Learning to predict: Neuronal signatures of auditory expectancy in human event-related potentials

Yonatan I. Fishman[a,b,*],

Wei-Wei Lee[b],

Elyse Sussman[b]

[a]Department of Neurology, Albert Einstein College of Medicine, United States

[b]Dominick P. Purpura Department of Neuroscience, Albert Einstein College of Medicine, United States

Abstract

Learning to anticipate future states of the world based on statistical regularities in the environment is a key component of perception and is vital for the survival of many organisms. Such statistical learning and prediction are crucial for acquiring language and music appreciation. Importantly, learned expectations can be implicitly derived from exposure to sensory input, without requiring explicit information regarding contingencies in the environment. Whereas many previous studies of statistical learning have demonstrated larger neuronal responses to unexpected versus expected stimuli, the neuronal bases of the expectations themselves remain poorly understood. Here we examined behavioral and neuronal signatures of learned expectancy via human scalp-recorded event-related brain potentials (ERPs). Participants were instructed to listen to a series of sounds and press a response button as quickly as possible upon hearing a target noise burst, which was either reliably or unreliably preceded by one of three pure tones in low-, mid-, and high-frequency ranges. Participants were not informed about the statistical contingencies between the preceding tone 'cues' and the target. Over the course of a stimulus block, participants responded more rapidly to reliably cued targets. This behavioral index of learned expectancy was paralleled by a negative ERP deflection, designated as a neuronal contingency response (CR), which occurred immediately prior to the onset of the target. The amplitude and latency of the CR were systematically modulated by the strength of the predictive relationship between the cue and the target. *Re*-averaging ERPs with respect to the latency of behavioral responses revealed no consistent relationship between the CR and the motor response, suggesting that the CR represents a neuronal signature of learned expectancy or anticipatory attention. Our results demonstrate that statistical regularities in an auditory input stream can be implicitly learned and exploited

to influence behavior. Furthermore, we uncover a potential 'prediction signal' that reflects this fundamental learning process.

1. Introduction

In everyday life we routinely learn that there are contingencies between certain events and form expectations about future events based on this learning. For example, when listening to speech or music, we readily learn the phonotactics of our language and the structure of musical phrases. This learning can generate expectations about what phonemes or musical phrases are expected to come next based on what we have heard (Rohrmeier and Koelsch, 2012; Koelsch et al., 2019; Saffran et al., 1996, 1999; Warker and Dell, 2006; Bonte et al., 2005; Pelucchi et al., 2009). Importantly, we often form these expectations implicitly from exposure to sensory input, without being explicitly informed about the statistical contingencies in the environment. For instance, infants learn word boundaries from the continuous stream of speech they encounter putatively by passive computation of the transitional probabilities between phonemes (Saffran et al., 1999, 1996). Similar implicit learning capabilities are observed in human adults and non-human animals in non-linguistic contexts (Moldwin et al., 2017; Lu and Vicario, 2014; Heimbauer et al., 2018; Schiavo and Froemke, 2019; Saffran et al., 1999). This type of statistical learning, herein called 'learned expectancy', is crucial for the survival of many organisms, enabling them to extract regularities in the environment and thereby anticipate future events. These processes are important for adaptive decision-making (Summerfield and de Lange, 2014).

Despite its critical role in everyday cognition and behavior, the neuronal bases of learned expectancy are still not well understood. Many studies of statistical learning investigate differences in the neuronal responses to expected versus unexpected stimuli with a diversity of results (Auksztulewicz et al., 2017, 2018; Barascud et al., 2016; Southwell et al., 2017; Hsu et al., 2015; Heilbron and Chait, 2018; Symonds et al., 2017; Summerfield and de Lange, 2014; Richter and de Lange, 2019; den Ouden et al., 2012; Garrido et al., 2013; Todorovic et al., 2011). The most general finding of these studies is that unexpected stimuli elicit larger neuronal responses in sensory cortical areas than expected stimuli, consistent with predictive coding models of cortical processing that postulate top-down brain mechanisms which suppress responses to expected stimuli. The larger response to unexpected stimuli is often interpreted as a prediction 'error' signal that has been proposed to reflect the brain's prioritization of unanticipated sensory inputs (Press et al., 2020a; Friston, 2005). These previous studies have largely focused on the brain's increased response to unexpected stimuli as an indicator of surprise, or novelty. However, much less is known about the neuronal underpinnings of the expectations themselves, which emerge from and guide statistical learning (Heilbron and Chait, 2018).

