








Not All Predictions Are Equal: “What” and “When” Predictions Modulate Activity in Auditory Cortex through Different Mechanisms

 Ryszard Auksztulewicz,^{1,2,3,11}  Caspar M. Schwiedrzik,^{4,5,6,7} Thomas Thesen,⁸  Werner Doyle,⁸  Orrin Devinsky,⁸  Anna C. Nobre,² Charles E. Schroeder,^{9,10}  Karl J. Friston,¹ and  Lucia Melloni^{8,11}

¹Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, London, United Kingdom WC1N 3BG, ²Oxford Centre for Human Brain Activity, Department of Psychiatry, University of Oxford, Oxford, United Kingdom OX3 7JX, ³Department of Biomedical Sciences, City University of Hong Kong, Hong Kong, ⁴Laboratory of Neural Systems, Rockefeller University, New York, New York 10065, ⁵Neural Circuits and Cognition Laboratory, European Neuroscience Institute, Göttingen, Germany 37077, ⁶University Medical Center Goettingen, 37075 Göttingen, Germany, ⁷Cognitive Neuroscience Laboratory, German Primate Center, Göttingen, Germany 37077, ⁸Department of Neurology, New York University Langone Medical Center, New York, New York 10017, ⁹Departments of Neurological Surgery and Psychiatry, Columbia University College of Physicians and Surgeons, New York, New York 10032, ¹⁰Cognitive Neuroscience and Schizophrenia Program, Nathan Kline Institute, Orangeburg, New York 10962, and ¹¹Department of Neuroscience, Max-Planck-Institute for Empirical Aesthetics, Frankfurt am Main, Germany 60322

Using predictions based on environmental regularities is fundamental for adaptive behavior. While it is widely accepted that predictions across different stimulus attributes (e.g., time and content) facilitate sensory processing, it is unknown whether predictions across these attributes rely on the same neural mechanism. Here, to elucidate the neural mechanisms of predictions, we combine invasive electrophysiological recordings (human electrocorticography in 4 females and 2 males) with computational modeling while manipulating predictions about content (“what”) and time (“when”). We found that “when” predictions increased evoked activity over motor and prefrontal regions both at early (~180 ms) and late (430–450 ms) latencies. “What” predictability, however, increased evoked activity only over prefrontal areas late in time (420–460 ms). Beyond these dissociable influences, we found that “what” and “when” predictability interactively modulated the amplitude of early (165 ms) evoked responses in the superior temporal gyrus. We modeled the observed neural responses using biophysically realistic neural mass models, to better understand whether “what” and “when” predictions tap into similar or different neurophysiological mechanisms. Our modeling results suggest that “what” and “when” predictability rely on complementary neural processes: “what” predictions increased short-term plasticity in auditory areas, whereas “when” predictability increased synaptic gain in motor areas. Thus, content and temporal predictions engage complementary neural mechanisms in different regions, suggesting domain-specific prediction signaling along the cortical hierarchy. Encoding predictions through different mechanisms may endow the brain with the flexibility to efficiently signal different sources of predictions, weight them by their reliability, and allow for their encoding without mutual interference.

Key words: associative learning; biophysical modeling; electrocorticography; prediction; predictive coding

Significance Statement

Predictions of different stimulus features facilitate sensory processing. However, it is unclear whether predictions of different attributes rely on similar or different neural mechanisms. By combining invasive electrophysiological recordings of cortical activity with experimental manipulations of participants’ predictions about content and time of acoustic events, we found that the two types of predictions had dissociable influences on cortical activity, both in terms of the regions involved and the timing of the observed effects. Further, our biophysical modeling analysis suggests that predictability of content and time rely on complementary neural processes: short-term plasticity in auditory areas and synaptic gain in motor areas, respectively. This suggests that predictions of different features are encoded with complementary neural mechanisms in different brain regions.

Introduction

A central aspect of brain operation is the extraction of regularities from the sensorium to form expectations about imminent sensations, used to guide behavior (Friston, 2010). Predictive capacities are regarded as a general principle of brain function (Knill and Pouget, 2004; Summerfield and de Lange, 2014), often proposed to rest on domain-general mechanisms: that is, neuronal message passing in canonical microcircuits (Bastos et al., 2012) along the cortical hierarchy (Kiebel et al., 2008). However, expectations about upcoming stimuli can regard different environmental features: stimulus timing (Griffin et al., 2002), location (Chikkerur et al., 2010), or identity itself (Arnal and Giraud, 2012). Are all predictions implemented by the same neural mechanism? Separately encoding predictions regarding uncorrelated aspects of the environment would entail functional specialization (Zeki and Shipp, 1988; Friston and Buzsáki, 2016), with dissociable sources of predictability modulating neural activity through different mechanisms (e.g., neurotransmitters and neuromodulators) (Schultz and Dickinson, 2000; Yu and Dayan, 2005; Baldeweg, 2007; Iglesias et al., 2013; Marshall et al., 2016), possibly in distinct regions (Kanai et al., 2015; Friston and Buzsáki, 2016). A key advantage of encoding predictions through complementary mechanisms is that uncorrelated, unequally reliable, or even opposing predictions from different domains can be signaled without interference.

Evoked EEG/MEG responses have been shown to differ for different kinds of predictions (Doherty et al., 2005; SanMiguel et al., 2013; Lau and Nguyen, 2015), providing initial evidence for their domain specificity. Here we independently manipulated content-based (“what”) and time-based (“when”) predictions in an audiovisual associative learning task to test mechanistic hypotheses of how predictions might be implemented at the level of interactions between brain regions and neuronal populations. We combine direct electrophysiological recordings from the cortex of epilepsy patients with extensively validated biophysical models (Moran et al., 2011; Papadopoulou et al., 2015) to explain the observed effects on evoked responses and thereby shed light on the plausible neurophysiological mechanisms subserving “what” and “when” predictions.

“What” predictability, often studied in associative learning paradigms where paired stimuli form a fixed contingency (Jiang et al., 2012; McGann, 2015; Schwiedrzik and Freiwald, 2017), has been shown to induce anticipatory signals in low-level sensory cortex (den Ouden et al., 2009; Turk-Browne et al., 2010; Luft et al., 2015). Associative learning is thought to rely on synaptic plas-

Table 1. Demographic data of the participants enrolled in the study^a

Patient no.	Age (yr)	Gender	Handedness	Implant hemisphere	Seizure type	Age of onset (yr)	VCI	WMI	POI	PSI
1	24	Female	R	L	CPS	15	108	89	107	84
2	27	Male	R	L	CPS	19	100	92	107	97
3	27	Male	L	L/R	CPS	17	NA	NA	NA	100
4	25	Female	R	L/R	CPS	18	102 ^b			85 ^c
5	56	Female	R	L	CPS	43	93	89	96	92
6	24	Female	R	L	CPS	2	98	105	96	105

^aCPS, Complex Partial Seizure; VCI, Verbal Comprehension Index; WMI, Working Memory Index; POI, Perceptual Organisation Index; PSI, Processing Speed Index.

^bCompound Verbal IQ scale.

^cCompound Perceptual IQ scale.

ticity mediated by activity-dependent (i.e., voltage-dependent) NMDA signaling (Xia et al., 2005) and local disinhibition (Letzkus et al., 2015) of principal cells, consistent with a high concentration of voltage-sensitive NMDA receptors in superficial layers targeted by descending connections (Rosier et al., 1993). Thus, we hypothesized that “what” predictability of acoustic stimuli might increase neural sensitivity in auditory regions due to activity-dependent gain modulation, whereby the postsynaptic responsiveness of principal cells is modulated by descending inputs from other regions, enabling short-term plasticity.

“When” predictability has also been linked to sensory gain modulation (Schroeder and Lakatos, 2009; Rohenkohl et al., 2012). Converging evidence suggests that this modulation has motor, rather than sensory, origins, especially in rhythmic contexts (Schubotz et al., 2000; Cravo et al., 2011; Morillon et al., 2015). Dopaminergic manipulations affect timing and “when” predictability, partly relying on the nigrostriatal motor pathway (Coull et al., 2012; Narayanan et al., 2012; Parker et al., 2013), consistent with the dual-stream hypothesis and the involvement of motor regions in the coding of sound sequences (Leaver et al., 2009; Rauschecker, 2012; Bornkessel-Schlesewsky et al., 2015). Because such classical neuromodulatory effects are not necessarily voltage-dependent (Formenti et al., 1998; Gorelova et al., 2002), we hypothesized that “when” predictability might be mediated by activity-independent gain modulation (whereby the gain of principal cells is not directly modulated by descending cortical inputs, but regulated by putative classical neuromodulators, and thus prone to more distal and subcortical influences) expressed in motor and/or sensory regions.

Thus, beyond showing regionally specific effects of “what” and “when” predictability, we sought to dissociate activity-dependent from activity-independent gain modulation using biophysically realistic neural mass (dynamic causal) models.

Materials and Methods

Participants. Six individuals with pharmacologically intractable epilepsy (4 female, 2 male; mean age, 30.5 years; age range, 24–56 years; mean number of years since epilepsy diagnosis, 11.5 years; range, 9–22 years; all right-handed except for 1 left-handed patient; compare Table 1) participated in this study. Table 1 presents demographic data for each participant. All patients had electrocorticographic (ECoG) electrodes implanted as part of presurgical diagnosis of epilepsy. Data collection was performed at the Comprehensive Epilepsy Center of New York University Langone Health and was approved by the Institutional Review Board at New York University Langone Health. All patients provided oral and written informed consent before participation in the study, in accordance with the Declaration of Helsinki.

ECoG recordings. All patients had 8 × 8 grids of subdural platinum-iridium electrodes embedded in Silastic sheets (2.3-mm-diameter contacts, Ad-Tech Medical Instruments) with a minimum 10 mm

Received Feb. 9, 2018; revised July 22, 2018; accepted July 26, 2018.

Author contributions: R.A., C.M.S., T.T., W.D., O.D., A.C.N., C.E.S., K.J.F., and L.M. wrote the first draft of the paper; R.A., C.M.S., A.C.N., C.E.S., K.J.F., and L.M. edited the paper; R.A., C.M.S., C.E.S., and L.M. designed research; T.T., W.D., O.D., and L.M. performed research; T.T., W.D., O.D., and L.M. contributed unpublished reagents/analytic tools; R.A., A.C.N., K.J.F., and L.M. analyzed data; R.A., C.M.S., and L.M. wrote the paper.

This work was supported by a Marie Curie International Outgoing Fellowship within the 7th European Community Framework Programme to L.M., Human Frontier Science Program Long-Term Fellowship LT001118/2012-L to C.M.S., European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant Agreement 706519 to C.M.S., and National Institutes of Health Grants MH103814 and EY024776 to C.E.S. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Health. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We thank Callah Boomhaur and Preet Minhas for help with data collection; and Hugh Wang for providing the electrode localization.

The authors declare no competing financial interests.

Correspondence should be addressed to either of the following: Dr. Ryszard Aukstulewicz, Department of Biomedical Sciences, 31 To Yuen St, Kowloon Tong, Hong Kong, E-mail: rauksztu@cityu.edu.hk; or Dr. Lucia Melloni, Department of Neurology, New York University Langone Medical Center, 222 East 41st Street - 14th Floor, New York, NY 10017, E-mail: lucia.melloni@nyumc.org.

DOI:10.1523/JNEUROSCI.0369-18.2018

Copyright © 2018 the authors 0270-6474/18/388681-14\$15.00/0

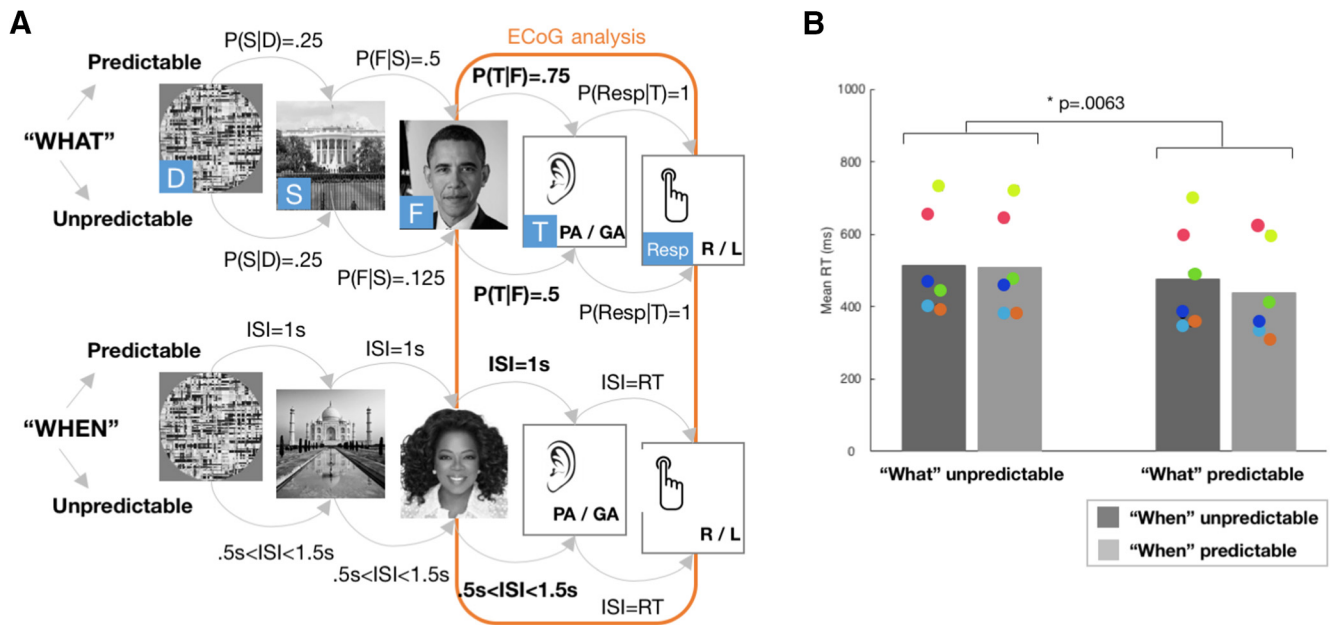


Figure 1. Experimental paradigm and behavioral results. **A**, In each trial, participants categorized auditory syllables, which could be predictable or unpredictable with respect to their content (“what”) and onset (“when”). Each trial contained a dummy image (D), followed by a scene image (S), face image (F), and a target syllable (T), followed by a response (Resp). “What” predictability was manipulated as the probability of specific stimuli being presented, given previous stimuli. “When” predictability was manipulated by fixing or jittering the interstimulus intervals (ISI) between each two stimuli. **B**, Analysis of RTs revealed that participants were faster in categorizing syllables when their contents were predictable. Additionally, under “what” predictability, syllables with predictable temporal onsets were categorized faster than those with unpredictable onsets.

center-to-center distance implanted over the temporal/frontal cortices (1 in right hemisphere, 5 in left hemisphere), with additional linear strips of electrodes ($1 \times 8/12$ contacts), or depth electrodes ($1 \times 8/12$ contacts), or combinations thereof. Recordings from grid, strip, and depth electrode arrays were made using a Nicolet ONE clinical amplifier (Natus), bandpass filtered from 0.5 to 250 Hz, and digitized at 512 Hz. Signals were online referenced to a screw bolted to the skull and common-average referenced offline. The analysis presented below focused only on grid electrodes (see Fig. 2A). The number of electrodes recorded in each patient’s grid varied between 31 and 58 electrodes (mean, 40 electrodes). Electrode localization followed previously described procedures (Yang et al., 2012). In brief, for each patient, we obtained preoperative and postoperative T1-weighted MRIs, which were coregistered with each other and normalized to an MNI-152 template, allowing the extraction of the electrode location in MNI space. Electrode labels were assigned using FreeSurfer cortical parcellation (RRID:SCR_001847) (Fischl et al., 2004) based on the Desikan-Killiany atlas (Desikan et al., 2006).

Experimental design and statistical analysis. The behavioral paradigm (Fig. 1A) was based on a 2×2 factorial design with factors “when” predictability and “what” predictability, resulting in 4 different conditions presented in separate runs. Each run consisted of 96 trials. A fixation cross shown for a variable duration of 1.5–2 s marked the beginning of the trial. The fixation cross was followed by a dummy image (composed of grayscale random horizontal and vertical lines) that marked the beginning of the sequence followed by a picture of a scene, a picture of a face, and an auditory syllable. Each image was presented for 210 ms. The order of stimulus types was fixed, but different stimulus exemplars were presented across trials, selected from 4 different scene images, 8 different face images, and 2 different syllables (“PA” and “GA”). Participants were instructed to categorize the last syllable (“PA” or “GA”) with a speeded button press using the index and middle fingers of the right hand. Syllables were produced by a male speaker and recorded in-house. For the experiment, they were played with speakers at 70 dB or levels comfortable for the patient and controlled by Presentation (Neurobehavioral Systems). Visual stimuli were displayed foveally on a laptop placed at the bedside (~70 cm distance). Visual stimulation was included in the current paradigm because a subset of patients in the larger participant pool had additional electrode grids implanted over their occipital regions;

these data will be analyzed and reported separately. A key feature of our cross-modal design is that it allows us to test for the effect of predictability in the baseline, without any confound from preceding stimuli. In particular, for the analysis reported here, it is possible to study responses in auditory cortex independently from the activity elicited by the predictor stimulus, in this case the face. This is important because stimulus effects persist for a substantial amount of time and thus can depend on responses to the previous stimuli (e.g., via adaptation effects).

Temporal (“when”) predictability was manipulated using regular or random temporal intervals between the stimuli, respectively. In the temporally predictable conditions, we presented four stimuli in a sequence with a fixed stimulus onset asynchrony (SOA) at 1 s between each two stimuli. In the temporally unpredictable conditions, the SOA for each two stimuli in the sequence varied from 0.5 s to 1.5 s (i.e., with a mean SOA of 1 s but with a random jitter $\pm 0-0.5$ s), in line with previous manipulations of temporal predictability (e.g., Besle et al., 2011; Cravo et al., 2013; Morillon et al., 2016).

Content (“what”) predictability was manipulated using contingencies between individual stimuli. In the content-predictable condition, there was a nonarbitrary semantic relationship between the content of the scene image and the content of the face image. For instance, an image of the White House predicted with 100% probability that the face image would be an American president (Barack Obama or George W. Bush) with 50% presentation of each individual image. Each face image in turn was associated with a 75% presentation probability of one syllable over the other. Supramodal serial cueing was used because, in a subset of patients, the grid and/or strip electrodes extended into the occipital regions; these data will be reported separately. The stimulus–stimulus contingencies were fixed for each participant throughout the experiment. In the content-unpredictable condition, the presentation of each stimulus was equiprobable and no stimulus–stimulus contingencies were defined. Four different scene images and eight different face images were used per condition, respectively.

The four conditions were tested in separate runs after participants had completed a practice session of the completely predictable condition. The practice session served to instruct the participants and help them learn the “when” and “what” associations between the different stimuli. No explicit instructions regarding stimulus predictability were given in

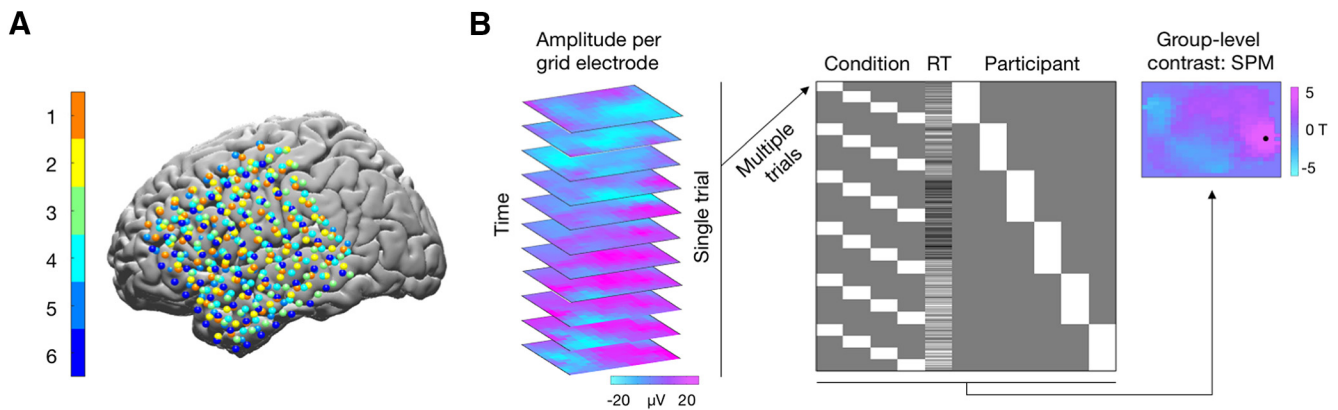


Figure 2. ECoG analysis. *A*, Location of ECoG grid electrodes per participant. *B*, Analysis pipeline. Single-trial data were converted from electrode \times time into grid \times time images and entered into a general linear model (GLM) coding for condition (in a 2×2 design with factors “what” and “when” predictability), single-trial RT, and participant. Group-level contrasts were designed to obtain statistical parametric maps (grid \times time; here only plotted for one time point).

this block or in the four subsequent blocks in which predictability was manipulated. It is worth noting that the task was simply to categorize syllables; thus, participants did not need to rely on preceding images to solve the task correctly, but doing so would be optimal. In total, patients completed five blocks (1 practice and 4 experimental sessions) of 96 trials, each lasting just <10 min. In total, with short breaks between the blocks, the entire experiment lasted ~ 1 h.

Single-trial reaction time (RT) data were analyzed using mixed-effects modeling with fixed factors “what” predictability and “when” predictability, and a mixed factor participant (Fig. 2*B*). Trials with no button press within the first 2000 ms after syllable onset were discarded from further analysis ($<1\%$ of all trials). RTs were log-transformed before analysis to correct for the skewness of their distribution. Accuracy data were analyzed in a binomial logistic regression (dependent variable: hit/miss) with predictors: “when” predictability, “when” predictability, and subject. Additionally, we analyzed the effect of “what” prediction validity (i.e., the difference between the 75% of the tones that could be correctly predicted and the 25% of the tones that were mispredicted) in the “what”-predictable conditions in another binomial logistic regression (dependent variable: hit/miss) with predictors: “when” predictability, validity, and subject.

Electrophysiological data analysis was performed in SPM12 (Wellcome Trust Centre for Neuroimaging, University College London; RRID: SCR_007037) for MATLAB (The MathWorks; RRID: SCR_001622) and using custom MATLAB code. Continuous signals were low-pass filtered at 240 Hz and notch-filtered at 58–62 Hz and harmonics using zero-phase Butterworth filters. Data were then downsampled to 200 Hz to speed up calculation time in the statistical parametric mapping analysis.

Continuous data were epoched into segments from 100 ms before auditory syllable onset to 500 ms after syllable onset. Epochs were baseline-corrected to the prestimulus period (-100 – 0 ms relative to syllable onset). For the “what”-predictable conditions (independent of their “when” predictability), we restricted analyses to those trials containing valid prediction (75% of the trials). To maintain the same proportion of trials across all conditions, in the respective content-unpredictable conditions, we selected a random subset of 75% trials. To reject trials contaminated by artifacts and ictal activity, a channel with the maximum amplitude of the auditory evoked response (averaged across trials) was selected per participant (see Fig. 3*D*). In all participants, these electrodes were implanted over, or adjacent to, the superior temporal gyrus (STG). Given the high signal-to-noise ratio of auditory evoked responses, these electrodes were used to reject trials in which no (stereotypical) auditory evoked response was observed. Operationally, single trials with both auditory evoked extrema at latencies ± 1 SD outside of the average latency were discarded from further analyses. By adopting this criterion, we rejected trials in which the auditory evoked response had, on average, its maximum (minimum) peak at least 154.1 ms (103.1 ms) earlier or later than the mean peaks for the given participant. The

remaining artifacts were rejected upon a visual inspection of single trials. On average, 12.5% (SD 7.03%) trials were rejected per participant and condition.

In the analysis of evoked responses (see Fig. 3*B*), rather than testing for the effects of “what” and “when” predictability on signals acquired from single electrodes, we created 2D maps of stimulus-evoked ECoG amplitude by projecting the 3D MNI coordinates of the electrode grid onto the sagittal plane and converting them into images of 32×32 pixels (removing the laterality of grid location). For group-level inference, single-trial data were entered into a factorial design with participant as a fixed effect, and “when” predictability and “what” predictability as random effects. Group-level statistical parametric maps were thresholded at a family-wise error (FWE)-corrected $p = 0.05$ (peak-level) (Kilner et al., 2005), thereby implementing a correction for multiple comparisons while taking into account the spatiotemporal smoothness of these data features. In an additional *post hoc* analysis, we tested for the effects of “what” prediction validity (i.e., the difference between the 75% of the tones that could be correctly predicted and the 25% of the tones that were mispredicted) in another factorial design with participant as a fixed factor, and “when” predictability and validity as random factors.

Dynamic causal modeling (DCM). DCM for evoked responses (Pinotsis et al., 2012) was used to model the effects of “when” and “what” predictability on the amplitude and dynamics of the evoked response in terms of the underlying neurophysiology. DCM enables the fitting of observed neural responses using biologically realistic mean-field models of coupled dynamical systems, and the modeling of differences between experimental conditions in terms of specific model parameters. Here (robust-averaged) single-subject evoked responses were modeled in a network of three interacting cortical sources, corresponding to three subdural electrodes, chosen on the basis of the significant main and interaction effects of “what” and “when” predictability at the single-subject level, as described in Results (see also Fig. 3*E, F*). Specifically, per participant, we chose electrodes lying closest to the spatial coordinates of the group-level significant effect peaks. To obtain a coherent model space for our participant sample, we only included regions consistently identified as sensitive to at least one experimental manipulation at the group level. Selecting different regions per participant would have made it unfeasible to integrate the results across single participants’ datasets. Because the aim of applying DCM to explain the observed effects was to infer the putative physiological mechanisms mediating our experimental manipulations (i.e., “what” and “when” predictability), we used evoked responses averaged over trials, which increases the signal-to-noise ratio and ensures that the models can be efficiently fitted to the data. Each cortical source was modeled with a canonical microcircuit (Bastos et al., 2012; Pinotsis et al., 2012) comprising four neural populations: pyramidal cells in supragranular and infragranular layers, spiny stellate cells, and inhibitory interneurons (see Fig. 4*B*). The dynamics at each source were modeled with the following coupled differential equations:

$$\dot{V}_{SS} = I_{SS} \quad (1)$$

$$\begin{aligned} \dot{I}_{SS} = & \kappa_{SS}(A^F\sigma(V_{SP}) - \gamma_{SS\rightarrow SS}\sigma(V_{SS}) - \gamma_{SP\rightarrow SS}\sigma(V_{SP}) \\ & - \gamma_{II\rightarrow SS}\sigma(V_{II})E(x)) - 2\kappa_{SS}V_{SS} - \kappa_{SS}^2I_{SS} \quad (2) \end{aligned}$$

$$\dot{V}_{II} = I_{II} \quad (3)$$

$$\begin{aligned} \dot{I}_{II} = & \kappa_{II}(-A^B\sigma(V_{DP}) + \gamma_{SS\rightarrow II}\sigma(V_{SS}) + \gamma_{DP\rightarrow II}\sigma(V_{DP}) \\ & - \gamma_{II\rightarrow II}\sigma(V_{II})) - 2\kappa_{II}V_{II} - \kappa_{II}^2I_{II} \quad (4) \end{aligned}$$

$$\dot{V}_{SP} = I_{SP} \quad (5)$$

$$\begin{aligned} \dot{I}_{SP} = & \kappa_{SP}(-A^B\sigma(V_{DP}) + \gamma_{SS\rightarrow SP}\sigma(V_{SS}) - \gamma_{SP\rightarrow SP}\sigma(V_{SP})) \\ & - 2\kappa_{SP}V_{SP} - \kappa_{SP}^2I_{SP} \quad (6) \end{aligned}$$

$$\dot{V}_{DP} = I_{DP} \quad (7)$$

$$\begin{aligned} \dot{I}_{DP} = & \kappa_{DP}(A^F\sigma(V_{SP}) - \gamma_{DP\rightarrow DP}\sigma(V_{DP}) - \gamma_{II\rightarrow DP}\sigma(V_{II})) - 2\kappa_{DP}V_{DP} \\ & - \kappa_{DP}^2I_{DP} \quad (8) \end{aligned}$$

$$\dot{\gamma}_{SP\rightarrow SP} = \gamma_{SP\rightarrow SP}M\sigma(V_{DP}) \quad (9)$$

Here, neuronal populations are denoted by subscripts (SS, Spiny stellate cells; II, inhibitory interneurons; SP, superficial pyramidal cells; DP, deep pyramidal cells). V_m and I_m denote the voltage and current, respectively, of each population m , characterized by a synaptic rate constant κ_m . A sigmoid operator σ transforms the postsynaptic potential into firing rate. Forward and backward (extrinsic) connections between regions are denoted by A^F and A^B , and the (intrinsic) connections from population m to n within a region are denoted by $\gamma_{m\rightarrow n}$. Spiny stellate cells are assumed to receive inputs $E(x)$. Equations 1–8 describe the dynamics of voltage and current in all four neuronal populations. In each population, the change in voltage depends on the current, whereas the change in current is a nonlinear function of voltage. For instance, the change in current of the superficial pyramidal cells (Eq. 6) depends on the depolarization of deep pyramidal cells in hierarchically higher regions, weighted by the strength of the descending connections from those regions; the depolarization of spiny stellate cells, weighted by the strength of the intrinsic connection between the two populations; and the voltage of superficial pyramidal cells themselves, weighted by their self-inhibitory intrinsic connections. Additionally, Equation 9 models activity-dependent modulation of the gain of superficial pyramidal cells. Per cortical region included in the model, the observation model mapped the estimated time-series of source-level time-series (a linear weighting of deep and superficial pyramidal cell activity) onto the observed sensor-level data, as in previous DCM work using ECoG (Pinotsis et al., 2012; Phillips et al., 2016). This model has been used in several other DCM studies of evoked responses (Brown and Friston, 2012; Moran et al., 2013; Auksztulewicz and Friston, 2015), and neurophysiological inference using DCM has been validated in animal models (Moran et al., 2011) and invasive recordings in humans (Papadopoulou et al., 2015).

The DCM analysis aims to disambiguate between alternative hypotheses regarding the mechanisms underlying “when” and “what” predictability. Thus, alternative models were designed, each allowing for a different subset of connections to be modulated by “when” and/or “what” predictability. Specifically, we asked whether the synaptic gain in any of three cortical sources (i.e., sensory, motor or frontal cortex [FC]) or any combination thereof needed to be modulated by “when” and/or “what” predictability to best explain the observed evoked responses. Furthermore, two candidate mechanisms were compared against each other: (1) the gain modulation is activity-independent, due to putative classical neuromodulatory (e.g., dopaminergic) effects ($M = 0$ in Eq. 9); and (2) the gain modulation is activity-dependent, due to putative NMDA-mediated short-term plasticity ($M \neq 0$ in Eq. 9; compare Fig. 4A). In the context of this DCM, activity-dependent changes in synaptic connectivity are modeled by intrinsic (self-inhibitory) connections that are a sigmoid function of neuronal activity (here, of the population in question).

Thus, activity-dependent gain modulation is scaled by neuronal inputs from other regions. Activity-independent gain modulation, on the other hand, translates into disinhibition of a given region without additional weighting of the strength of this disinhibition by inputs from other regions. We focused on those three cortical regions, given the often conflicting evidence for the effects of predictability at the level of sensory (Griffin et al., 2002; den Ouden et al., 2009; Alink et al., 2010; Turk-Browne et al., 2010; Arnal and Giraud, 2012; Lakatos et al., 2013; Summerfield and de Lange, 2014; Luft et al., 2015), motor (Morillon et al., 2015), and FC, including the inferior and middle frontal gyri (den Ouden et al., 2009; Turk-Browne et al., 2010; Coull et al., 2011); and also because our electrophysiological results showed modulations of those areas as a function of “what” and “when” predictions, a prerequisite for DCM.

The resulting model space allowed us to infer which of the distinct neural mechanisms subserve “what” and “when” predictability, respectively (Coull et al., 2011, 2012; Narayanan et al., 2012; Parker et al., 2013). Because we tested for orthogonal effects of “when” and “what” predictability, the model space (see Fig. 4A) comprised 208 alternative models (8×8 models with all combinations of activity-independent gain modulation by two factors at all three sources, 8×8 models with all combinations of activity-dependent gain modulation in STG and activity-independent gain modulation in the remaining regions, 8×8 models with all combinations of activity-dependent gain modulation in precentral gyrus (PG) and activity-independent gain modulation in the remaining regions, and 8×8 models with all combinations of activity-dependent gain modulation in FC and activity-independent gain modulation in the remaining regions, minus 48 duplicate models). In all models, extrinsic connections (between regions) could be modulated by both experimental factors (predictability of “what” and “when”).

The three-source models were fitted to the poststimulus window (0–500 ms relative to syllable onset) per participant. Dynamic causal models are generative models: that is, they can generate simulated sensor-level data given model parameters, such as connection strengths, gain parameters, and time constants. By fitting these models to the observed sensor-level data (i.e., inverting the models), one can estimate posterior model parameters, which can reproduce the observed data as closely as possible. Each model was fitted to the observed data using a variational Bayes scheme (variational class), an iterative model inversion procedure akin to the expectation-maximization algorithm, to obtain the free-energy approximation to its log-evidence (Friston et al., 2007) and the posterior parameter estimates (together with their posterior covariance). The free-energy approximation provides a lower bound on model evidence and is expressed as a sum of accuracy and complexity (i.e., models are scored as having higher evidence when they accurately fit the data but are penalized for model complexity). To account for uncertainty about posterior parameters inherent in performing Bayesian model selection across a large number of similar models, we used random-effects Bayesian model averaging (BMA) (Penny et al., 2010). BMA uses the entire model space within a family of models (in our case, all models) by assigning a weight to the parameters of each model according to the model’s log evidence. Thus, it optimally uses all available information, weighting across models according to their reliability; additionally, it has the advantage of accommodating uncertainty over models when no single model is clearly winning. Parameters were only considered significant when differing from baseline with >99.9% posterior probability (based on their posterior variance estimated during model inversion). It is worth noting that only a subset of parameters were significantly different from baseline; this means that the optimized model did not correspond to either the most parsimonious or most complex model, and thus the model average likely represents models characterized by an optimal combination of accuracy and complexity, given our dataset. Posterior parameter estimates are reported on a logarithmic scale: positive (negative) parameter estimates correspond to increasing (decreasing) connectivity or gain relative to baseline. To convert the reported values into percentage modulation, posterior parameter estimates need to be exponentiated (e.g., the posterior mean of activity-dependent STG gain modulation by “what” predictability 0.1325 corresponds to $\exp(0.1325) = 114.17\%$ of the baseline). Thus, STG gain under “what” predictability is 14.17% stronger than under no predictability. In participants whose individual neuronal

responses did not show robust modulations by “what” or “when” predictability, null models could be chosen as best describing the data. Finally, to quantify the differential effects of specific parameters (e.g., activity-dependent vs activity-independent gain), we performed a contribution analysis of single parameters on simulated source-space responses. Specifically, we used the posterior parameter estimates of the optimized model and assessed how changes in a single parameter would translate into changes in source activity in each region (see Fig. 5B). These sensitivity profiles illustrate that different parameters explain the variance of evoked responses at different latencies and in different regions.

Results

Behavioral results

We obtained direct recordings of brain activity with electrodes implanted on the cortical surface of 6 patients undergoing pre-surgical diagnosis of medication-resistant epilepsy, while they performed an audiovisual associative learning task in which “what” and “when” predictions were independently manipulated. We first investigated whether predicting the content and/or the timing at which the stimulus would be presented leads to changes in RT in a syllable categorization task. Mixed-effect modeling of the RT data (fixed effects: “what” and “when” predictability; random effect: participant) revealed that predicting the content of the stimulus (“what predictions”) speeded syllable categorization by 38 ms on average (Fig. 1B; main effect of “what” predictions: $F_{(1,1857)} = 16.42, p = 0.009$). In contrast, predicting the onset of a stimulus (“when” predictions) did not alter RT for syllable categorization (no main effect of “when” predictions: $F_{(1,1857)} = 2.07, p = 0.209$). “When” predictions, however, speeded syllable categorization by 18 ms on average if “what” predictions were also established (interaction effect, $F_{(1,1857)} = 5.44, p = 0.019$; *post hoc* “when”: $F_{(1,802)} = 6.43, p = 0.011$). This effect disappeared if “what” predictions could not be established (“when”: $F_{(1,1050)} = 0.07, p = 0.784$). Thus, predicting the timing and the content of upcoming stimuli leads to faster syllable categorization than just anticipating the content of syllables alone (Fig. 1B).

Further evidence for the effect of content predictions in behavior was obtained by investigating the effect of validity on RT. Participants categorized syllables on average 57 ms faster in trials in which the content prediction was valid than when the content prediction was invalid (mixed-level modeling with “when” predictability and validity as fixed effects, and participant as a random effect: $F_{(1,1071)} = 8.96; p = 0.029$). We also investigated whether the validity of content predictions affected accuracy. A binomial logistic regression on responses in the “what”-predictable condition (predictors: “when” predictability, validity, and subject; dependent variable: hit/miss) revealed a significant effect of validity on response accuracy ($p < 0.001$; mean accuracy 94.79% and 85.75% for valid and invalid syllables, respectively). The main effects of “what” and “when” predictability on accuracy were not significant, likely due to ceiling-level performance because syllables were easily perceivable. Together, participants optimally adjusted their behavior such that responses speeded up when syllable categorization could be anticipated based on the availability of “what” and “when” predictions; and RTs were slowed down and accuracy was lower when content predictions were violated.

Evoked neural responses

For each participant, we analyzed the auditory evoked response amplitude at grid electrodes implanted over temporal, frontal, and parietal lobes. Evoked (low-frequency) intracranial re-

sponses provide information about neural activity that is dissociable from higher-frequency responses (Lindén et al., 2011; Lachaux et al., 2012). We focused on the evoked responses as we aimed to use biophysically realistic computational models to shed light onto the underlying physiology of those responses. These models, however, have currently only been extensively validated for evoked responses (Brown and Friston, 2012; e.g., Moran et al., 2013; Aukstulewicz and Friston, 2015), and a comparison of identical generative models of evoked ECoG signals and MEG data yielded convergent results (Phillips et al., 2016), thus allowing our results to connect to previous studies on evoked EEG/MEG investigating predictions (Doherty et al., 2005; SanMiguel et al., 2013; Lau and Nguyen, 2015). Biophysically realistic models of non-stationary-induced responses, such as high-frequency gamma responses, do exist; however, they have only been validated at time scales much longer than those used in our paradigm (Papadopoulou et al., 2015).

First, we investigated whether “what” and “when” predictions affect the amplitude of evoked responses, aiming to identify the cortical locus and latency of predictability effects. We compared the evoked response between “what” and “when” conditions in a factorial design across the entire time course (100 ms prestimulus to 500 ms poststimulus) and all electrodes (projected onto 2D maps; see Materials and Methods) by means of statistical parametric mapping, followed by correction for multiple comparisons. At earlier (165 ms) latencies, we observed an interaction between “what” and “when” predictability over the posterior STG (Fig. 3A,E). Thus, the STG was sensitive to the combination of “what” and “when” predictability. Furthermore, we found that “when” predictions selectively increased the amplitude of the evoked response over the PG, the supramarginal gyrus, and the rostral middle frontal gyrus (Fig. 3C,E; for test statistics, see Table 2) both at early (~180 ms) and late (430–450 ms) latencies. “What” predictability, however, increased evoked amplitudes over inferior and middle frontal gyri only at late latencies (420–460 ms) (Fig. 3B). These significant main or interaction effects of “what” and “when” predictability were identified in every participant; albeit, results showed some spatial heterogeneity characteristic of intracranial data (Boatman et al., 2005; Nourski et al., 2014).

“What” predictability effects emerged relatively late (after 400 ms), close to the mean RTs, raising the possibility that motor preparation might explain the previous results. To directly evaluate this possibility, we performed a control analysis on the evoked responses in which we included an additional regressor in the GLM, coding for log RT at the single-trial level. Critically, the addition of this regressor (with one value per trial) did not change any of the previously identified spatiotemporal clusters of significant effects of “what” and/or “when” predictability (thresholded at $p < 0.005$, peak level; corrected for multiple comparisons at FWE < 0.05 , cluster level, Fig. 3E,F), suggesting that motor preparation does not explain our results.

We also addressed the possibility that our results could reflect differences in hazard rates across conditions, especially for “when” predictions, as previous studies have shown that the probability of an event occurring at a particular time point modulates both the baseline and evoked activity locked to this event (Cravo et al., 2011; Mento et al., 2015). To that end, we included a second regressor coding for foreperiod duration: that is, the trial-by-trial SOA between visual (face) stimulus and the auditory target stimulus. Including this regressor did not qualitatively change the effects of “what” and “when” predictability. Indeed, no significant clusters of evoked amplitudes modulated by foreperiod were

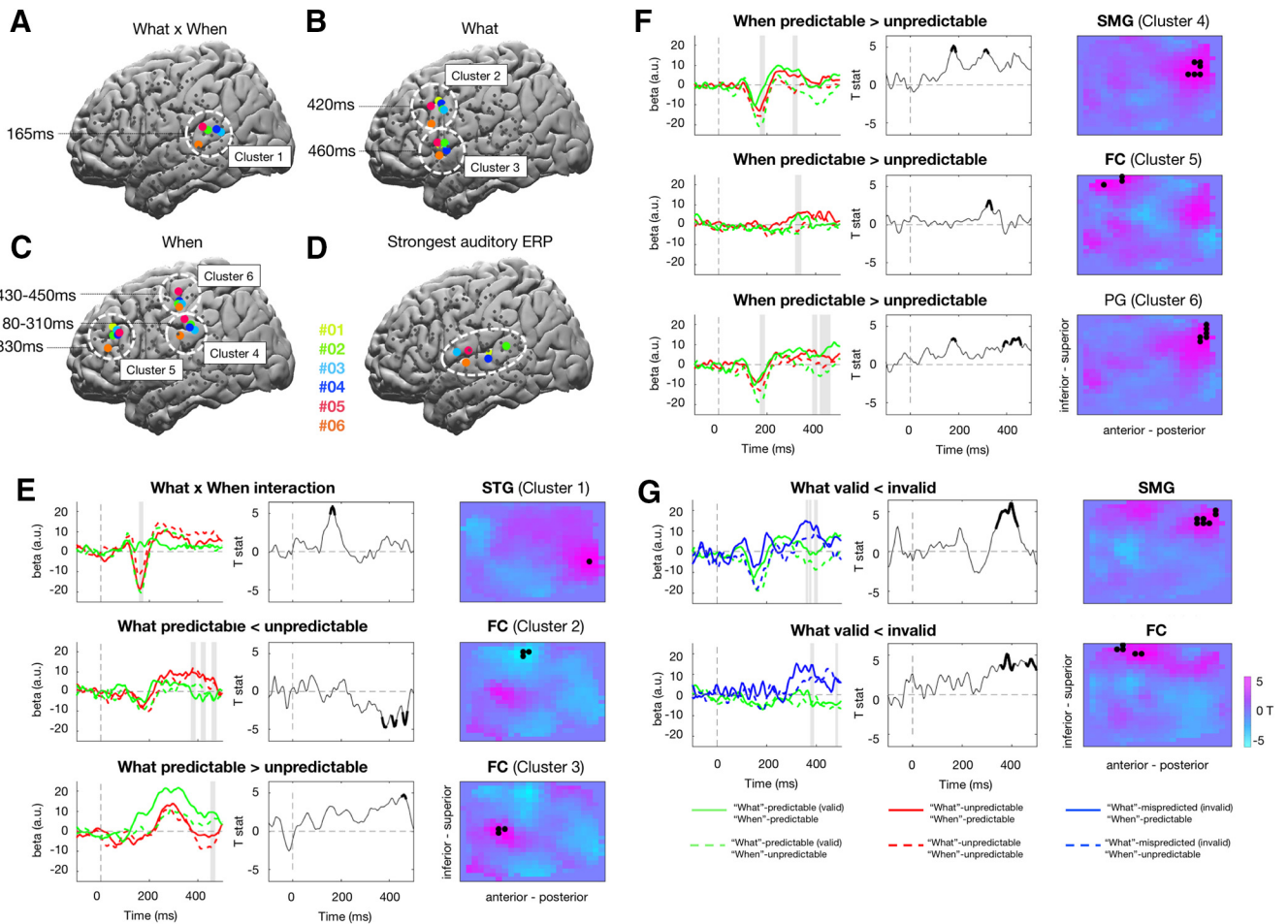


Figure 3. Group-level main and interaction effects of “what” and “when” predictability. Plots represent single participants’ (color-coded) electrodes (gray dots indicate all recording sites) adjacent to the peak group effects (used in subsequent analyses), significant at the FWE-corrected peak-level threshold $p < 0.05$. **A**, The amplitude of the auditory evoked response in STG showed an interaction effect of “what” and “when” peaking at 165 ms after stimulus. Colored dots indicate individual subjects. **B**, “What” predictability increased evoked amplitudes at frontal electrodes in two clusters peaking at 420 and 460 ms. **C**, “When” predictability increased evoked amplitudes over the supramarginal gyrus and motor regions in two clusters peaking at 180–310 ms and 430–450 ms, and at frontal electrodes (over the middle frontal gyrus) peaking at 330 ms. **D**, Location of electrodes with strongest auditory evoked responses (see Materials and Methods). **E, F**, Significant effects of “what” and/or “when” predictability while controlling for reaction speed. Rows represent individual spatiotemporal clusters with significant main or interaction effects of “what” and/or “when” predictability. Left column represents individual evoked responses. Green represents “what” predictable. Red represents “what” unpredictable. Solid line indicates “when” predictable. Dashed line indicates “when” unpredictable. Gray line indicates effect significant at FWE < 0.05 . Middle columns represent t statistic time course. Bold represents significant at FWE < 0.05 . Right columns represent topography of the effect. Color shading represents T statistic. Black dots indicate individual selected electrodes. Panel labels represent the location of the significant cluster. SMG, Supramarginal gyrus. **G**, Significant effects of “what” prediction validity while controlling for reaction speed. Legend as above. Blue lines in the left columns indicate invalid “what” predictions.

Table 2. Summary statistics of the evoked response analysis^a

Contrast	Location	Latency (ms poststimulus)	p_{FWE} cluster level	F peak level	z peak level
“What” predictable versus unpredictable	Middle frontal gyrus	420	0.017	24.57	4.80
	Inferior frontal gyrus	460	0.017	23.33	4.67
“When” predictable versus unpredictable	Supramarginal gyrus	180	< 0.001	26.85	5.03
		310	0.004	26.25	4.97
	Rostral middle frontal gyrus	330	0.017	24.82	4.82
	Precentral gyrus	450	0.009	24.73	4.81
		430	0.012	23.76	4.71
“What” × “When” interaction	Superior temporal gyrus	165	0.001	34.37	5.71
“What” valid versus invalid	Middle frontal gyrus	360–400	0.001	43.64	6.42
	Supramarginal gyrus	380–385	0.006	30.00	5.31
		480	0.027	23.05	4.63

^a“What” and “when” predictability was treated as factors in a flexible factorial ANOVA over single-trial evoked responses. Data show all significant clusters (corrected at a peak-level $p_{FWE} < 0.05$).

observed (thresholded at $p < 0.005$, peak-level; corrected for multiple comparisons at FWE < 0.05 , cluster-level).

In an additional *post hoc* analysis, we tested for the effects of “what” prediction validity, corresponding to the effects of con-

tent predictions identified in the analysis of behavioral accuracy and RT. Given the significant effects on behavior, we have included a regressor coding for log RT, as above. This analysis revealed two clusters of activity: in the FC and supramarginal

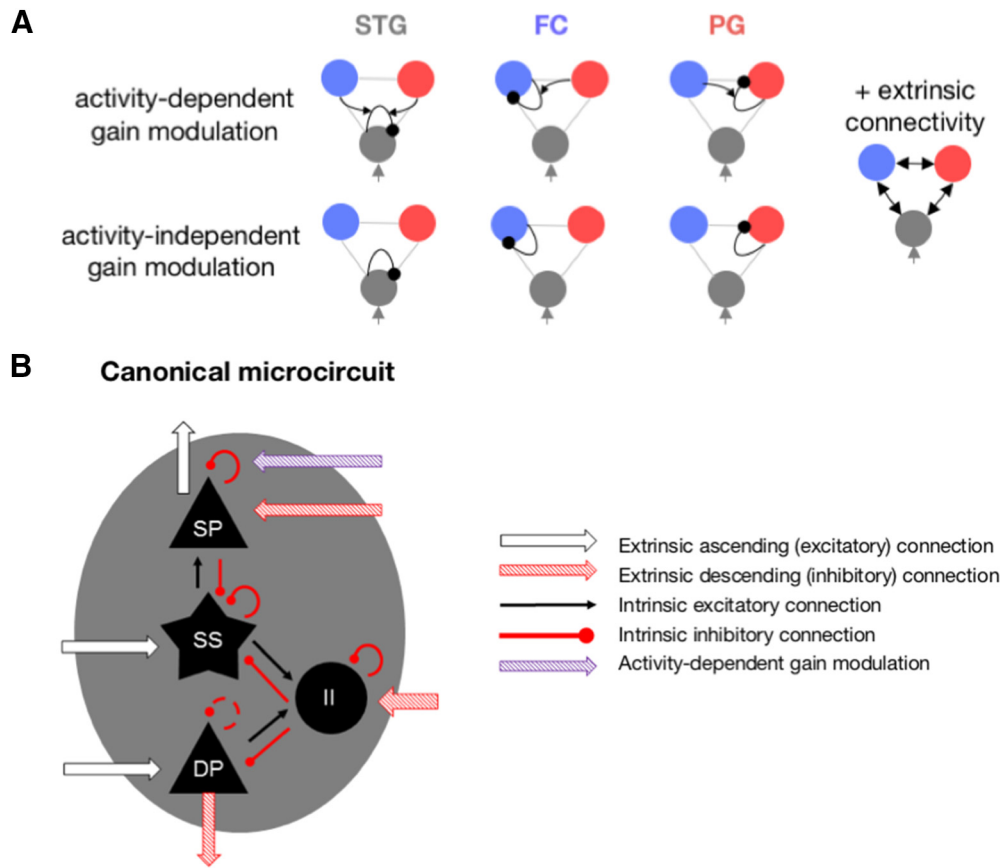


Figure 4. DCM: methods. **A**, Model space. The factorially designed model space allowed disambiguation of hierarchical neuromodulatory effects of “what” and “when” predictability. Alternative models allowed for recurrent connections (self-inhibition) in different neuronal populations (modeling their gain) to be modulated by “what” and/or “when” predictability. Gain was modulated either in an activity-dependent or activity-independent manner. **B**, Each source was modeled using four neuronal populations (SP, Superficial pyramidal cells; DP, deep pyramidal cells; SS, spiny stellate cells; II, inhibitory interneurons) forming a canonical microcircuit. Ascending connections mediate bottom-up flow of auditory activity, whereas descending connections mediate top-down influences of higher regions.

gyrus (Fig. 3G; for test statistics, see Table 2), where mispredicted tones (following invalid cues) were associated with a higher ECoG amplitude than tones that could be correctly predicted (following valid cues). No significant interaction effect between “what” validity and “when” predictability has been observed. However, given a low number of trials in the “what”-invalid condition (14.85 trials on average, after artifact rejection; SD = 3.88 trials), the lack of significant effects might be due to a lower statistical power than in the main analysis of “what” and “when” predictability. For the same reasons, we chose not to model the event-related potentials corresponding to tones following valid and invalid “what” predictions using subsequent DCM, focusing instead on the primary research questions regarding the mechanisms of “what” and “when” predictability.

Predictability effects have also been observed in anticipation of the onset of predictable events (Volosin et al., 2016). Thus, to test whether predictability affects the evoked responses in the prestimulus period, we repeated the analyses with an extended time window, ranging from 250 ms before to 500 ms after syllable onset. To identify prestimulus effects, instead of using the 100 ms baseline period immediately preceding the tone onset, we baseline-corrected the data relative to the 100 ms period preceding the onset of a trial (i.e., the dummy image). No significant clusters of main or interaction effects were found in the prestimulus period, even at a liberal threshold $p < 0.05$ (after correcting at a cluster-level $p_{FWE} < 0.05$), suggesting that baseline activity in our data was not modulated by “what” or “when” predictability.

DCM

So far, we found that predictability affects evoked responses, and that “what” and “when” predictions arise in partially nonoverlapping brain networks and at different latencies. To better understand the neurophysiological mechanism by which “what” and “when” predictions alter neural activity, we applied DCM to the evoked responses in each participant. DCM rests upon significant physiological effects to model their underlying activity, and reproduces observed electrophysiological signals using a forward or generative model using biophysically realistic neuronal architecture (here: neural mass models with each source modeled as a canonical microcircuit; Fig. 4B), and Bayesian statistics. A major advantage of DCM is that it allows inferences regarding putative neuronal mechanisms (i.e., effective connectivity and gain modulation from observed electrophysiological responses), as validated in previous studies (e.g., Garrido et al., 2007; Moran et al., 2011; Papadopoulos et al., 2015; Phillips et al., 2016) (for more details, see Materials and Methods).

For modeling analyses, we selected electrodes in regions identified at the group level, focusing on the most robust effects: that is, STG, PG, and FC (including inferior and middle frontal gyri), respectively (Fig. 4). The STG region was selected based on the significant interaction effect of “what” and “when” predictability at the group level. The PG region incorporated the early and late effects of “when” predictability found both in the PG per se and in the supramarginal gyrus (with the two clusters overlapping in most participants; Fig. 3F). Similarly, the FC region incorporated

the late effects of both “what” and “when” predictability in inferior and middle frontal gyri (overlapping in most participants). Thus, in all three regions modeled by the DCM, the evoked responses were significantly modulated by “what” and/or “when” predictability, which allowed us to obtain a coherent model space for the entire participant sample. We performed subsequent modeling steps on data from individual participants to account for intersubject variability.

We evaluated and compared several alternative models to explain cortical responses to auditory stimuli and their modulation by predictions using a Bayesian approach (see Materials and Methods). Our model space (Fig. 4A) allowed disambiguation of hierarchical neuromodulatory effects of “what” and “when” predictability in a factorial fashion. Specifically, models allowed for gain in different regions (modeled as recurrent self-inhibitory connections of their superficial pyramidal populations) to be independently modulated by “what” and/or “when” predictability. Furthermore, gain could occur either in an activity-independent manner (modeling classical neuromodulatory, e.g., dopaminergic effects) or in an activity-dependent manner (modeling NMDA-mediated short-term plasticity). By virtue of its factorial design, the model space included null models in which different subsets of connections were fixed without allowing “what” and/or “when” predictability to modulate their strength. This way, we could arbitrate among different neural mechanisms: that is, top-down-dependent (NMDA-mediated) plasticity, and classical, unspecific, neuromodulatory gain control, mediating predictability of different stimulus attributes (Coull et al., 2011, 2012; Narayanan et al., 2012; Parker et al., 2013). Instead of selecting a single winning model, we used random-effects BMA (compare Penny et al., 2010) (see Materials and Methods) to infer the posterior parameters explaining the observed neural effects. BMA uses the entire model space by assigning a weight to the parameters of each model according to its log model evidence. Parameters were considered significant when different from baseline with >99.9% posterior probability.

We found that the electrophysiological profile of responses for “what” and “when” predictions were best explained by distinct changes in gain (Fig. 5A; Table 3): “what” predictability was best explained by increases in gain in an activity-dependent manner in STG (reflecting putative short-term plasticity). In contrast, “when” predictability appeared to augment gain in an activity-independent manner (reflective putative classical neuromodulation) in motor areas and in sensory areas (STG). The distinct effects of activity-dependent and activity-independent gain modulation on simulated evoked responses are shown in Figure 5B. These simulations suggest that the effects of activity-dependent STG gain modulation are visible in all three regions in the network (i.e., STG, PG, and FC), whereas the effects of activity-independent gain modulation are more local and largely constrained to the region whose gain is being modulated. The network-wide effect of activity-dependent STG gain modulation can be explained by a positive feedback mechanism, whereby a small initial increase in gain weighted by inputs from higher regions increases the efferent output to these regions and subsequent descending input. This process may further amplify the activity-dependent gain modulation. Because the descending inputs are both modulatory (increasing the gain) and inhibitory (Fig. 4B), the STG itself might not be as sensitive to activity-dependent gain modulation as the areas downstream. Activity-independent gain modulation, on the other hand, appears to primarily affect the output of the source region. The secondary effects on other regions are weaker, appear later in time, and are

overall better explained by modulations in specific connections between regions (see below). Together, this result suggests that, whereas both “what” and “when” predictability modulates neural gain, they engage qualitatively distinct neurophysiological mechanisms (i.e., activity-dependent vs activity-independent gain operating across different cortical regions).

We also investigated whether effective connectivity between areas is modulated through similar or different mechanisms for “what” and “when” predictions. The DCM revealed that the effects of “what” and “when” predictions on the connections between regions were also dissociable: inputs from the STG had more influence on FC activity following stimuli with unpredictable content (“what” predictions). In contrast, inputs from the STG had more influence on PG activity following stimuli with unpredictable onset (“when” predictions). Also, “what” predictions increased the effective connectivity between sensory and precentral regions. Simulations of the network-wide effects of modulating connections between regions indicated that they have a high degree of regional specificity and primarily affect activity in the target regions (Fig. 5B). Model fits showed good correspondence of observed and simulated data for peak amplitudes and latencies in addition to condition-specific effects (Fig. 5C). The optimized models explained on average 92.5% (SD 6.6%) of single subjects’ variance in ECoG data across cortical regions and experimental conditions.

Discussion

The goal of our study was to establish whether predictability of independent aspects of the environment (i.e., the contents and timing of auditory stimuli) is mediated by shared or dissociable neural mechanisms. Our results demonstrate that predictions had a sizable impact on behavior: “what” predictability of syllables led to shorter RTs in the syllable categorization task. Additionally, “when” predictability speeded up categorization of syllables when their contents were predictable (Fig. 1). This modulatory behavioral effect of temporal predictability is consistent with previous findings indicating that temporal predictability more likely influences behavior when combined with other sources of predictability (e.g., with spatial orienting) (Rohenkohl et al., 2014; but see Lau and Nguyen, 2015) or predictable spectral content of acoustic stimulation (Hsu et al., 2013).

At the neuronal level, “what” and “when” predictability had both interactive and dissociable effects. At relatively early latencies (165 ms), “what” and “when” predictions jointly modulated the amplitude of the auditory-evoked response (Fig. 3A) in posterior STG, specifically over locations most sensitive to auditory inputs (Fig. 3D). This suggests that sensory processing of stimuli is modulated based on whether their contents and their timing are predictable, extending previous finding showing synergistic effects of temporal and spatial predictability on visually evoked potentials (Doherty et al., 2005) and of temporal and content-based predictability on auditory omission responses (SanMiguel et al., 2013). Because both behavioral (Rohenkohl et al., 2014) and early evoked neural responses (Doherty et al., 2005; SanMiguel et al., 2013) in unimodal visual and auditory paradigms bring converging evidence for modulatory effects of temporal predictability, it is unlikely that the lack of a main effect of temporal predictability on behavioral or early evoked responses in our study were due to the cross-modal character of “what” predictability. Later responses (e.g., the N400), in contrast, may be more susceptible to additive effects, as suggested by a study combining temporal and semantic predictions (Lau and Nguyen, 2015).

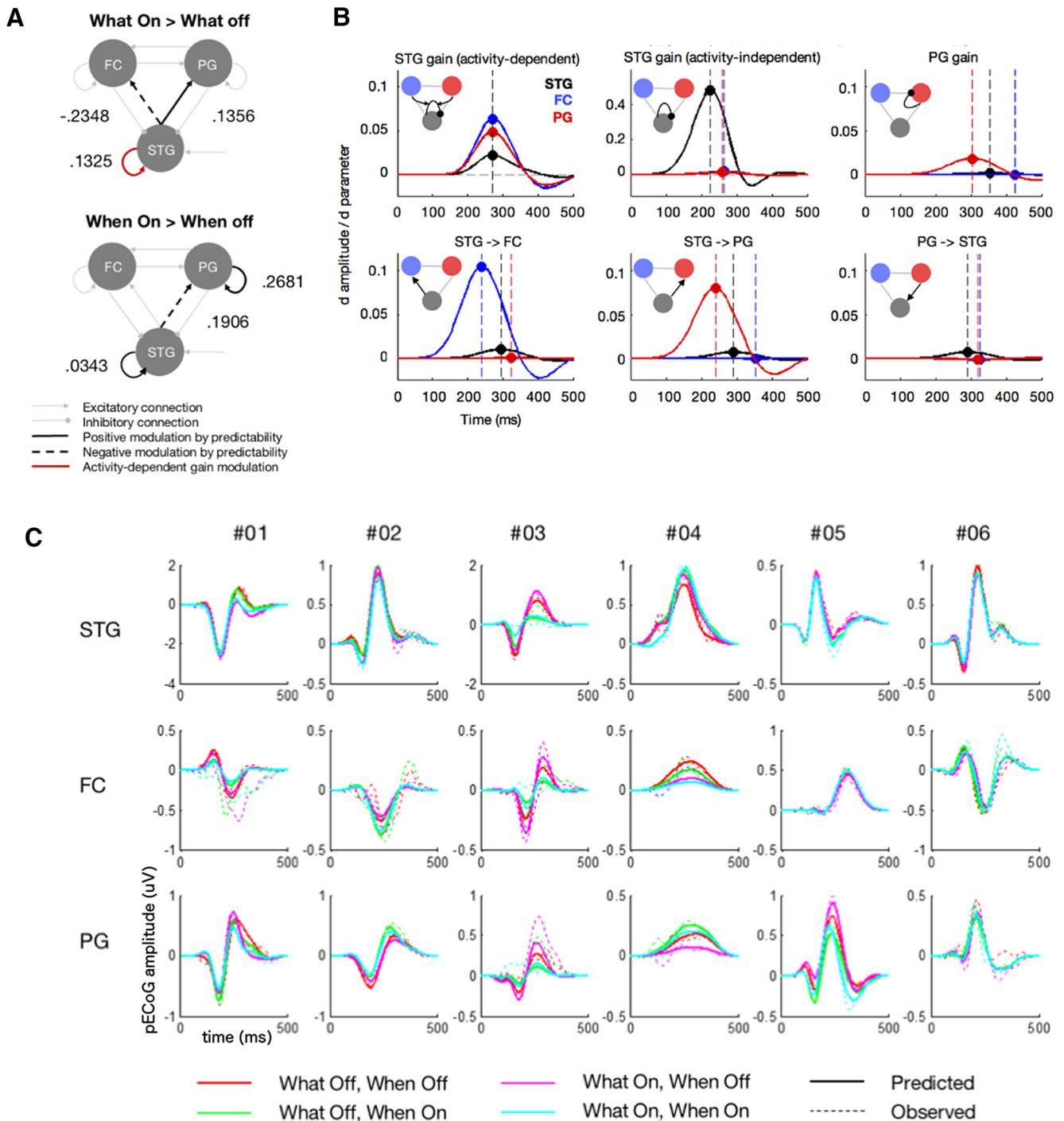


Figure 5. DCM: results. **A**, Parameters of the optimized model, revealing dissociable effects of “what” and “when” expectation on intrinsic and extrinsic connectivity. Values displayed next to connections indicate modulatory parameter; estimates are on a logarithmic scale, relative to baseline (e.g., the STG self-connection increase by 0.1325 under “what” predictability is equivalent to the exponential of 0.1325 = 14.17% gain modulation for “what” predictable vs unpredictable stimuli). **B**, Parameter-specific simulations showing different effects of intrinsic (activity-dependent and -independent gain modulation) and extrinsic (connections between regions) parameters on predicted evoked signal. Each panel represents the simulated effect of changes in a single parameter estimate (connection strength; “*d* parameter”) on changes in source activity amplitude (“*d* amplitude”). Insets, Connection being modulated. Dashed vertical lines indicate peak latency per region. These simulations show that different connections have specific influence on network activity and that two types of gain modulation (activity-dependent and -independent) have distinct effects across the network. **C**, Model fits per participant and condition. Dotted lines indicate individual participants’ evoked responses.

“What” and “when” predictions also had dissociable effects: whereas “what” predictability increased amplitudes of late (420–460 ms) evoked responses over PFC (inferior and middle frontal gyri) only (Fig. 3B), “when” predictability modulated evoked activity over both motor regions and the middle frontal gyrus at both earlier (~180 ms) and later (430–450 ms) latencies (Fig.

3C). These results suggest that predictability of different stimulus features might exert their effect at different stages of processing. While the different effect latencies for “what” and “when” predictability might to some extent be explained by differences in our experimental manipulations (i.e., “when” predictability could be established earlier in a trial, after the first couple of

Table 3. DCM parameters^a

Modulatory factor	Parameter name	Description	Posterior mean	Posterior variance
"What" predictability	N{1}(1,1)	Activity-dependent STG gain modulation	0.1325*	0.0301
	B{1}(2,1)	STG->FC connection	-0.2348*	0.0665
	B{1}(3,1)	STG->PG connection	0.1356*	0.0434
	B{1}(1,2)	FC->STG connection	-0.0484	0.0517
	B{1}(2,2)	Activity-independent FC gain modulation	0.1352	0.0612
	B{1}(1,3)	PG->STG connection	0.1494	0.0702
	B{1}(3,3)	Activity-independent PG gain modulation	-0.1412	0.0469
"When" predictability	B{2}(1,1)	Activity-independent STG gain modulation	0.0343*	0.0105
	B{2}(2,1)	STG->FC connection	-0.1270	0.0643
	B{2}(3,1)	STG->PG connection	-0.1906*	0.0558
	B{2}(1,2)	FC->STG connection	0.1848	0.0551
	B{2}(2,2)	Activity-independent FC gain modulation	0.0659	0.0489
	B{2}(1,3)	PG->STG connection	0.0358	0.0847
	B{2}(3,3)	Activity-independent PG gain modulation	0.2681*	0.0446

^aAll parameters had a prior mean of 0 and prior log variance of 0.125.

*Posterior probability >99.9%.

images, and in a deterministic fashion, whereas "what" predictability could only be established after the final face image and in a probabilistic fashion), a similar pattern of results has also been reported by Hsu et al. (2013), who used a probabilistic manipulation of temporal expectations and a deterministic manipulation of spectral acoustic expectations and found that "when" predictability affects event-related potential amplitudes at earlier latencies than "what" predictability. In other studies of (rhythmic) temporal expectations, "when" predictability has previously been linked to motor cortical activity (Schubotz et al., 2000; Saleh et al., 2010; Cravo et al., 2011; Coull et al., 2012; Morillon et al., 2015), stressing the importance of anticipatory motor activity in encoding temporal expectations. It is worth noting that, in our study, even late activity (430–450 ms) in motor areas was not explained by trial-by-trial RTs (Fig. 3E,F), and, on average, preceded the participants' responses (Fig. 1B), suggesting that modulations in motor cortex might reflect inherent encoding of predictions by motor cortex (Morillon et al., 2015) as opposed to preparation and execution of responses (e.g., the P300 and the lateralized readiness potential), which can also be modulated by target onset probability (Müller-Gethmann et al., 2003; Los and Helsenfeld, 2005; Hackley et al., 2007). "When" predictability also modulated prefrontal activity ~330 ms after stimulus onset. In our study, prefrontal activity (in the middle frontal gyrus) was also modulated by content predictability (Fig. 3C), suggesting a more general role in encoding high-level contextual expectation (Dürschmid et al., 2016; Phillips et al., 2016) as opposed to a specific involvement of the PFC in temporal predictability (Triviño et al., 2010; Phillips et al., 2015).

These results also fit well with notions of distinct auditory ventral and dorsal pathways (for reviews, see Recanzone and Cohen, 2010; Arnott and Alain, 2011), whereby the anteroventral pathway, including the ventral PFC, is prominently involved in sound identity processing ("what"), whereas the posterodorsal pathway likely mediates processing of other (e.g., spatial, "where") auditory attributes (Tian et al., 2001). Interestingly, the latter study suggests that the segregation is not complete, as a subset of neurons in CL (in the posterior STG) was highly selective for both content and space. Studies in monkeys have further revealed that the dorsal stream is also more selective for temporal ("when") features of sounds (Camalier et al., 2012; Kusmierek and Rauschecker, 2014), consistent with postulates of a dorsal "when" pathway (e.g., Schubotz et al., 2003). Whether and how the encoding of "where" and "when" in the dorsal pathway inter-

acts are still debated. Rauschecker and Scott (2009), for instance, have proposed that the dorsal stream constitutes a pathway for sensorimotor integration and control that subsumes "where," "how," and "when" functions of auditory processing. In the same context, our finding indicating modulations of the inferior FC to "what" predictability is consistent with previous reports that the ventrolateral PFC is the endpoint of the auditory ventral stream (Rauschecker and Scott, 2009).

Beyond the cortical and latency differences observed in the evoked responses to "what" and "when" predictions, our modeling results suggest that "what" and "when" predictions tap onto qualitatively distinct gain modulating mechanisms operating in different cortical regions (Fig. 5A): "what" predictability increased the gain in an activity-dependent manner in STG, consistent with an involvement of NMDA-dependent sensory short-term plasticity in associative learning. Thus, our modeling results support the hypothesis that "what" predictions are mediated by voltage-dependent NMDA signaling (Xia et al., 2005), leading to a local disinhibition (Letzkus et al., 2015) of principal cells encoding the prediction errors of stimulus contents, when the respective predictions are violated. In contrast, "when" predictability augmented the gain in an activity-independent manner in motor and sensory regions, consistent with classical neuromodulatory (e.g., dopaminergic) effects (Narayanan et al., 2012; compare Coull et al., 2012; Parker et al., 2013). These findings, in line with previous studies (Doherty et al., 2005; San-Miguel et al., 2013; Rohenkohl et al., 2014), support the hypothesis that temporal predictability has a marked modulatory effect. Our modeling results suggest that rhythmic temporal predictability increases the sensorimotor gain, likely mediated by voltage-independent mechanisms (Formenti et al., 1998; Gorelova et al., 2002) and thus unspecific to inputs received from hierarchically higher regions. This is consistent with other models according to which motor and sensory regions entrain to the rhythmic structure of the environment and dynamically adjust the gain to any incoming stimuli expected within a given time window (Lakatos et al., 2013; Morillon et al., 2015). If no "what" prediction can be established, the gain effect induced by temporal predictability will be nonselective, and as a result may not translate into a behavioral benefit. However, in the presence of "what" predictions, this dynamic gain modulation will additionally boost the responsiveness of neurons whose activity-dependent gain has been amplified by "what" predictions, mediated by descending connections and short-term plasticity in sensory re-

gions. Thus, given strong “what” predictions, the temporal gain effect might be more readily observed at the behavioral level. Our results may also explain the absence of effects of “what” or “when” predictions in the prestimulus window, such as the contingent negative variation, occasionally measured using ECoG (Hamano et al., 1997) and more typically observed in MEG/EEG studies as a marker of stimulus prediction (e.g., Chennu et al., 2013; Breska and Deouell, 2017): in our study, while gain modulation may have plausibly occurred already before target onset, the prestimulus window was silent (lacking any stimulation); thus, the effects of increased gain may best be revealed in the response evoked to the target stimulus.

The effects of “what” and “when” predictability on connectivity between regions were also dissociable. Stimuli with unpredictable contents were linked to increased excitability in frontal region to ascending drive from the STG, whereas stimuli with unpredictable onsets were associated with increased excitability in motor region to ascending inputs from the STG. These results indicate that “what” prediction errors (Friston, 2005), likely signaled in the unpredictable condition, propagate primarily to prefrontal regions, whereas “when” prediction error signaling relies on sensorimotor processing, likely involving indirect subcortical or cortical connections linking auditory and motor cortices (Morillon et al., 2015). Content predictability also increased the strength of connections from sensory to precentral regions, possibly facilitating transmission of information to downstream areas (Engel et al., 2001).

While in our study both kinds of predictability were linked to decreased RTs, increased amplitude of early evoked responses, and plausibly explained as mediated by different kinds of synaptic gain control, previous studies suggest that similar facilitating effects of predictability might be mediated by attention (Lakatos et al., 2013; Zhao et al., 2013). In our paradigm, for instance, “when” predictability can result in more efficient temporal orienting to specific intervals. However, it has been recently shown that the effects of predictability itself can be modulated by top-down attention to specific features (e.g., temporal information) that are relevant in a given context (Aukstulewicz et al., 2017). Although in the current paradigm we did not explicitly manipulate top-down attention, it has been previously shown that attention can be linked to sensory gain modulation (Hillyard et al., 1998; Gould et al., 2011; Aukstulewicz and Friston, 2015). In the study of Aukstulewicz and Friston (2015), for instance, attention was defined as task relevance and manipulated orthogonally to predictability defined as the difference between repeated and novel stimuli. In contrast, the key comparison in our study was between predictable (75% “standards”) and unpredictable syllables for which no content-based expectation could be formed. Thus, our experimental manipulation (“predictable” vs “unpredictable”) was of contextual (second-order) predictions, likely mediated by gain control modulation (Summerfield and de Lange, 2014; Kanai et al., 2015; Aukstulewicz and Friston, 2016), as opposed to first-order predictions (“predicted” vs “unpredicted”) about a particular stimulus, previously linked to descending connections (Murray et al., 2002; Sanders et al., 2014; Aukstulewicz and Friston, 2015). Although available electrode coverage precluded direct recordings from primary auditory cortex, the presence of modulatory effects in the STG at early latencies (165 ms), in combination with the modeling results indicating that predictability modulate excitability in STG, raise the possibility that contextual expectation, as opposed to attention, modulates the gain at higher levels of auditory processing (Aukstulewicz and Friston, 2016).

Together, our electrophysiological and modeling results suggest that the predictability of different stimulus attributes modulates neural gain at different levels of the processing hierarchy (i.e., sensory vs motor and through activity-dependent vs activity-independent gain control). This factorization of “what” and “when” processing might endow the system with flexibility to combine and segregate the effects of different kinds of predictability in the dynamic modulation of sensory processing, such that independent sources of information can be weighted by their reliability, thereby optimizing processing.

References

- Alink A, Schwiedrzik CM, Kohler A, Singer W, Muckli L (2010) Stimulus predictability reduces responses in primary visual cortex. *J Neurosci* 30:2960–2966. [CrossRef Medline](#)
- Arnal LH, Giraud AL (2012) Cortical oscillations and sensory predictions. *Trends Cogn Sci* 16:390–398. [CrossRef Medline](#)
- Arnott SR, Alain C (2011) The auditory dorsal pathway: orienting vision. *Neurosci Biobehav Rev* 35:2162–2173. [CrossRef Medline](#)
- Aukstulewicz R, Friston K (2015) Attentional enhancement of auditory mismatch responses: a DCM/MEG study. *Cereb Cortex* 25:4273–4283. [CrossRef Medline](#)
- Aukstulewicz R, Friston K (2016) Repetition suppression and its contextual determinants in predictive coding. *Cortex* 80:125–140. [CrossRef Medline](#)
- Aukstulewicz R, Friston KJ, Nobre AC (2017) Task relevance modulates the behavioural and neural effects of sensory predictions. *PLoS Biol* 15:e2003143. [CrossRef Medline](#)
- Baldeweg T (2007) ERP repetition effects and mismatch negativity generation: a predictive coding perspective. *J Psychophysiol* 21:204–213. [CrossRef](#)
- Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ (2012) Canonical microcircuits for predictive coding. *Neuron* 76:695–711. [CrossRef Medline](#)
- Besle J, Schevon CA, Mehta AD, Lakatos P, Goodman RR, McKhann GM, Emerson RG, Schroeder CE (2011) Tuning of the human neocortex to the temporal dynamics of attended events. *J Neurosci* 31:3176–3185. [CrossRef Medline](#)
- Boatman DF, Miglioretti DL (2005) Cortical sites critical for speech discrimination in normal and impaired listeners. *J Neurosci* 25:5475–5480. [CrossRef Medline](#)
- Bornkessel-Schlesewsky I, Schlesewsky M, Small SL, Rauschecker JP (2015) Neurobiological roots of language in primate audition: common computational properties. *Trends Cogn Sci* 19:142–150. [CrossRef Medline](#)
- Breska A, Deouell LY (2017) Neural mechanisms of rhythm-based temporal prediction: delta phase-locking reflects temporal predictability but not rhythmic entrainment. *PLoS Biol* 15:e2001665. [CrossRef Medline](#)
- Brown HR, Friston KJ (2012) Dynamic causal modelling of precision and synaptic gain in visual perception: an EEG study. *Neuroimage* 63:223–231. [CrossRef Medline](#)
- Camalier CR, D’Angelo WR, Sterbing-D’Angelo SJ, de la Mothe LA, Hackett TA (2012) Neural latencies across auditory cortex of macaque support a dorsal stream supramodal timing advantage in primates. *Proc Natl Acad Sci U S A* 109:18168–18173. [CrossRef Medline](#)
- Chennu S, Noreika V, Gueorguiev D, Blenkmann A, Kochen S, Ibáñez A, Owen AM, Bekinschtein TA (2013) Expectation and attention in hierarchical auditory prediction. *J Neurosci* 33:11194–11205. [CrossRef Medline](#)
- Chikkerur S, Serre T, Tan C, Poggio T (2010) What and where: a Bayesian inference theory of attention. *Vision Res* 50:2233–2247. [CrossRef Medline](#)
- Coull JT, Cheng RK, Meck WH (2011) Neuroanatomical and neurochemical substrates of timing. *Neuropsychopharmacology* 36:3–25. [CrossRef Medline](#)
- Coull JT, Hwang HJ, Leyton M, Dagher A (2012) Dopamine precursor depletion impairs timing in healthy volunteers by attenuating activity in putamen and supplementary motor area. *J Neurosci* 32:16704–16715. [CrossRef Medline](#)
- Cravo AM, Rohenkohl G, Wyart V, Nobre AC (2011) Endogenous modulation of low frequency oscillations by temporal expectations. *J Neurophysiol* 106:2964–2972. [CrossRef Medline](#)

- Cravo AM, Rohenkohl G, Wyart V, Nobre AC (2013) Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex. *J Neurosci* 33:4002–4010. [CrossRef Medline](#)
- den Ouden HE, Friston KJ, Daw ND, McIntosh AR, Stephan KE (2009) A dual role for prediction error in associative learning. *Cereb Cortex* 19:1175–1185. [CrossRef Medline](#)
- Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany RJ (2006) An automated labelling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31:968–980. [CrossRef Medline](#)
- Doherty JR, Rao A, Mesulam MM, Nobre AC (2005) Synergistic effect of combined temporal and spatial expectations on visual attention. *J Neurosci* 25:8259–8266. [CrossRef Medline](#)
- Dürschmid S, Edwards E, Reichert C, Dewar C, Hinrichs H, Heinze HJ, Kirsch HE, Dalal SS, Deouell LY, Knight RT (2016) Hierarchy of prediction errors for auditory events in human temporal and frontal cortex. *Proc Natl Acad Sci U S A* 113:6755–6760. [CrossRef Medline](#)
- Engel AK, Fries P, Singer W (2001) Dynamic predictions: oscillations and synchrony in top-down processing. *Nat Rev Neurosci* 2(10):704–716.
- Fischl B, van der Kouwe A, Destrieux C, Halgren E, Ségonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy D, Caviness V, Makris N, Rosen B, Dale AM (2004) Automatically parcellating the human cerebral cortex. *Cereb Cortex* 14:11–22. [CrossRef Medline](#)
- Formenti A, Martina M, Plebani A, Mancia M (1998) Multiple modulatory effects of dopamine on calcium channel kinetics in adult rat sensory neurons. *J Physiol* 509:395–409. [CrossRef Medline](#)
- Friston K (2005) A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci* 360:815–836. [CrossRef Medline](#)
- Friston K (2010) The free-energy principle: a unified brain theory? *Nat Rev Neurosci* 11:127–138. [CrossRef Medline](#)
- Friston K, Buzsáki G (2016) The functional anatomy of time: what and when in the brain. *Trends Cogn Sci* 20:500–511. [CrossRef Medline](#)
- Friston K, Mattout J, Trujillo-Barreto N, Ashburner J, Penny W (2007) Variational free energy and the Laplace approximation. *Neuroimage* 34:220–234. [CrossRef Medline](#)
- Garrido MI, Kilner JM, Kiebel SJ, Stephan KE, Friston KJ (2007) Dynamic causal modelling of evoked potentials: a reproducibility study. *Neuroimage* 36:571–580. [CrossRef Medline](#)
- Gorelova N, Seamans JK, Yang CR (2002) Mechanisms of dopamine activation of fast-spiking interneurons that exert inhibition in rat prefrontal cortex. *J Neurophysiol* 88:3150–3166. [CrossRef Medline](#)
- Gould IC, Rushworth MF, Nobre AC (2011) Indexing the graded allocation of visuospatial attention using anticipatory alpha oscillations. *J Neurophysiol* 105:1318–1326. [CrossRef Medline](#)
- Griffin IC, Miniussi C, Nobre AC (2002) Multiple mechanisms of selective attention: differential modulation of stimulus processing by attention to space or time. *Neuropsychologia* 40:2325–2340. [CrossRef Medline](#)
- Hackley SA, Schankin A, Wohlschlaeger A, Wascher E (2007) Localization of temporal preparation effects via trisected reaction time. *Psychophysiology* 44:334–338. [CrossRef Medline](#)
- Hamano T, Lüders HO, Ikeda A, Collura TF, Comair YG, Shibasaki H (1997) The cortical generators of the contingent negative variation in humans: a study with subdural electrodes. *Electroencephalogr Clin Neurophysiol* 104:257–268. [CrossRef Medline](#)
- Hillyard SA, Vogel EK, Luck SJ (1998) Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philos Trans R Soc Lond B Biol Sci* 29:1257–1270. [CrossRef Medline](#)
- Hsu YF, Hämäläinen JA, Waszak F (2013) Temporal expectation and spectral expectation operate in distinct fashion on neuronal populations. *Neuropsychologia* 51:2548–2555. [CrossRef Medline](#)
- Iglesias S, Mathys C, Brodersen KH, Kasper L, Piccirelli M, den Ouden HE, Stephan KE (2013) Hierarchical prediction errors in midbrain and basal forebrain during sensory learning. *Neuron* 80:519–530. [CrossRef Medline](#)
- Jiang J, Schmajuk N, Egnér T (2012) Explaining neural signals in human visual cortex with an associative learning model. *Behav Neurosci* 126:575–581. [CrossRef Medline](#)
- Kanai R, Komura Y, Shipp S, Friston K (2015) Cerebral hierarchies: predictive processing, precision and the pulvinar. *Philos Trans R Soc Lond B Biol Sci* 370:1668. [CrossRef Medline](#)
- Kiebel SJ, Daunizeau J, Friston KJ (2008) A hierarchy of time-scales and the brain. *PLoS Comp Biol* 4:e1000209. [CrossRef Medline](#)
- Kilner JM, Kiebel SJ, Friston KJ (2005) Applications of random field theory to electrophysiology. *Neurosci Lett* 374:174–178. [CrossRef Medline](#)
- Knill DC, Pouget A (2004) The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci* 27:712–719. [CrossRef Medline](#)
- Kusmirek P, Rauschecker JP (2014) Selectivity for space and time in early areas of the auditory dorsal stream in the rhesus monkey. *J Neurophysiol* 111:1671–1685. [CrossRef Medline](#)
- Lachaux JP, Axmacher N, Mormann F, Halgren E, Crone NE (2012) High-frequency neural activity and human cognition: past, present and possible future of intracranial EEG research. *Prog Neurobiol* 98:279–301. [CrossRef Medline](#)
- Lakatos P, Musacchia G, O’Connell MN, Falchier AY, Javitt DC, Schroeder CE (2013) The spectrotemporal filter mechanism of auditory selective attention. *Neuron* 77:750–761. [CrossRef Medline](#)
- Lau EF, Nguyen E (2015) The role of temporal predictability in semantic expectation: an MEG investigation. *Cortex* 68:8–19. [CrossRef Medline](#)
- Leaver AM, Van Lare J, Zielinski B, Halpern AR, Rauschecker JP (2009) Brain activation during anticipation of sound sequences. *J Neurosci* 29:2477–2485. [CrossRef Medline](#)
- Letzkus JJ, Wolff SB, Lüthi A (2015) Disinhibition, a circuit mechanism for associative learning and memory. *Neuron* 88:264–276. [CrossRef Medline](#)
- Lindén H, Tetzlaff T, Potjans TC, Pettersen KH, Grün S, Diesmann M, Einevoll GT (2011) Modeling the spatial reach of the LFP. *Neuron* 72:859–872. [CrossRef Medline](#)
- Los SA, Helsenfeld DJ (2005) Intentional and unintentional contributions to nonspecific preparation: electrophysiological evidence. *J Exp Psychol Gen* 134:52–72. [CrossRef Medline](#)
- Luft CD, Meeson A, Welchman AE, Kourtzi Z (2015) Decoding the future from past experience: learning shapes predictions in early visual cortex. *J Neurophysiol* 113:3159–3171. [CrossRef Medline](#)
- Marshall L, Mathys C, Ruge D, de Berker AO, Dayan P, Stephan KE, Bestmann S (2016) Pharmacological fingerprints of contextual uncertainty. *PLoS Biol* 14:e1002575. [CrossRef Medline](#)
- McGann JP (2015) Associative learning and sensory neuroplasticity: how does it happen and what is it good for? *Learn Mem* 22:567–576. [CrossRef Medline](#)
- Mento G, Tarantino V, Vallesi A, Bisiacchi PS (2015) Spatiotemporal neurodynamics underlying internally and externally driven temporal prediction: a high spatial resolution ERP study. *J Cogn Neurosci* 27:425–439. [CrossRef Medline](#)
- Moran RJ, Jung F, Kumagai T, Endepols H, Graf R, Dolan RJ, Friston KJ, Stephan KE, Tittgemeyer M (2011) Dynamic causal models and physiological inference: a validation study using isoflurane anaesthesia in rodents. *PLoS One* 6:e22790. [CrossRef Medline](#)
- Moran RJ, Campo P, Symmonds M, Stephan KE, Dolan RJ, Friston KJ (2013) Free energy, precision and learning: the role of cholinergic neuromodulation. *J Neurosci* 33:8227–8236. [CrossRef Medline](#)
- Morillon B, Hackett TA, Kajikawa Y, Schroeder CE (2015) Predictive motor control of sensory dynamics in auditory active sensing. *Curr Opin Neurobiol* 31:230–238. [CrossRef Medline](#)
- Morillon B, Schroeder CE, Wyart V, Arnal LH (2016) Temporal prediction in lieu of periodic stimulation. *J Neurosci* 36:2342–2347. [CrossRef Medline](#)
- Müller-Gethmann H, Ulrich R, Rinkenauer G (2003) Locus of the effect of temporal preparation: evidence from the lateralized readiness potential. *Psychophysiology* 40:597–611. [CrossRef Medline](#)
- Murray SO, Kersten D, Olshausen BA, Schrater P, Woods DL (2002) Shape perception reduces activity in human primary visual cortex. *Proc Natl Acad Sci U S A* 99:15164–15169. [CrossRef Medline](#)
- Narayanan NS, Land BB, Solder JE, Deisseroth K, DiLeone RJ (2012) Prefrontal D1 dopamine signalling is required for temporal control. *Proc Natl Acad Sci U S A* 109:20726–20731. [CrossRef Medline](#)
- Nourski KV, Steinschneider M, Oya H, Kawasaki H, Jones RD, Howard MA (2014) Spectral organization of the human lateral superior temporal gyrus revealed by intracranial recordings. *Cereb Cortex* 24:340–352. [CrossRef Medline](#)
- Papadopoulou M, Leite M, van Mierlo P, Vonck K, Lemieux L, Friston K, Marinazzo D (2015) Tracking slow modulations in synaptic gain using

- dynamic causal modelling: validation in epilepsy. *Neuroimage* 15:117–126. [CrossRef Medline](#)
- Parker KL, Alberico SL, Miller AD, Narayanan NS (2013) Prefrontal D1 dopamine signalling is necessary for temporal expectation during reaction time performance. *Neuroscience* 255:246–254. [CrossRef Medline](#)
- Penny WD, Stephan KE, Daunizeau J, Rosa MJ, Friston KJ, Schofield TM, Leff AP (2010) Comparing families of dynamic causal models. *PLoS Comput Biol* 6:e1000709. [CrossRef Medline](#)
- Phillips HN, Blenkmann A, Hughes LE, Bekinschtein TA, Rowe JB (2015) Hierarchical organization of frontotemporal networks for the prediction of stimuli across multiple dimensions. *J Neurosci* 35:9255–9264. [CrossRef Medline](#)
- Phillips HN, Blenkmann A, Hughes LE, Kochen S, Bekinschtein TA, Cam-Can, Rowe JB (2016) Convergent evidence for hierarchical prediction networks from human electrocorticography and magnetoencephalography. *Cortex* 82:192–205. [CrossRef Medline](#)
- Pinotsis DA, Schwarzkopf DS, Litvak V, Rees G, Barnes G, Friston KJ (2012) Dynamic causal modelling of lateral interactions in the visual cortex. *Neuroimage* 66C:563–576. [CrossRef Medline](#)
- Rauschecker JP (2012) Ventral and dorsal streams in the evolution of speech and language. *Front Evol Neurosci* 4:7. [CrossRef Medline](#)
- Rauschecker JP, Scott SK (2009) Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci* 12:718–724. [CrossRef Medline](#)
- Recanzone GH, Cohen YE (2010) Serial and parallel processing in the primate auditory cortex revisited. *Behav Brain Res* 206:1–7. [CrossRef Medline](#)
- Rohenkohl G, Cravo AM, Wyart V, Nobre AC (2012) Temporal expectation improves the quality of sensory information. *J Neurosci* 32:8424–8428. [CrossRef Medline](#)
- Rohenkohl G, Gould IC, Pessoa J, Nobre AC (2014) Combining spatial and temporal expectations to improve visual perception. *J Vis* 14:8. [CrossRef Medline](#)
- Rosier AM, Arckens L, Orban GA, Vandesande F (1993) Laminar distribution of NMDA receptors in cat and monkey visual cortex visualized by [³H]-MK-801 binding. *J Comp Neurol* 335:369–380. [CrossRef Medline](#)
- Saleh M, Reimer J, Penn R, Ojakangas CL, Hatsopoulos NG (2010) Fast and slow oscillations in human primary motor cortex predict oncoming behaviourally relevant cues. *Neuron* 65:461–471. [CrossRef Medline](#)
- Sanders LL, Auksztulewicz R, Hohlefeld FU, Busch NA, Sterzer P (2014) The influence of spontaneous brain oscillations on apparent motion perception. *Neuroimage* 102:241–248. [CrossRef Medline](#)
- SanMiguel I, Saupé K, Schröger E (2013) I know what is missing here: electrophysiological prediction error signals elicited by omissions of predicted “what” but not “when.” *Front Hum Neurosci* 7:407. [CrossRef Medline](#)
- Schroeder CE, Lakatos P (2009) Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci* 32:9–18. [CrossRef Medline](#)
- Schubotz RI, Friederici AD, von Cramon DY (2000) Time perception and motor timing: a common cortical and subcortical basis revealed by fMRI. *Neuroimage* 11:1–12. [CrossRef Medline](#)
- Schubotz RI, von Cramon DY, Lohmann G (2003) Auditory what, where, and when: a sensory somatotopy in lateral premotor cortex. *Neuroimage* 20:173–185. [CrossRef Medline](#)
- Schultz W, Dickinson A (2000) Neuronal coding of prediction errors. *Annu Rev Neurosci* 23:473–500. [CrossRef Medline](#)
- Schwiedrzik CM, Freiwald WA (2017) High-level prediction signals in a low-level area of the macaque face-processing hierarchy. *Neuron* 96:89–97.e4. [CrossRef Medline](#)
- Summerfield C, de Lange FP (2014) Expectation in perceptual decision making: neural and computational mechanisms. *Nat Rev Neurosci* 15:745–756. [CrossRef Medline](#)
- Tian B, Reser D, Durham A, Kustov A, Rauschecker JP (2001) Functional specialization in rhesus monkey auditory cortex. *Science* 292:290–293. [CrossRef Medline](#)
- Triviño M, Correa A, Arnedo M, Lupiáñez J (2010) Temporal orienting deficit after prefrontal damage. *Brain* 133:1173–1185. [CrossRef Medline](#)
- Turk-Browne NB, Scholl BJ, Johnson MK, Chun MM (2010) Implicit perceptual anticipation triggered by statistical learning. *J Neurosci* 30:11177–11187. [CrossRef Medline](#)
- Volosin M, Grimm S, Horváth J (2016) Exploiting temporal predictability: event-related potential correlates of task-supportive temporal cue processing in auditory distraction. *Brain Res* 1639:120–131. [CrossRef Medline](#)
- Xia S, Miyashita T, Fu TF, Lin WY, Wu CL, Pyzocha L, Lin IR, Saitoe M, Tully T, Chiang AS (2005) NMDA receptors mediate olfactory learning and memory in *Drosophila*. *Curr Biol* 15:603–615. [CrossRef Medline](#)
- Yang AI, Wang X, Doyle WK, Halgren E, Carlson C, Belcher TL, Cash SS, Devinsky O, Thesen T (2012) Localization of dense intracranial electrode arrays using magnetic resonance imaging. *Neuroimage* 63:157–165. [CrossRef Medline](#)
- Yu AJ, Dayan P (2005) Uncertainty, neuromodulation, and attention. *Neuron* 46:681–692. [CrossRef Medline](#)
- Zeki S, Shipp S (1988) The functional logic of cortical connections. *Nature* 335:311–317. [CrossRef Medline](#)
- Zhao J, Al-Aidroos N, Turk-Browne NB (2013) Attention is spontaneously biased toward regularities. *Psychol Sci* 24:667–677. [CrossRef Medline](#)