of the United States of America, 89(1), 60–64.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Massons method. Tutorials in Quantitative Methods for Psychology, 1(1), 42–45.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods, 134(1), 9–21.

Doherty, J. R., Rao, A., Mesulam, M. M., & Nobre, A. C. (2005). Synergistic effect of combined temporal and spatial expectations on visual attention. The Journal of Neuroscience, 25(36), 8259–8266.

Edelman, S., & Bulthoff, H. H. (1992). Orientation dependence in the recognition of familiar and novel views of three-dimensional objects. Vision Research, 32(12), 2385–2400.

Logothetis, N., Pauls, J., Bulthoff, H., & Poggio, T. (1994). View-dependent object recognition by monkeys. Current Biology, 4(5), 401–414.

Logothetis, N. K., Pauls, J., & Poggio, T. (1995). Shape representation in the inferior temporal cortex of monkeys. Current Biology, 5(5), 552–563.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Computational Intelligence and Neuroscience, 2011.

Pelli, D. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spatial Vision, 10(4), 437–442.

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. Electroencephalography and Clinical Neurophysiology, 72(2), 184–187.

Rohenkohl, G., & Nobre, A. C. (2011). Alpha oscillations related to anticipatory attention follow temporal expectations. The Journal of Neuroscience, 31(40), 14076–14084.

Sinha, P., & Poggio, T. (1996). Role of learning in three-dimensional form perception. Nature, 384(6608), 460–463.

Townshend, J. T., & Ashby, F. G. (1983). Stochastic Modeling of Elementary Psychological Processes. Cambridge: Cambridge University Press.

Townshend, J. T., & Ashby, F. G. (2005). Methods of modeling capacity in simple processing systems. In J. N. Castellan Jr., & F. Restle (Eds.), Cognitive Theory: Volume 3 (pp. 200–239). Hillsdale, NJ: Lawrence Erlbaum Associates.