One potential neuronal signature of expectancy is the contingent negative variation (CNV). The CNV is a long-duration event-related brain potential (ERP) that is initiated by a stimulus (S1) when a behavioral response to a second stimulus (S2) is dependent upon maintaining information about S1. A negative-going ERP deflection is observed between S1 and S2 which indicates that a brain process was joining the two stimuli. Studies of the CNV suggest that it is a composite neurophysiological phenomenon which represents, in addition to a motor preparatory response, a neuronal correlate of expectancy for the S2 stimulus that follows the presentation of the S1 stimulus (Mento, 2013; Mento et al., 2015; van Boxtel and Brunia, 1994a, Bickel et al., 2012; Chennu et al., 2013; Tecce, 1972; Donchin et al., 1972; Naatanen and Gaillard, 1974; Hamano et al., 1997; Walter et al., 1964).

Another potential neuronal signature of expectancy has been assessed with omission paradigms in which an expected stimulus is unexpectedly omitted. The unexpected omission elicits a neuronal response that has been localized to sensory cortical areas involved in processing the features of the expected stimulus (Schröger et al., 2015; Kok et al., 2014; Arnal and Giraud, 2012; SanMiguel et al., 2013a; SanMiguel et al., 2013b; Bendixen et al., 2012; Hughes et al., 2001; Wacongne et al., 2011). These omission studies suggest that the brain activates a feature-specific 'template' of the expected stimulus like that evoked by the corresponding actual stimulus. Hence, the omission response has been interpreted by a number of investigators as a 'prediction signal' reflecting prior expectations (Heilbron and Chait, 2018; Kok et al., 2014; Schröger et al., 2015; de Lange et al., 2018; SanMiguel et al., 2013b; Chennu et al., 2016; Demarchi et al., 2019).

A third potential neuronal signature of expectancy has been observed in brain oscillations recorded from the frontal cortex of patients with epilepsy (Durschmid et al., 2019). Durschmid et al., 2019 found a reduction in high-frequency gamma activity originating from frontal cortex immediately prior to the onset of expected sounds. It remains unclear, however, whether and how this reduction in high-frequency activity relates to other putative expectancy signals in lower frequency bands (e.g., theta).

Prior research thus suggests a variety of findings with no real consensus. Moreover, several outstanding issues remain. First, the relationships between behavior and the neuronal correlates of learned expectancy have been studied in only a small number of studies (e.g., Mento, 2017; Coull, 2009; Hillyard, 1969). Second, many studies using the omission paradigm have compared neuronal responses to expected versus unexpected stimulus omissions (Wacongne et al., 2011; Hughes et al., 2001). However, it is ambiguous whether the omission response revealed through this comparison represents a 'pure' prediction signal or alternatively a prediction 'error' signal reflecting the unexpectedness of the omission (Schröger et al., 2015). Third, most previous CNV studies kept features of S1 fixed and did not investigate whether the participants' learned expectations were specific to the sensory attributes of the stimuli (e.g., their shape or pitch). Finally, many previous CNV studies either provided explicit information to participants regarding the behavioral relevance of the predictive association between S1 and S2 (e.g., Perchet and Garcia-Larrea, 2005; Gomez et al., 2019; Arjona et al., 2016; Mento, 2017; Posner, 1980), or it was unclear what specific information was given to participants when instructing them how to perform the behavioral task (cf. Mento et al., 2013). Thus, it remains unresolved whether a neuronal correlate

of expectancy can arise from implicit learning of stimulus contingencies, similar to the statistical learning that is prevalent in everyday life.

The current study was designed to address some of these issues by characterizing behavioral and neuronal signatures of learned expectancy. We used an implicit learning paradigm to study the buildup of expectations that occurs by repetition of an arbitrary association between two sounds. Unlike many previous studies which measured expectancy by the strength of the neuronal response to violations of an expected stimulus, our paradigm does not contain violations of established expectations. Instead, our experimental design is intended to evaluate the neuronal signatures of the expectation itself. In addition, we sought to determine whether the behavioral and neuronal signatures of this learned expectancy would track specific features of the sounds (e.g., their frequency) and reflect the probabilistic relationships between them. To this end, we recorded EEG from the scalp while participants listened to a series of sounds consisting of three pure tones in low-, mid-, and high-frequency ranges which either reliably or unreliably preceded a target noise burst. Participants were instructed to press a response button as quickly as possible upon hearing the target. They were not informed about the statistical contingencies between the tone 'cues' and the target or the behavioral relevance of the tones. We hypothesized that participants would implicitly learn the association between the cues and the target, and that a distinct neuronal response would emerge reflecting the strength of this relationship. Furthermore, we predicted that this implicit learning would facilitate participants' behavioral performance on the target detection task.

2.   Materials and methods

2.1.   Participants

18 healthy adults (10 females; 8 males; all right-handed, ranging in age from 22– 44 years; $M = 28$, SD=6) participated in the study. All participants passed a hearing screening at 20 dB HL or better at 500, 1000, 2000, and 4000 Hz bilaterally and had no reported history of neurological disorders. All procedures were approved by the Institutional Review Board (IRB) of the Albert Einstein College of Medicine. The study was carried out in accordance with the Declaration of Helsinki. Participants gave written consent prior to the study and were paid for their participation.

2.2.   Stimuli and procedures

Stimuli: Fig. 1 provides a schematic diagram of the stimuli. Four stimuli, three pure tones, serving as cues, and one noise burst, serving as the target, were generated and presented using Neuroscan hardware and software (Compumedics, Inc.). All stimuli were 200 ms in duration with 7.5 rise/fall time (Hanning window). Sounds were presented bilaterally via insert earphones at a comfortable listening level of approximately 70 dB SPL. Sound level was calibrated using a Brüel & Kjaer (Naerum, Denmark) sound level meter with an artificial ear. The three pure tones differed only in frequency: 440 Hz (low, A_4), 1244.5 Hz (mid, E^b_5), and 2793 Hz (high, F_7) and were presented with equal probability (33%) in random order. The stimulus onset asynchrony (SOA) was 1500 ms. The noise burst target was presented 100 ms after the offset of one of the three tones.

Procedures: Participants sat in a sound attenuated booth (IAC, Bronx, NY). There were three conditions that varied in the predictability of the target noise burst occurring after a pure tone. In Condition 1 the target occurred with 100% probability following the low tone (Predictable Target). In Condition 2 the target occurred with 100% probability following the high tone (Predictable Target). In Condition 3 the target occurred randomly after either the low or the high tone, with equal 50% probability (Unpredictable Target). The target never followed the middle tone (Fig. 1). Participants were instructed to press a designated response button as rapidly as possible upon detecting the target. Conditions were presented in a Latin Square design. Stimuli were presented in 8 stimulus blocks in each of the 3 conditions (24 blocks in total) which were randomly distributed across participants. Each block was approximately 2.5 min in duration (20 min total). Behavioral and EEG data were averaged across the 8 blocks, separately by condition. Tones of each frequency (low, middle, and high) were presented 280 times in each experimental condition. Breaks were provided at the mid-point and participants were unhooked from the amplifiers so that they could walk around and have a snack. Shorter breaks were given as needed. Total session time, including instructions, cap placement, recording time, and breaks was approximately 2 h.

Importantly, subjects were only instructed to respond to the target and were not given any prior information about the cue-target contingencies. Indeed, subjects were not informed that the cues were relevant in any way to the target detection task. Rather, the cue-target contingencies could only be learned implicitly through listening to the sound sequences, thus modeling the implicit learning (i.e., occurring in the absence of explicit prior information) that is common in daily life.

2.3. EEG recording

EEG was recorded with a 32-channel electrode cap (Electro-Cap International Inc., Eaton, OH), using the International modified 10–20 system, including electrodes placed on the left and right mastoids. The tip of the nose was used for the reference electrode and P09 for the ground electrode. A bipolar configuration was used to monitor the horizontal electrooculogram (EOG), between electrodes F7 and F8, and between FP1 and an external electrode placed below the left eye to monitor the vertical EOG. Impedances were maintained below 5 kOhms. EEG was collected with a bandpass filter of 0.05–200 Hz and a sampling rate of 1000 Hz. Offline, epochs −100 to 1000 ms from tone onset were created for each tone response. Epochs were bandpass filtered at 0.1–30 Hz and baseline-corrected prior to subsequent analyses. Artifacts due to movement, eye-blinks, or other noise sources were eliminated by removing epochs containing voltage deflections exceeding +/− 75 microvolts. Event-related potentials (ERPs) under each experimental condition and for each subject were then generated by averaging clean EEG epochs across trials.

2.4. Data and statistical analyses

Behavior: Reaction times (RTs) were recorded by computer for offline analysis and measured relative to the onset of the *cue* tones. Hits were counted as responses that occurred 0–1200 ms from the onset of the target stimulus. Hit rate was calculated as the total number of correct responses to the target divided by the number of targets, in each condition separately. False alarm rate was calculated as the number of button-presses to any non-target

sound divided by the number of non-target sounds. Misses were counted as the absence of a response to a target.

Behavioral reaction times generally do not exhibit a normal distribution of values. This deviation from normality was ameliorated by inverse-transforming reaction times prior to conducting statistical analyses (Ratcliff, 1993). Studentized *t*-tests for dependent samples were conducted to assess effects of predictability on hit rate and reaction time.

ERPs: Peak amplitude of each obligatory and endogenous ERP component was visually identified in the grand-averaged mean ERPs in each condition. These components included the P1, N1, and P2 components elicited by the cue tone, a negative-going component elicited prior to the onset of the target (described below), and the N1 and P3b elicited by the target. The peak latency identified in the grand-mean waveforms was used to determine the window for measuring the mean amplitude of each component at the electrode where the components had the largest amplitude: the Fz electrode for the P1 and N1, the Cz electrode for the P2 and the negative-going component, and the Pz electrode for the P3b.

Mean amplitudes were calculated for each participant within a time window that was determined from the grand-mean peak of each ERP component, separately for each stimulus in each condition. For the P1 and N1, we used a 30 ms window centered on the grand-mean peak, for P2 a 40 ms window, and for the negative-going component and the P3b, a 60 ms window. Latencies of each component were defined by the time corresponding to the largest value occurring within the aforementioned windows. We report latency statistics only for those components that were significantly elicited, as determined by one-sample t-tests on the mean amplitude.

ERP amplitudes and latencies were statistically compared using repeated-measures t-tests or ANOVA (rmANOVA). Where data violated the assumption of sphericity, Greenhouse-Geisser corrections were applied and epsilon values reported. Effect sizes were computed as Cohen's *d* and Eta-squared (η^2), with Cohen's *d* = mean difference between groups/standard deviation of difference between groups, or *t*/ N, and $\eta^2 = SS_{effect}/SS_{total}$ (Lakens, 2013). For post-hoc analyses, Tukey HSD for repeated measures was conducted on pairwise contrasts only when main effects or interactions of the rmANOVA were significant. Contrasts were reported as significantly different at *p* < .05. Statistical analyses were performed using Statistica 12 software (Statsoft, Inc., Tulsa, OK) and ProStat (Poly Software, Inc., NY). Our initial analysis determined no effect of tone frequency (low cue vs. high cue) on target responses. Therefore, we collapsed the data across the low and high cue tone responses. Except for analyses of responses to the middle 'control' tone, all analyses and results presented in this report are based upon the combined low and high cue tone responses, with target predictability serving as the main independent variable.

3. Results

3.1. Behavior

Reaction time was modulated by the predictability of the targets. RT was significantly shorter for Predictable targets (*M* = 505 ms, SD=78) than for Unpredictable targets

(M = 590 ms, SD=51), $t(17)$ = 10.5, p <0.0001). The significantly shorter RT in the Predictable condition indicates that subjects implicitly learned the cue-target contingencies and exploited this implicit knowledge to improve their behavioral performance. Hit rate was significantly higher for Unpredictable targets than for Predictable targets, although the difference was small (M = 0.99, SD = 0.01 and M = 0.98, SD=0.02, respectively, $t(17)$=3.2, p = .005; difference in hit rate: M = 0.015, SD=0.02). The distribution of differences in hit rate between Predictable and Unpredictable conditions did not show a significant deviation from normality as measured by a Kolmogorov—Smirnov test (D = 0.2238; p > .20).

Fig. 2A displays the mean RTs of all participants for each of the first 25 trials of the Predictable (red line) and Unpredictable (blue line) target blocks. The difference in RT between Predictable and Unpredictable conditions significantly increased from the 1st to the 2nd half of these trials (trials 1–13 and 14–25, respectively; mean difference between 1st and 2nd half trials=26 ms, SD=19.5), thus indicating that learning of the cue-target contingencies improved over the course of a stimulus block (paired, two-tailed t-test $t(17)$ = 5.6; p < .00005; Cohen's d = 1.3; Fig. 2B).

3.2. ERP results

Fig. 3 displays the grand-mean ERPs elicited in the 100% (Predictable, red trace), 50% (Unpredictable, blue trace), and 0% (Predictable, black trace) target probability conditions at Fz, Cz, and Pz electrodes. Figs. 4 and 5 show corresponding ERP component amplitudes and latencies, respectively, measured in individual subjects.

There were no significant differences in the amplitudes of the P1 and N1 components across stimulus conditions (P1: $F_{2,34}$ = 1.3, p = .33; N1: $F_{2,34}$ <1, p = .93). Effects of target predictability emerged in the amplitude of the positive-going P2 component, which varied systematically with the conditional probability of the target (Fig. 3, Cz electrode). Specifically, P2 was largest, intermediate, and smallest in amplitude when the conditional probability of the target was 0, 50, and 100%, respectively (P2: $F_{2,34}$= 10.9, ε = 0.71, p < .001, η_p^2 = 0.39).

Furthermore, when the conditional probability of the target was 100%, when the target always followed a specific cue tone, a distinct negative-going deflection was observed during the delay period between cue offset and target onset, which peaked at approximately 300 ms (Fig. 3, red trace). This negative component was absent when the conditional probability of the target was 0%, that is, when the target never followed that specific cue tone (Fig. 3, black trace). Thus, henceforth, we provisionally refer to this component as a 'contingency response' (CR). There was a main effect of target predictability on the CR component amplitude ($F_{2,34}$ = 13.7, ε = 0.90, p < .0001, η_p^2 =0.45). Post-hoc analysis revealed no significant difference in the amplitude of the CR between Predictable (100%) and Unpredictable (50%) conditions (p = .92), and that the CR amplitude was significantly larger when the conditional probability of the target was 100% and 50% than when the conditional probability of the target was 0% (p < .001).

With the exception of the N1 component, the peak latency of all components elicited by the cue tone decreased with increasing conditional probability of the target, consistent with the

degree to which subjects could expect the target to appear based on the frequency of the cue tone (P1: $F_{2,34} = 4.5$, $p = .02$, $\eta^2 = 0.07$; N1: $F_{2,34} = 3.2$, $p = .05$; P2: $F_{2,34} = 16.0$, $p < .0001$, $\eta^2 = 0.23$;. The peak latency of the CR in the Predictable (100%) condition was significantly shorter than in the Unpredictable (50%) condition ($t(17) = 3.68$, $p = .002$).

The amplitude of the N1 evoked by the target was significantly larger when the conditional probability of the target was 50% compared to when the conditional probability was 100% (Fig. 3, Target N1 blue vs red traces; $t(17) = 3.0$, $p = .008$, Cohen's $d = 0.71$). The peak latency of the target N1 was significantly shorter in the Predictable than in the Unpredictable condition ($t(17) = 7.0$, $p < .0001$, Cohen's $d = 1.65$). The amplitude of the P3b evoked by the target was similar between the Predictable and Unpredictable conditions ($t(17) = 1.7$, $p = .11$). However, the peak latency of the P3b was significantly shorter when the target was predictable than when it was unpredictable ($t(17) = 18.55$, $p < .0001$, Cohen's $d = 4.36$).

Correlations between behavioral RTs and ERP amplitudes and latencies are summarized in Table 1. There were no significant relationships between RT and the P1 and N1 components. RT was significantly correlated with the amplitudes of the P2 and the CR, but not with those of either of the target-related components (N1 and P3b). RT was significantly correlated with the latencies of the CR and both target-related components.

To further test the feature-specificity of putative neuronal signatures of learned expectancy, we compared the amplitude of the CR elicited by the mid-frequency 'control' tone, which was never followed by the target in any of the conditions, with that of the CR elicited by the low-and high-frequency tones. We reasoned that if the CR is feature-specific, for example, tracking only those features that serve as cues to the target, then the CR elicited by the low and high tones should be significantly larger than the CR elicited by the middle tone. Furthermore, the effect size corresponding to this difference should be larger in the Predictable (0 and 100%) than in the Unpredictable (50%) condition. Consistent with the feature-specificity hypothesis, the amplitude of the CR elicited by the low and high tones was significantly larger (more negative) than the amplitude of the ERP elicited by the middle tone within the same time period in both the Predictable and Unpredictable conditions (Fig. 6A; CR amplitude Predictable: $t(17) = 4.4$, $p < .0005$, Cohen's $d = 1.04$; CR amplitude Unpredictable: $t(17) = 3.4$, $p = 0.003$, Cohen's $d = 0.8$). Moreover, the effect size was larger in the Predictable than in the Unpredictable condition.

Because the mid-frequency tone was not behaviorally relevant, we expected that ERPs elicited by this tone would not vary with the predictability of the target. However, the response to the control tone was modulated by the predictability of the target (Fig. 6B). The amplitude of the P1 was significantly reduced and the waveform in the latency range of the CR was significantly more positive when the target was predictable (100%) compared with when the target was unpredictable (50%) (P1: $t(17) = 2.4$, $p = .03$, Cohen's $d = 0.57$; CR: $t(17) = 3.5$, $p = .003$, Cohen's $d = 0.83$). Additionally, the latency of the P2 component was significantly shorter when the target was unpredictable compared with when it was predictable ($t(17) = 2.5$, $p = .02$, Cohen's $d = 0.59$).

3.3. ERP scalp topography

Grand-mean ERP voltage distributions under the three conditional probability conditions for the N1, CR, and the P3b components and corresponding derived scalp current density (SCD) topographies are shown in Fig. 7. SCD is a reference-independent measure of the distribution of current sources and sinks contributing to the distribution of scalp ERPs, which mainly reflects the activity of cortical generators (Pernier et al., 1988; Tenke and Kayser, 2012). The N1 displayed a fronto-central distribution of SCD with bilateral negative peaks consistent with neuronal dipole sources in the right and left auditory cortex along the superior temporal plane (Giard et al., 1994; Naatanen and Picton, 1987). In contrast, the CR component displayed a single fronto-central negative peak. The P3b component showed a single focus with a broad distribution centered around the Pz electrode. Given its posterior-parietal scalp distribution (Figs. 3 and 7), we identified this component as a P3b, as described in previous studies (Wronka et al., 2012; Polich, 2007; Volpe et al., 2007). Consistent with amplitude data shown in Fig. 4, the peak value of SCD for the N1 component elicited by the low and high tones did not vary with the conditional probability of the target. Similarly, the positive peak of SCD corresponding to the P3b component did not differ as a function of conditional probability. In contrast, the peak value of SCD for the CR increased with increasing conditional probability.

3.4. Assessment of motor contributions to the CR and P2

The amplitudes of the P2 and CR components were moderately correlated with individual subjects' reaction times. This raises the possibility that modulations of these components partly reflect processes involved in preparing to execute the behavioral task (van Boxtel and Brunia, 1994b; van Boxtel and Brunia, 1994a; Brunia and van Boxtel, 2001). Although we cannot completely rule out a motor contribution to these components, two lines of evidence suggest that these responses cannot be explained solely by motor preparatory activity. First, motor preparatory activity would be expected to be similar in the Predictable (100%) and Unpredictable (50%) conditions. Specifically, in the Unpredictable condition, subjects would have learned that the target could follow either the high tone cue or the low tone cue and hence prepared to respond upon hearing one or the other (without being certain whether or not the target would follow them). Similarly, in the Predictable condition, subjects would have learned that the target always follows the low tone cue in one block and the high tone cue in the other block and prepared to respond accordingly. Despite this putative similarity in subjects' motor preparation, the amplitude and latency of the CR and P2 components differed depending on the predictability of the target. Second, we more directly tested whether the increased CR amplitude and the reduced P2 amplitude reflected subjects' preparation to make a motor response, by re-averaging the EEG epochs with respect to the latency of the button press on each trial for each subject in the 50% and 100% conditions rather than stimulus onset. If the increased CR amplitude and reduced P2 amplitude were exclusively related to motor response preparation and execution, then a prominent negative deflection (accounting for the CR and reduction in P2) consistently locked to the timing of the subjects' button press should be evident in the re-averaged ERP within the time frame preceding the button press. On the other hand, if the CR and reduced P2 amplitude reflect subjects' expectation for the target's appearance, there should be no such negativity in the re-averaged ERP-or at least it should be markedly reduced compared with the difference

in amplitude between the CR in the 0% conditional probability condition (wherein there were no targets and button press responses) and that in the other conditions (50% and 100% conditional probability) when ERPs are averaged with respect to the onset of the cue and target (as shown in Fig. 3). Importantly, no negative deflections were evident in the grand-mean re-averaged ERP waveforms that could account for the CR and reduced P2 amplitudes (Fig. 8). However, the re-averaged ERPs did display a prominent positive component following the button press, which likely corresponds with the P3b, consistent with the significant correlation observed between P3b latency and subjects' reaction times.

Based on these lines of evidence, we conclude that modulations in the amplitude of the CR and P2 components elicited in our implicit learning paradigm predominantly reflected learned expectancy for, or anticipatory attention to, the target rather than motor preparatory activity.

4. Discussion

We investigated behavioral and neuronal correlates of learned expectancy. Results demonstrated implicit learning of the arbitrary association between a target sound and a cue tone that reliably predicted the target's appearance. Implicit learning was revealed by shorter reaction times to the target when it was reliably predicted compared with when it was unreliably predicted. In addition, we observed a negative-going ERP component (which we denote here as a contingency response [CR]), and a positive-going component (P2) whose amplitudes and latencies were systematically modulated by the degree to which the cue tone predicted the target. That is, CR amplitude was largest and P2 amplitude was smallest when the target always followed a given cue (100% conditional probability) compared to when the target never followed the cue (0% conditional probability). Additionally, the peak latencies of both components were shortest when the target was fully predicted by the cue. It is likely that the roughly inverse relationship observed between modulations of the P2 and CR amplitude and latency by target predictability is due to partial overlap of these temporally contiguous components. Moreover, the amplitude of the response to the mid-frequency tone, which never served as a cue to the target, was significantly reduced (more positive) in the CR latency range. Thus, the P2 and CR tracked probabilistic cue-target relationships that participants implicitly derived from the stimulus statistics. These findings suggest that the P2 and CR constitute neuronal signatures of learned expectancy insofar as their amplitude and latency reflect the degree to which participants implicitly learn the cue-target contingencies occurring within a stimulus block that allow them to predict when a target will occur. The P2 and CR components may represent ERP correlates of a 'prediction signal' in line with the predictive coding framework (Friston, 2005; Durschmid et al., 2019). Alternatively, the modulation of the amplitude and latency of these components may represent the *confidence* with which subjects could expect the target to appear (Sherman et al., 2016).

Learning of cue-target contingencies was observed in real time, as demonstrated by the increase in the difference in reaction times between Predictable and Unpredictable conditions over trials. Expectancies were learned implicitly in that participants were instructed only to respond to the target stimulus and were not given explicit information about the cue-target contingencies. Although it is possible that participants ultimately

became aware of the specific cue-target contingencies over the course of a given stimulus block, the learning of the association was derived implicitly because it occurred in the absence of explicit information about any contingencies in the stimulus sequence and about the behavioral relevance of the cue tones for their response to the target. Hence, the modulations of the CR and P2 by target predictability and their behavioral correlates reflect the type of learned expectancy that is common in everyday life and which emerges spontaneously by exposure to statistical patterns in the environment. However, an alternative interpretation is that the CR and P2 represent an effect of attention, where the listener is learning what to attend to- the cue that reliably predicts the target- rather than reflecting expectation for the target per se.

Consistent with earlier CNV studies, the CR component had a fronto-central scalp distribution compatible with sources within the supplementary motor area and frontal cortex (Rosahl and Knight, 1995; Hamano et al., 1997; Chennu et al., 2013; Mento et al., 2013). Based on this anatomical and functional similarity, we tentatively identify the CR as a type of CNV. The supplementary motor area, in particular, is thought to play an important role in processing temporal information and generating temporal expectations (Mento et al., 2013; Wiener et al., 2010). It is noteworthy that the CR was highly circumscribed in time, occurring largely within the short 100 ms delay between the offset of the cue tones and the onset of the target. This short duration contrasts with the much longer development of the CNV, typically on the order of seconds, reported in previous work, and more closely parallels the considerably shorter timescales over which perceptual expectations operate in everyday life, such as when processing phonemic transitions in speech and melodic transitions in music.

The reduced P2 amplitude and the negativity associated with the CR were not observed when ERPs were re-averaged with respect to the latency of subjects' button press responses. This indicates that these ERP components reflect expectancy for, or anticipatory attention to, the target rather than only motor preparatory activity. Thus, insofar as the CR reported here is related to the CNV described in previous studies, the present work adds to the evidence that the CNV is not purely a motor phenomenon (van Boxtel and Brunia, 1994b, van Boxtel and Brunia, 1994a; Brunia and van Boxtel, 2001; Mento et al., 2013). It is not clear how the CR relates to the reported prediction-related reduction in high-frequency gamma activity, recorded using ECoG from the frontal cortex of patients with epilepsy (Durschmid et al., 2019), except in the timing of the events, with both response modulations occurring immediately prior to expected sounds. As high-frequency activity is not readily detectable using scalp EEG, establishing a link between these electrophysiological indices will likely require investigations using invasive techniques.

We found that the latency of the P1 component elicited by the cue tones became progressively shorter as the conditional probability of the target increased from 0 to 100%. This suggests enhanced neuronal processing of cues which are highly predictive of the target even at the earliest stages of cortical processing. This early neuronal enhancement may reflect an allocation of selective attention to the frequency of the cue that is most behaviorally relevant in a given stimulus block (Woldorff et al., 1993; Rao et al., 2010; Fritz et al., 2007) and is consistent with the 'learned predictiveness' model of attention

and associative learning (Le Pelley et al., 2016; Mackintosh, 1975), whereby attention is preferentially allocated to cues that accurately predict behaviorally significant events. However, neither the amplitude nor the latency of the N1 component elicited by the cue tones was modulated by the predictability of the target. This result contrasts with earlier findings showing modulations of the N1 by both statistical learning and selective attention (Naatanen and Picton, 1987; Woldorff et al., 1993; Hillyard et al., 1973; Abla et al., 2008; Teinonen and Huotilainen, 2012). The lack of an effect of expectancy on the N1 elicited by the cues may be explained by the fact that in our paradigm expectancies were formed with respect to the target's appearance, given a particular cue, and not to the appearance of a particular cue tone, which was random and thus unpredictable. Because the cues were not predictable, it is perhaps not surprising that the obligatory N1 responses to the cues were not modulated by expectancy.

An unanticipated finding was that the response to the middle frequency 'control' tone, the tone that never cued the target, was modulated by target predictability. The more predictable the target was, the smaller the amplitude of the P1 and the more positive the waveform occurring within the time frame of the CR. Moreover, the latency of the P2 elicited by the control tone was significantly shorter when the target was less predictable. One possible interpretation of these results is that when the target followed the low and high tone cues less reliably, each with 50% probability, participants may have expected that the target could randomly follow any of the three tones in the sequence. Even though the target never followed the middle frequency tone, with only implicit information to rely on it was always possible that the middle tone could become behaviorally relevant, as supported by the fact that the ERP within the time frame of the CR was less positive in the Unpredictable condition (50%) than the Predictable condition (100%), in which the target exclusively followed one of the three tones. We speculate that modulation of the response occurred due to lack of explicit information about the middle tone's irrelevance to the performance of the target detection task.

Target-specific N1 and P3b responses were also modulated by predictability. The N1 evoked by the target was significantly larger when the target was unexpected (its conditional probability was 50%) compared with when it was expected (its conditional probability was 100%; see Figs. 3 and 4). This result is consistent with predictive coding models in which unexpected events elicit greater activity when a mismatch between the expected and the sampled input is detected (Friston, 2005), and also fits with models of 'expectation dampening' in which predicted sensory inputs are directly suppressed (Heilbron and Chait, 2018; Friston, 2005; Todorovic et al., 2011; Symonds et al., 2017; Han et al., 2019; Summerfield and de Lange, 2014; Richter and de Lange, 2019; den Ouden et al., 2012; Press et al., 2020a). Targets which were cued with less reliability (in the 50% probability condition) may have been perceived as more salient or 'surprising' than those which were more reliably associated with a preceding cue (in the 100% probability condition). A different interpretation of this result, however, is that the N1 to the target was actively enhanced under the 50% probability condition due to changes in sensory gain. In particular, in cases where predictive relationships are less certain (as in the 50% probability condition) agents may increase the weight they give to incoming signals to aid learning (Summerfield

and Egner, 2009; Frings et al., 2019; Press et al., 2020b). Unfortunately, we cannot distinguish between these interpretations in the present study.

The P3b has been variously considered an index of voluntary attention, context updating in working memory, stimulus uncertainty and classification (e.g., as a target), and of the link between stimuli and task-relevant responses (Sutton et al., 1965; Polich, 2007; Verleger, 2020). