Chapter 1

Spatial and temporal prediction during novel object recognition: Temporal and spectral signatures

1.1 Introduction

How does the brain integrate information from one moment to the next and use it it to drive predictions about what will happen? A number of models of general neocortical function have highlighted the centrality of prediction in neural processing (e.g., DayanHintonNealEtAl95RaoBallard99LeeMumford03Friston05GeorgeHawkins09; BastosUsreyAdamsEtAl12), but it is surprisingly but often overlooked in psychology and neuroscience investigations of sensory processing. Predictability might be useful in perceptual processes like object recognition, for example, by allowing better integration across features extracted over the course of several samples (Foldiak91StringerPerryRollsEtAl06WallisBaddeley97IsikLeiboPoggio12). The timing of predictions might also be important. The LeabraTI model (Chapter **??**) as well as several other theories of sensory prediction (ArnalGiraud12GiraudPoeppel12) emphasize that predictions occur in a pacemaker manner at regular intervals and thus temporal properties of endogenous processes as well as exogenous stimulation might also affect perceptual processing.

Previous work has implicated 7-13 Hz alpha oscillations in the allocation of spatial attention and anticipation of the temporal onset of stimuli (GouldRushworthNobre11; BelyusarSnyderFreyEtAl13; RohenkohlNobre11). Specifically, alpha oscillations desynchronize in the hemisphere contralateral to the attended region of visual space as well as in anticipation of the onset of a time-varying stimulus. Pre-stimulus alpha oscillations have also been related to the detection rate of at-threshold point-light targets (MathewsonGrattonFabianiEtAl09BuschDuboisVanRullen09), suggesting that spontaneous fluctuations in alpha oscillatory properties could also be related to the ability to properly anticipate a stimulus. Together, these results strongly implicate the role of alpha oscillations in hemifield-based spatial attention tasks and for prediction of relatively simple stimuli.

One issue with relating the extant literature on alpha oscillations to prediction is that is unclear whether the experiments described therein engaged actual predictive processing about what would happen or comparatively simple anticipatory attention mechanisms about *where* a stimulus might appear. For example, the alpha rhythm might simply correspond to shifts of spatial attention, which can be oriented approximately 10 times per second (VanRullenDubois11), and studies using a Posner spatial cueing task (Posner80) support this idea (CapotostoBabiloniRomaniEtAl09BuschVanRullen10). Other studies, however, suggest that prediction and attention are separable mechanisms with dissociable effects (KokRahnevJeheeEtAl12WyartNobreSummerfield12HorschigJensenVanSchouwenburgEtAl13). Experiments have generally not examined the role of alpha oscillations in complex perceptual tasks that require actual predictive processing, although it was recently demonstrated that an alpha-band presentation rate (12.5 Hz) is optimal for maximal fMRI BOLD response in a dynamic face recognition experiment (SchultzBrockhausBulthoffEtAl13).

The work described in this chapter investigated the role of predictive processing during a novel object recognition task. The experiment made use of novel three-dimensional stimuli that required integration over multiple sequential views to extract their three-dimensional structure. The stimuli were presented at central fixation and thus, manipulating the ordering of the

views could be used to test the effect of their predictability without confounding it with spatial attention. This is henceforth referred to as "spatially" predictable, although it is likely a combination of spatially- and featurally- predictive processing (i.e., the prediction of specific features at specific locations of a spatial map). To test the relationship between alpha oscillations in predictive processing, the experiment took advantage of the entrainability of endogenous alpha oscillations by exogenous rhythmic stimulation (SchroederLakatosKajikawaEtAl08CalderoneLakatosButlerEtAlInPress). Whereas previous experiments related pre-stimulus alpha to anticipation via a post-hoc grouping of detected and missed stimuli (MathewsonGrattonFabianiEtAl09BuschDuboisVanrullen09), the current experiment entrained alpha oscillations by presenting subsequent views of the stimulus at a regular 10 Hz rate to determine the causal effect of alpha oscillations on prediction (henceforth referred to as "temporally" predictable).

The results of the experiment indicated that spatial and temporal predictability of an entraining sequence enhanced discriminability of a subsequently presented probe stimulus. Temporal predictability also speeded response times for the probe judgement. EEG amplitude averaging indicated separable time courses for spatial and temporal prediction with temporal predictability always preceding the onset of stimuli and spatial predictability manifesting at the onset of stimuli and persisting for over 100 ms after in the case of the probe. Oscillatory analyses indicated strong bilateral alpha power and phase coherence modulation as a function of stimulus predictability, as well as similar and sometimes more prominent, effects in the lower frequency delta-theta (5 Hz) band.

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 investigations of three-dimensional object recognition (BulthoffEdelman92EdelmanBulthoff92LogothetisPaulsBulthoffEtAl94LogothetisPaulsPoggio95SinhaPoggio96) 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 (BalasSinha09b). 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 consist-
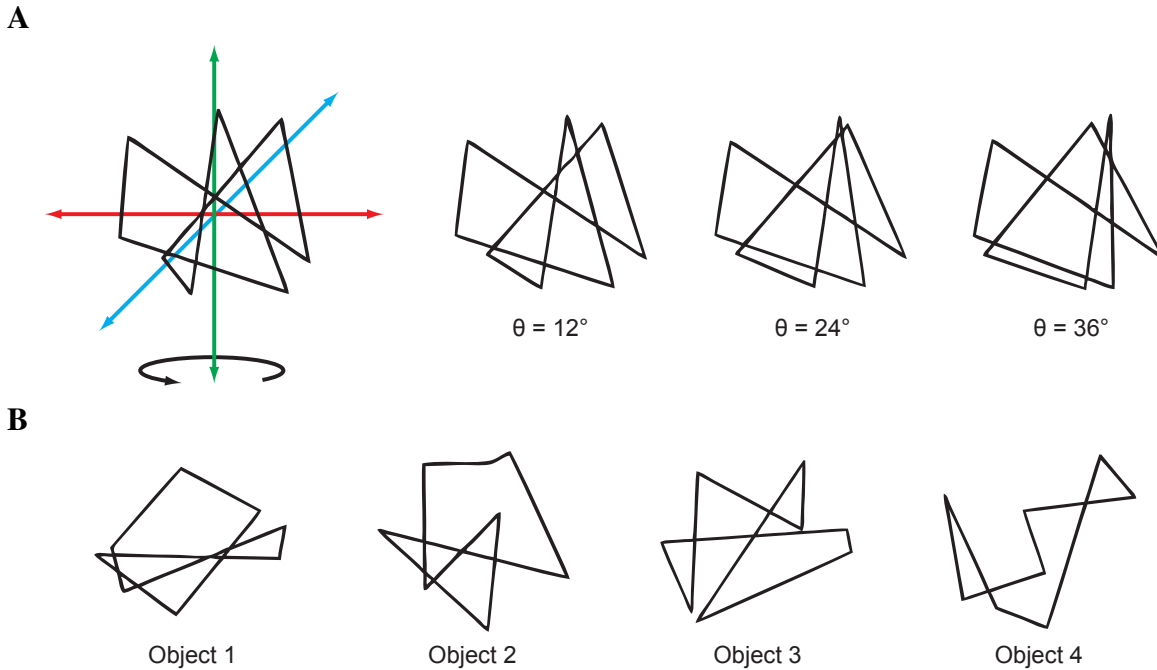
A

$\theta = 12°$ $\theta = 24°$ $\theta = 36°$

B

Object 1   Object 2   Object 3   Object 4

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.

ing 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 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.

The experiment was displayed on an LCD monitor at native resolution operating at 60 Hz using the Psychophysics Toolbox Version 3 (Brainard97Pelli97). Stimuli were presented at central fixation 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
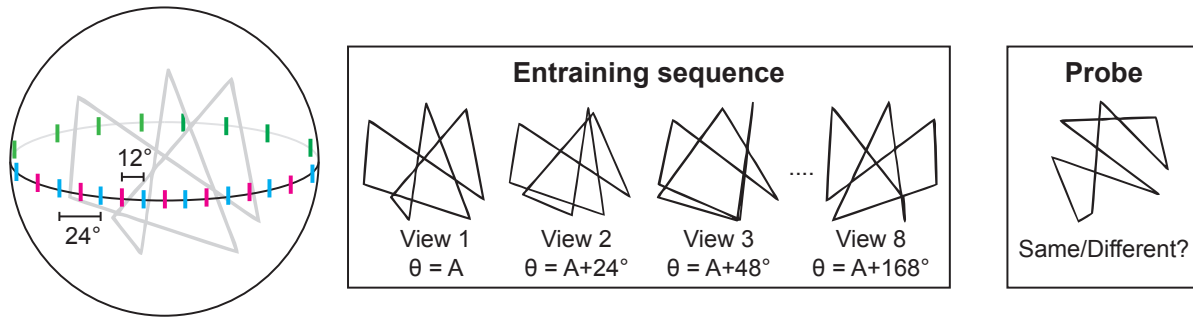
Figure 1.2: Experimental procedure

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.

system's clock before every trial during the experiment.

EEG data were preprocessed using the FieldTrip toolbox (OostenveldFriesMarisEtAl11). 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 (DelormeMakeig04) 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 components and any bad channels were replaced using spherical spline interpolation (PerrinPernierBertrandEtAl89). 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 due to differences in phase elicited by the entraining sequence as well as probe-evoked predictability effects.

All waveforms were averaged over a montage of 23 electrodes that covered the occipital and parietal cortices (Figure 1.3). The montage included locations from the 10-10 system that are commonly associated with perceptual processing (Oz, O1/O2, PO3/PO4, and PO7/PO8) (e.g., DohertyRaoMesulamEtAl05RohenkohlNobre11FahrenfortScholteLamme07).

1.2.6   Time-frequency analysis

Segmented data were used to compute time-frequency data for each trial. Data were first downsampled to 125 Hz and then used to compute the instantaneous Fourier coefficients at each time bin using a multi-taper approach. Hanning tapers were generated at 5-40 Hz and convolved with the data using a sliding time window (four cycles per frequency per time window). The relatively long time window required for low-frequency bands prevents computation of time-frequency
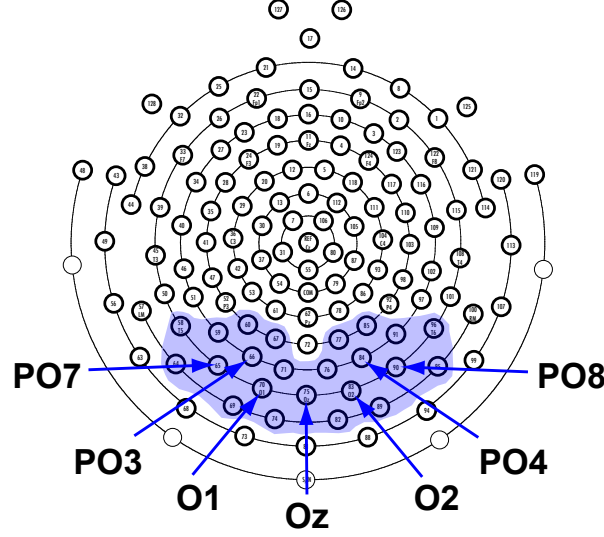
Figure 1.3: Electrode pooling for analyses

Blue shaded region denotes pooled electrodes. Locations from the 10-10 system are indicated.

data for short time segments, such as the 200 ms blank before the probe. To address this issue, time-frequency data was computed over the entire 2350 ms trial epoch and then cropped to investigate temporal regions of interest.

Power was computed from the magnitude of the instantaneous Fourier coefficients at each frequency (f) and time bin (t):w

$$Power(f,t) = \frac{1}{n} \sum_{k=1}^{n} |F_k(f,t)|^2$$

Phase, which is normally defined as the angle of the complex Fourier coefficients, cannot be averaged due to its circularity and thus standard statistical models cannot be applied to assess its significance. One solution to this problem is to compute inter-trial coherence (ITC) instead (LachauxRodriguezMartinerieEtAl99). ITC is averaged in the complex domain by first normalizing phase information to unit length by dividing off power and computing the magnitude:

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

ITC ranges between 0 and 1 and represents how systematic phase angles are across trials. A

value of 0 indicates that phase information is essentially uniformly distributed across trials while a value of 1 indicates a high degree of phase-locking at a particular frequency across trials.

All time-frequency analyses were averaged over the same montage of 23 occipitoparietal electrodes that was used to compute event-related averages (Figure 1.3).

1.3 Results

1.3.1 Behavioral measures of spatial and temporal predictability

Five subjects were excluded from behavioral analysis 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 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 (TownshendAshby78). 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 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 ($F(1, 51) = 6.13$, $p = 0.017$). The LeabraTI model (Chap-
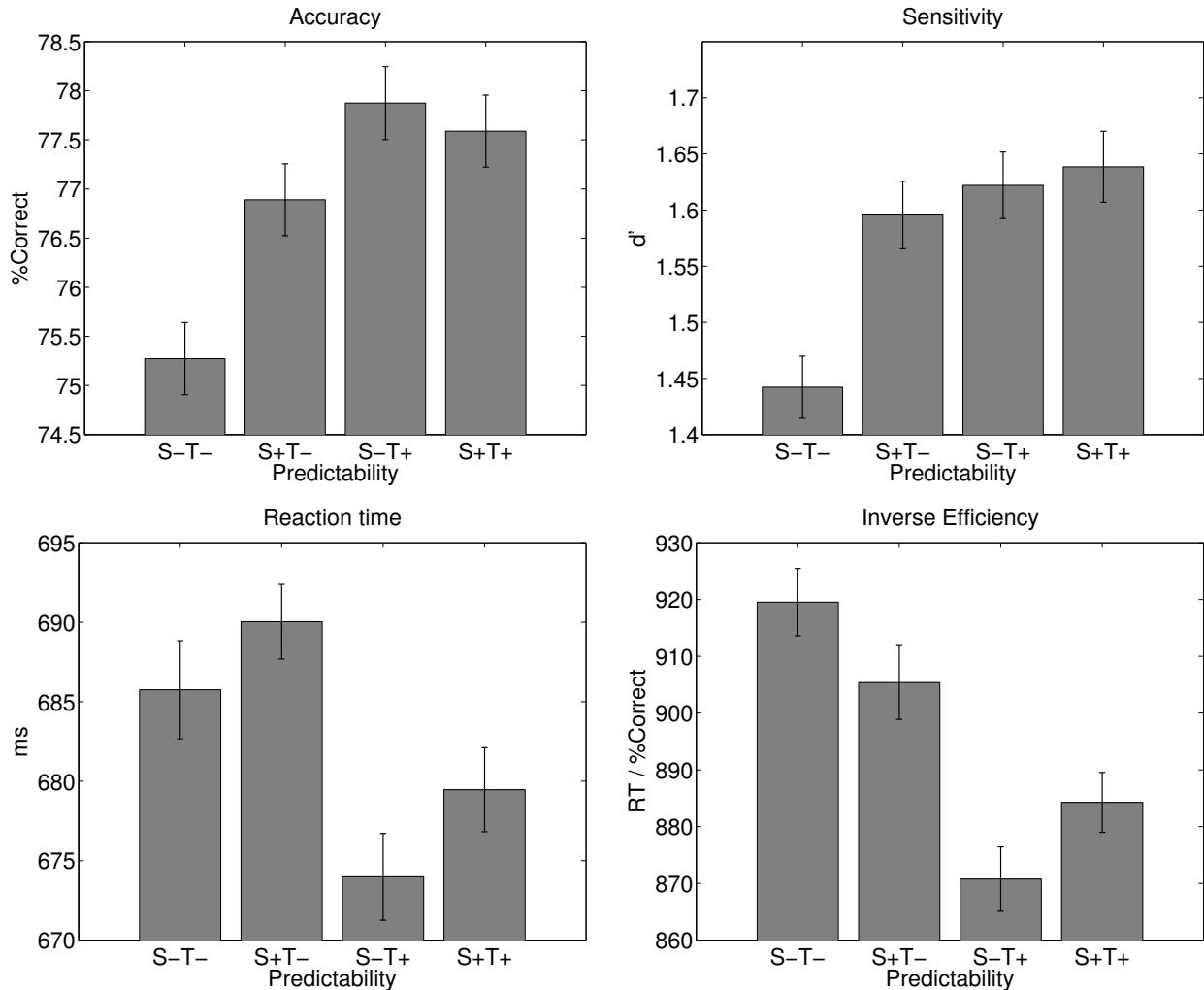
Figure 1.4: Behavioral measures of spatial and temporal predictability

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 Cousineau05) adapted for standard error.

ter **??**) as well as investigations predictability on attentional allocation (e.g., DohertyRaoMesulamEtAl05RohenkohlGouldPessoaEtAl14) suggest that spatial and temporal predictability should have a superadditive effect on behavioral outcomes. However, the combined spatial and temporal predictability condition here (denoted S+T+ in Figure 1.4) was subadditive. 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 at least partially explain why spatial predictability failed to reach significance for raw accuracy. 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 (TownshendAshby83). 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.

1.3.2   Time course of spatial and temporal predictability

A total of five subjects were excluded from EEG analysis – three for an overabundance of artifacts in the EEG recording resulting in low trial counts after rejection and two for accuracy 2.7σ (or further) below mean accuracy across subjects (these two subjects were also excluded from behavioral analyses, see preceding section). The remaining 24 subjects were included in all EEG analyses.

A 2x2 ANOVA with spatial and temporal predictability as within-subjects factors was used to assess statistical significance at each time bin of event-related averages. *p*-values were corrected for a maximum false discovery rate (FDR) of 5% using the method described in BenjaminiYeku-tieli01). Additionally, effects were only considered significant if they persisted for at least 16 ms.

To investigate the build-up of spatial and temporal predictability over the entraining sequence, activity from the second through final entraining views was averaged for each condition (the first entraining view is unpredictable, so it is omitted from the average). The results of this analysis are plotted in Figure 1.5. The first thing worth noting is that a large 10 Hz periodicity is present for the temporally predictable conditions (S-T+ and S+T+), phase-aligned approximately to the onset of each entrainer. Temporally unpredictable entrainers (S-T- and S+T-) are also approximately periodic. The reason for the 10 Hz periodicity in these conditions despite being temporally unpredictable is likely due to the 750 ms constant duration of the entraining sequence regardless of condition (Figure 1.2B). Presenting eight stimuli in 750 ms with a variable ISI is a 10 Hz presentation rate on average. Still, the temporally unpredictable entrainers exhibit markedly weaker amplitude and are approximately 180 degrees out of phase with the the temporally predictable entrainers.

The effect of spatial predictability manifested 26 ms after the onset of the entrainer and persisted for at least another 24 ms (one quarter of the 10 Hz period). Temporal predictability, in contrast, manifested prior to (-38 through -22 ms pre-stimulus) and at the onset of the entrainer (-6 ms pre-stimulus through 18 ms post-stimulus). The effect of temporal predictability appears to be
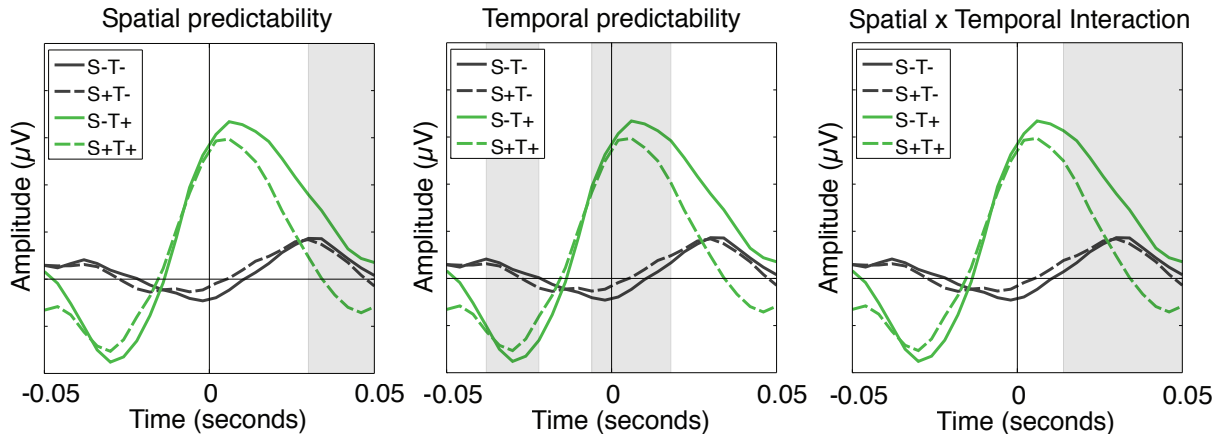
Figure 1.5: Entrainer-evoked activity

Grand averages for entrainers 2 through 8 as a function of entrainment condition. S-/+ refers to spatially unpredictable and predictable, T-/+ to temporally unpredictable and predictable. All plots depict the grand average with gray shaded regions denoting significant effects of spatial predictability (left), temporal predictability (center), and the interaction between these terms controlling for a maximum false discovery rate (FDR) of 5%.

driven by the antiphase relationship between T- and T+ conditions at these time points. Together, these effects demonstrate differential time courses for spatial and temporal predictability.

Spatial predictability was enhanced when stimuli were temporally predictable. This effect is characterized by the significant interaction between spatial and temporal predictability starting 14 ms after the onset of the entrainer and persisting for at least another 36 ms. This result indicates that the brain is more capable of differentiating between spatially coherent and random sequences of stimuli when it can properly anticipate the presentation of each stimulus (S-T+ versus S+T+) compared to when the onset is unpredictable.

To investigate the effect of spatial and temporal predictability on perception of the probe, waveforms were aligned to the probe onset onset averaged from 200 ms before through 400 ms after (Figure 1.6). The results of this analysis essentially mirror the entrainer-evoked effects albeit with a few key differences. Temporal predictability was again a purely anticipatory process, manifesting within the pre-stimulus period (-184 ms through -160 ms and -132 ms through -108 ms pre-stimulus). Spatial predictability also manifested briefly within the pre-stimulus period (-128

ms through -80 ms pre-stimulus), which was not seen for spatially predictable entrainers. The post-stimulus effects of spatial predictability occurred much earlier than for entrainers and persisted much longer, beginning approximately at the onset of the probe (-12 ms pre-stimulus) and lasting 136 ms through the P1 response with several transient effects after.
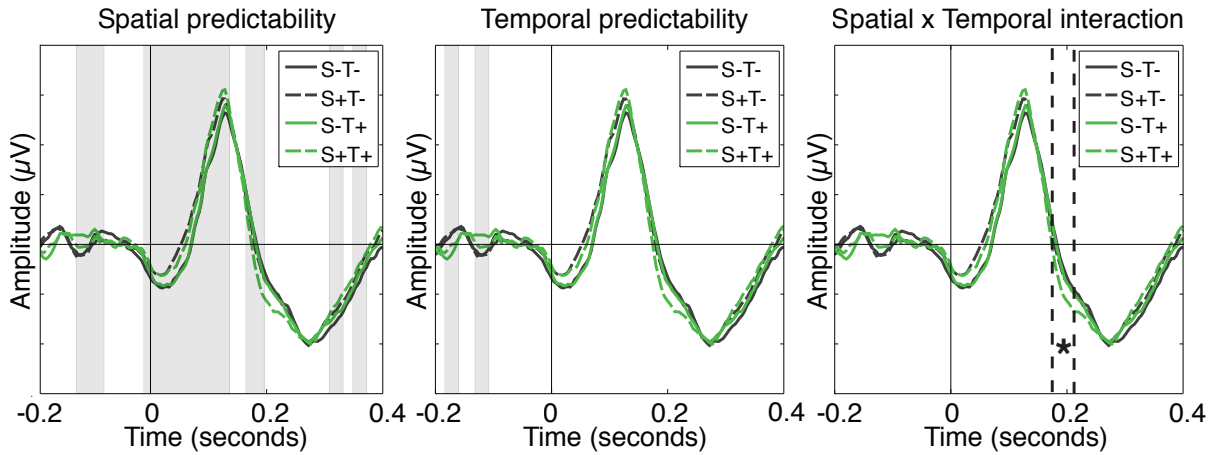

Figure 1.6: Probe-evoked activity

Grand averages for the probe stimulus, including the preceding 200 ms blank. Asterisk in the interaction plot indicates trending significance at the 5% level when averaging amplitude within the window defined by the dotted lines.

Unlike the significant interaction between spatial and temporal predictability for entrainers, the interaction for the probe failed to reach significance given the constraints of the statistical test (maximum FDR of 5%, 16 ms consecutive significance). However, averaging the amplitude within a window defined by exploratory analysis indicated a trending interaction from 170 ms to 210 ms ($F(1, 22) = 3.97$, $p = 0.059$). This effect was more pronounced and reached significance if only the right hemisphere channels were considered ($F(1, 22) = 5.61$, $p = 0.027$), but failed to reach significance in the left hemisphere ($F(1, 22) = 2.24$, $p = 0.148$), explaining the trending effect when both hemispheres are considered jointly.

1.3.3   Predictability entrains alpha oscillations

The same 24 subjects used in event-related analyses (see preceding section) were used in time-frequency analyses described here. Statistical methods were also identical, consisting of a 2x2 ANOVA with spatial and temporal predictability as within-subjects factors and *p*-values were corrected for a 5% maximum FDR (BenjaminiYekutieli01). Effects were only considered significant if they persisted for at least 16 ms.

Power and inter-trial coherence (ITC), a measure of phase angle consistency across trials (LachauxRodriguezMartinerieEtAl99) were computed over a 5-20 Hz frequency range to investigate the relationship between entrainer predictability and alpha oscillatory properties (Figures 1.7-1.8). Both spatial and temporal predictability had a significant effect on 10 Hz power and ITC beginning around 200 ms after the onset of the first entrainer (power: 160-168 ms after entrainer 1; ITC: 136-208 ms after entrainer 1). Together, these results indicate that 10 Hz entrainment affects both power and phase alignment and takes around 2-3 events to establish. Spatial predictability had a suppressive effect on 10 Hz power and phase alignment whereas temporal predictability had a positive effect. Any interactions between spatial and temporal predictability failed to reach significant levels altogether.

The effect of spatial predictability only persisted for three entrainers on average (duration ranges for 10 Hz power and ITC: 272-280 ms). Temporal predictability, in contrast, exhibited a sustained effect on 10 Hz power lasting nearly the entire entraining sequence. In the case of ITC, the effect of temporal predictability persisted throughout the blank period that separated the entraining sequence and the probe (Figure 1.9). This result indicates that alpha phase remained more aligned for temporally predictable stimuli even without exogenous entrainment.

Predictability effects on 10 Hz ITC failed to reach significance during the presentation of the probe or the following 400 ms period given the constraints of the statistical test (maximum FDR of 5%, 16 ms consecutive significance). An enhancement in ITC could be observed for the combined predictability (S+T+) condition and trended toward significance when averaged over a
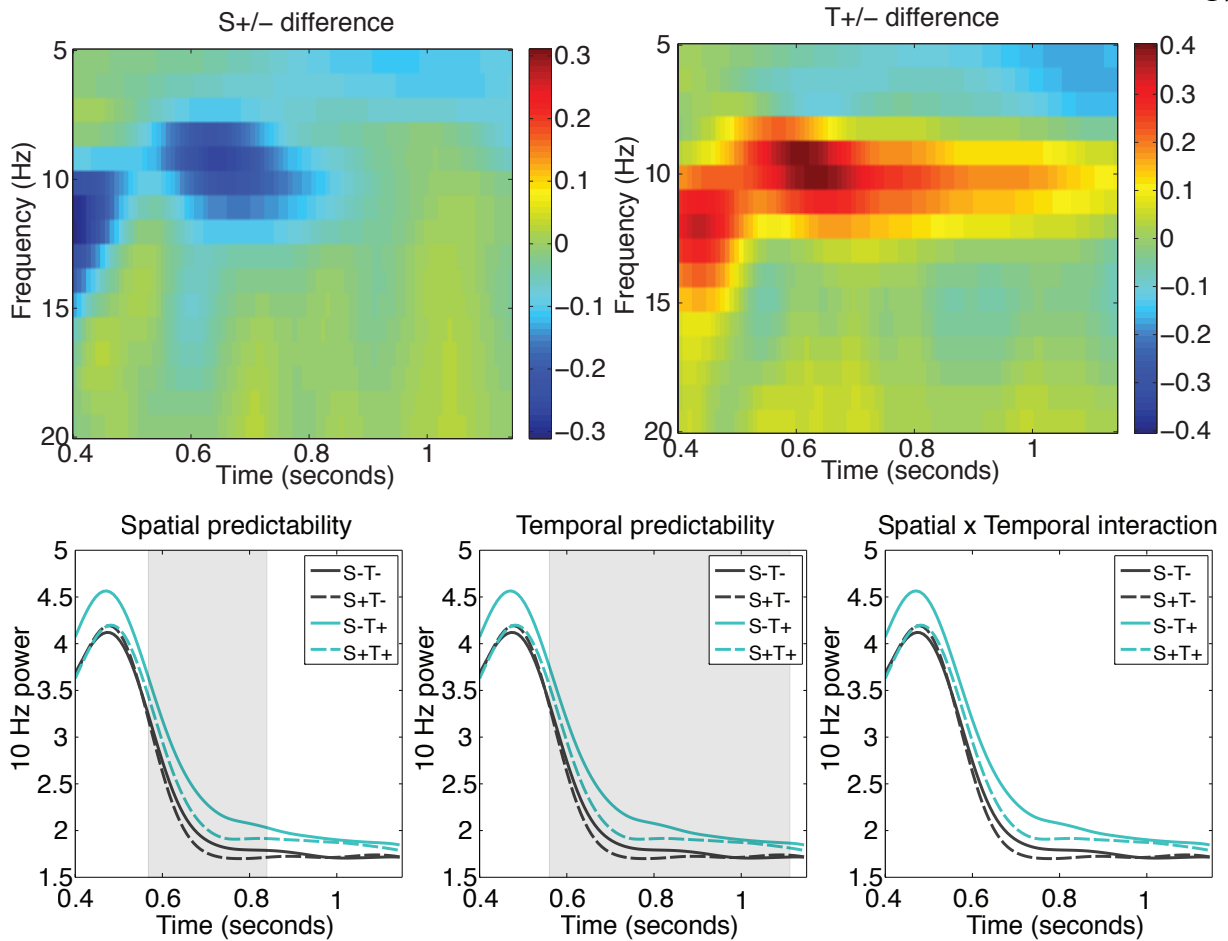
Figure 1.7: Effect of entrainer predictability on alpha power

Alpha-band power over the entraining sequence. **Top:** Main effects of spatial and temporal predictability on oscillatory power in the 5-20 Hz frequency range. **Bottom:** 10 Hz only effects of spatial predictability (left), temporal predictability (center), and the interaction between these terms with gray shaded ranges indicating significance while controlling for a maximum false discovery rate (FDR) of 5%. S-/+ refers to spatially unpredictable and predictable, T-/+ to temporally unpredictable and predictable. Time axes indicate total trial time after the initial fixation cross with 0.4 seconds corresponding to the first entrainer.

100 ms period beginning 155 ms after the onset of the probe ($F(1, 22) = 4.11$, $p = 0.055$). This effect was more pronounced and reached significance if only the right hemisphere channels were considered ($F(1, 22) = 7.45$, $p = 0.012$), but failed to reach significance in the left hemisphere ($F(1, 22) = 1.47$, $p = 0.237$), explaining the trending effect when both hemispheres are considered jointly.
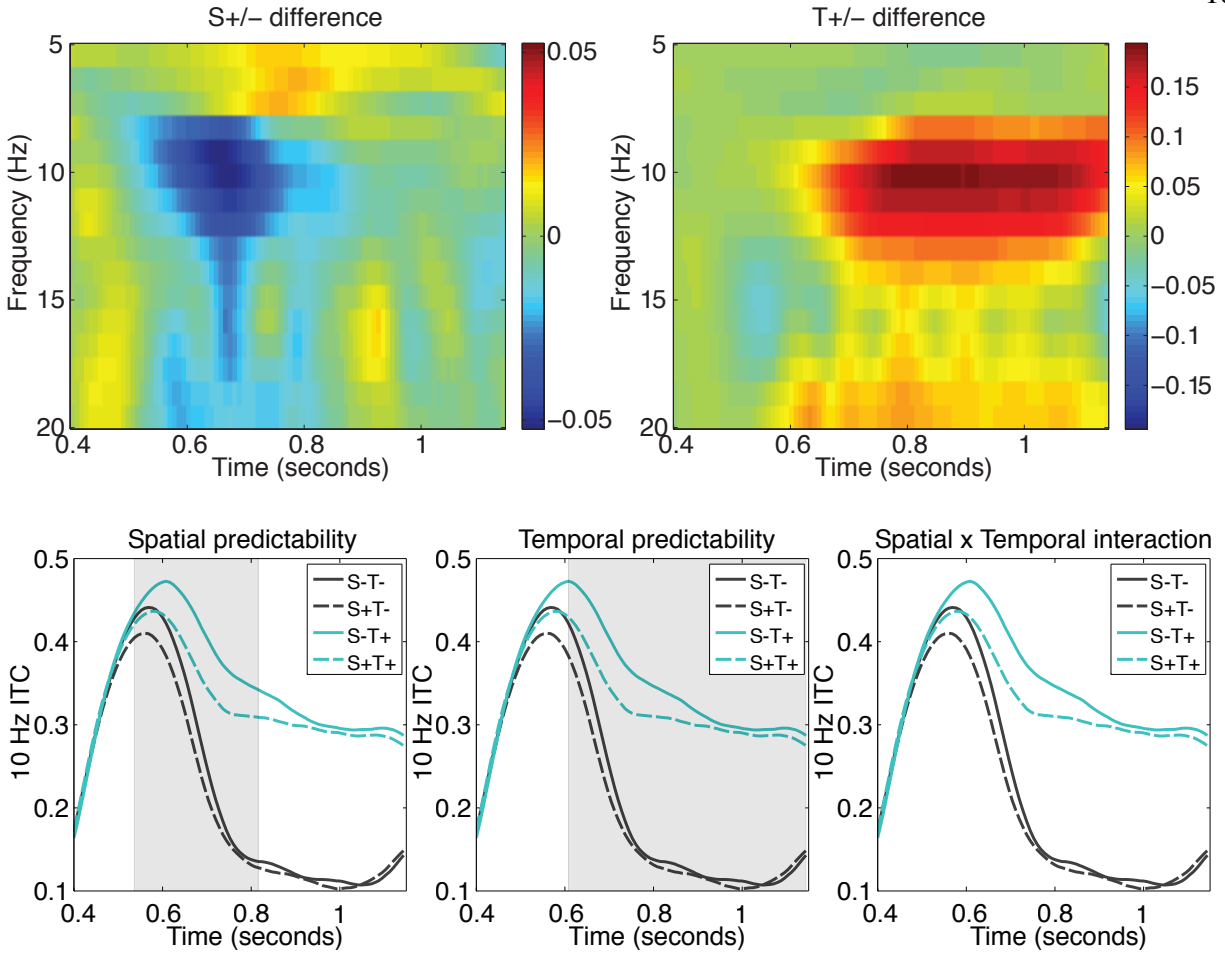
Figure 1.8: Effect of entrainer predictability on alpha phase coherence

Alpha-band inter-trial coherence (ITC) over the entraining sequence. Axes, legends, and shading for significant regions are the same as those described in Figure 1.7.

1.3.4   Predictability effects in delta-theta bands

The effects of spatial and temporal predictability on oscillatory properties during the probe period (-200 ms pre-stimulus through 400 ms after) were investigated in the delta-theta bands, centered around 5 Hz. This frequency was identified based on exploratory analysis and was also motivated by alternative models of sensory prediction (e.g., ArnalGiraud12GiraudPoeppel12). Power and ITC at 5 Hz are plotted in Figures 1.10-1.11.

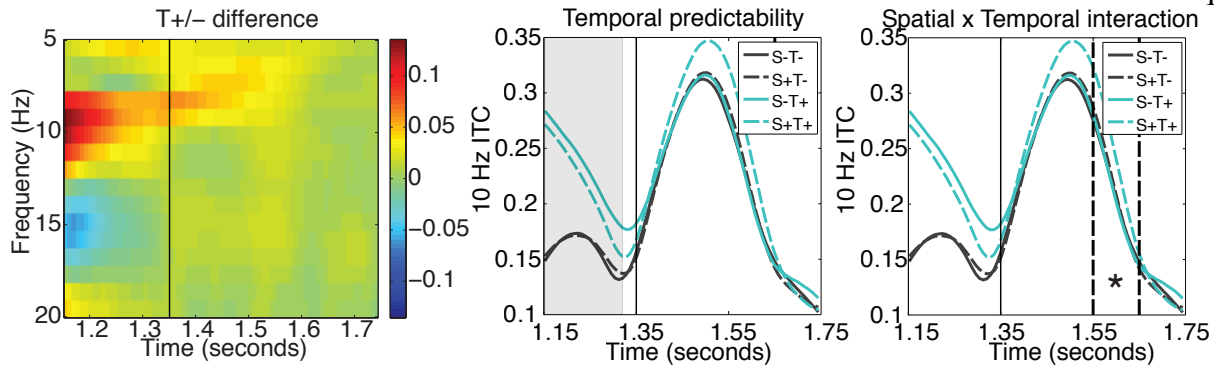Both spatial and temporal predictability had a significant effect on 5 Hz power during the

Figure 1.9: Alpha phase coherence before and after probe

Alpha-band inter-trial coherence (ITC) 200 ms preceding the probe and 400 ms following. Solid vertical line indicates probe onset. Asterisk indicates trending significance at the 5% level when averaging 10 Hz ITC within the window defined by the dotted lines. Time axes indicate total trial time after the initial fixation cross.
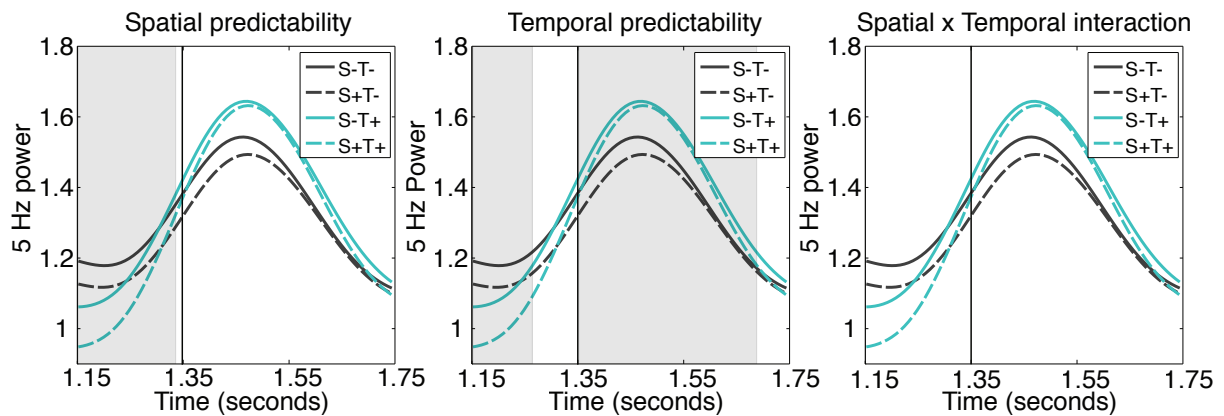

Figure 1.10: Delta-theta power before and after probe

Delta-theta power 200 ms preceding the probe and 400 ms following. Axes, legends, and shading for significant regions are the same as those described in Figure 1.9.

200 ms blank period preceding the probe. Temporal predictability had a suppressive effect on 5 Hz power, in contrast to the positive modulation found for 10 Hz power. This effect reversed following the presentation of the probe and persisted for over 300 ms. The interaction between spatial and temporal predictability failed to reach significance for 5 Hz power at any time points.

Temporal predictability also had a significant effect on 5 Hz ITC beginning during the 200
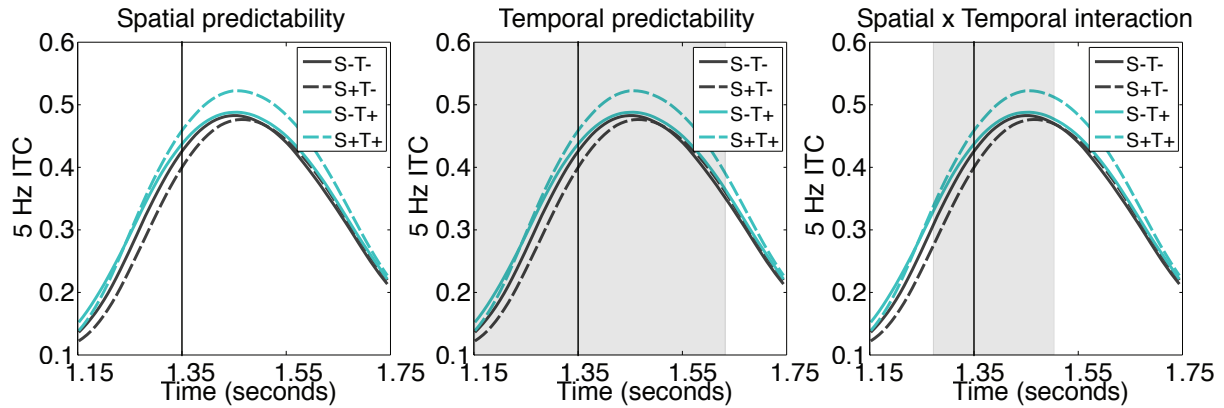
Figure 1.11: Delta-theta phase coherence before and after probe

Delta-theta inter-trial coherence (ITC) 200 ms preceding the probe and 400 ms following. Axes, legends, and shading for significant regions are the same as those described in Figure 1.9.

ms blank period preceding the probe and lasting nearly 300 ms after the presentation of the probe. Whereas 10 Hz ITC decreased during the blank period (yet remained significantly higher for temporally predictable stimuli), 5 Hz ITC increased, and continued to increase until approximately 100 ms after the onset of the probe. 5 Hz ITC was highest for the combined spatial and temporal predictability condition (S+T+), indexed by a significant interaction beginning before the probe onset (-78 ms pre-stimulus) and persisting 154 ms after.

1.4   Discussion

1.4.1   Summary of results

The work described in this chapter investigated how the brain integrates information from 100 ms samples and uses it to drive predictions about what will happen. The experimental paradigm used to address this question involved entraining alpha oscillatory activity to determine the effects of spatial and temporal predictability on a novel object recognition task. Behaviorally, spatial and temporal predictability increased probe discriminability on a same-different judgement as well as speeded response times. Inverse efficiency, which combines accuracy and reaction times and can be thought of as the amount of energy consumed by the system to produce a behavioral

outcome (TownshendAshby78TownshendAshby83), was lower for temporally predictable stimuli on average, but exhibited an increase for combined spatial and temporal predictability.

Event-related analysis of EEG data indicated that temporal predictability causes a strong periodicity (in this case, 10 Hz) phase aligned approximately to the onset of each temporally predictable entraining stimulus. This alignment approximately 180 degrees out of phase for temporally unpredictable stimuli and these differences in waveform alignment caused amplitude differences that preceded both entraining stimuli and the probe. The effects of spatial predictability generally manifested during the presentation of stimuli. There was an early divergence of probe-evoked activity caused by spatially predictable ordering of entrainers that over 100 ms through the P1 response with several transient effects after. There was evidence that spatial predictability was only effective when stimuli were also temporally predictable. This was evident for entrainers, and over right hemisphere channels during the probe. The effect failed to reach significance for left hemisphere channels and only trended toward significant levels when both hemispheres were considered jointly. It is likely that this is simply an issue of insufficient power, but other explanations such as hemispheric specialization (e.g., Dien09b) cannot be completely ruled out.

Both spatial and temporal predictability had effects on the power and phase coherence (indexed by inter-trial coherence, ITC) of neural oscillations. Spatially predictable entrainment caused a suppression of 10 Hz power with a lower degree of phase alignment than spatially random stimuli. Temporally predictable entrainment had the opposite effect, with increased 10 Hz power and phase alignment. Phase alignment due to temporal predictability remained elevated compared to temporally unpredictable stimuli during a 200 ms blank period between the entraining sequence and probe, indicating phase alignment could persist without exogenous entrainment. There was evidence of selective an enhancement in phase alignment for the combined spatial and temporal predictability case. As was the case in event-related analyses, this effect was significant in the right hemisphere, but not the left, and trended toward significance when both hemispheres were considered jointly. Power and phase coherence effects were also examined in the delta-theta bands (5 Hz) during the probe judgement. Results were similar to those found for 10 Hz, but more

robust with effects for both power and phase alignment.

1.4.2   Separate time courses for spatial and temporal prediction

Spatial and temporal predictability were characterized by distinct and generally non-overlapping time courses. Temporal predictability manifested solely before (or at) the onset of each stimulus and appeared to be driven by an approximate antiphase relationship between temporally predictable and unpredictable stimuli. Spatial predictability generally manifested during the presentation of stimuli, although in the case of the probe, showed transient difference without stimulation. Predictability effects were similar for entraining stimuli and the probe except that the effect of spatial predictability persisted for over 100 ms through the P1 response with several more transient effects after. The similarity between entrainers and the probe suggests that the brain might treat the probe as a continuation of the entraining sequence and process it in the same manner.

From these results, it can be concluded that temporal prediction is an anticipatory process, occurring during the absence of exogenous stimulation (in between entrainers or before the probe). The effect of temporal predictability extends 16 ms after the onset of each entrainer, but latencies for the first wave of responses in primary visual cortex (V1) are approximately 40-60 ms (NowakBullier97FoxeSimpson02) so there is no exogenous stimulation *per se* during the duration of the effect despite the stimulus being onscreen. Spatial prediction begins shortly before exogenous stimulation, but in the case of the probe, persists through the initial V1 responses. Spatial prediction, thus, might better be characterized as a post-stimulus process opposed to truly anticipatory process. The computation might involve a comparison between what is expected and what is actually coded by incoming spikes, consistent with the LeabraTI model (Chapter **??**) as well as predictive coding models (e.g., RaoBallard99Friston05DenOudenKokDeLange12).

1.4.3   Oscillatory mechanisms of spatial and temporal prediction

Spatial and temporal predictability had effects on both power and phase coherence of neural oscillations. In both cases, predictability took 2-3 entraining stimuli to establish, consistent

with previous investigations (MathewsonFabianiGrattonEtAl10; MathewsonPrudhommeFabiani-EtAl12). Spatial predictability was then characterized by a suppression of 10 Hz power and a phase angle variability causing a lower degree of alignment than spatially random stimuli. This effect is opposite than that of temporal predictability, which was characterized by increased 10 Hz power and phase angle alignment. Previous investigations have generally not simultaneously manipulated temporal rhythmicity and spatial coherence (although see DohertyRaoMesulamEtAl05RohenkohlGouldPessoaEtAl14) and thus, the relative suppression and decreased phase coherence during spatially predictable entrainment were unexpected results.

Successful oscillatory entrainment is thought to be a result of repeated phase resetting of endogenous oscillations causing phase to move into alignment with the frequency of exogenous stimulation (SchroederLakatosKajikawaEtAl08CalderoneLakatosButlerEtAlInPress). Oscillatory phase resetting has been shown to be caused by salient, unexpected events (FiebelkornFoxeButlerEtAl11; LandauFries12RomeiGrossThut12) and thus these unexpected results might be accounted for by considering successive views in terms of the amount of "surprise" (IttiBaldi09MeyerOlson11) they evoke. Subsequent views of the entraining sequence have significant feature overlap, characterized by the same populations of neurons spiking from one view to the next. Repeated spiking, especially as a function of expectation, has been shown to evoke rate suppression mechanisms (SummerfieldTrittschuhMontiEtAl08). When entraining views are presented out of order, feature overlap is minimized and each view is "surprising", causing an initial fast spiking burst response, which could lead to a higher degree of phase resetting.

The entrainment effects of spatial predictability were transient and dissipated before the end of the entraining sequence whereas the effects of temporal predictability persisted much longer. This might account for the null effects of spatial predictability on most behavioral measures. Temporal predictability effects persisted through the 200 ms blank period between the entraining sequence and probe and had robust effects on all behavioral measures. Thus, the present experiment could be modified to use a shorter entrainment sequence which would likely elicit successful spatial predictability for probe judgements.

The enhancement of 10 Hz phase angle alignment after the probe was presented did not reach significant levels assuming the FDR-corrected significance test at each time bin but did for 5 Hz phase alignment. Furthermore, temporal and spatial predictability main effects around the probe onset were more pronounced for 5 Hz oscillatory properties. One potential explanation for these effects is that reason for this is that the 200 ms blank period between the entraining sequence and the probe corresponded to two cycles at 10 Hz, but only one cycle at 5 Hz. Phase angles at 10 Hz exhibited significant dealignment over this period, and actually increased in alignment at 5 Hz. Thus, it is possible that this increase in phase alignment lead to a more pronounced selective enhancement for combined spatial predictability at 5 Hz, consistent with recent data (CravoRohenkohlWyartEtAl13). A more robust effect might be found for 10 Hz if the blank period between the entraining sequence and probe was only 100 ms in duration.

Another limitation of the present experimental paradigm is that it is unclear whether exogenous entrainment simply created new oscillations akin to steady state visually evoked potentials (SSVEP), or actually entrained existing endogenous oscillations. Thus, it might not be particularly surprising that temporally predictable stimuli caused increases in power and phase alignment. However, entrained alpha-band periodicity has been shown to correlate with individual resting alpha oscillation frequency (DeGraafGrossPatersonEtAl13) so it is likely that the paradigm recruited existing oscillations and caused them to align to the exogenous entrainment frequency. The fact that phase alignment continued through the 200 ms blank period without exogenous entrainment also supports this claim.

1.4.4   Alpha oscillations index stimulus-predictive processing

Alpha oscillations have previously been implicated in the allocation of hemifield-based spatial attention and anticipation of the temporal onset of relatively simple stimuli (GouldRushworthNobre11; BelyusarSnyderFreyEtAl13; RohenkohlNobre11MathewsonGrattonFabianiEtAl09BuschDuboisVanrullen09). It was unclear whether these results engaged actual predictive processing about what would happen or comparatively

simple anticipatory attention mechanisms about *where* a stimulus might appear.

The current experiment consisted of a relatively complex perceptual tasks that that required from integration across features extracted over the course of a sequence of stimulus presentations to perform a subsequent probe judgment. Both spatial and temporal predictability benefited probe discriminability and so it can reasonably be concluded that the experiment engaged actual predictive processing. Alpha oscillations indexed both the spatial and temporal predictability of the entraining stimuli as well as tracked the onset of the probe through phase alignment during the 200 ms blank period separating the entraining sequence and probe.

Cumulatively, these results suggest that alpha oscillations serve at least two roles in prediction that can roughly be characterized as "spatial" and "temporal" in nature. The first is a prediction about the spatio-featural content of upcoming sensory events. This prediction could simply be about where a stimulus might show up, irrespective of what features it contains. Importantly however, the prediction can also carry content about the probability of specific features showing up at a specific location in spatial map(KokRahnevJeheeEtAl12WyartNobreSummerfield12HorschigJensenVanSchouwenburgEtAl13). The second function of alpha oscillations is a pacemaker function that allows predictions to be made at regular intervals. This pacemaker function can phase align to the onset of stimuli (SchroederLakatosKajikawaEtAl08CalderoneLakatosButlerEtAlInPress) so that predictions can reliably be made in advance of actual sensory events. Overall, these suggested roles of alpha oscillations are compatible with the LeabraTI model (Chapter **??** and other models of sensory predictions that highlight the role of oscillations in making about making predictions about incoming sensory information (ArnalGiraud12GiraudPoeppel12).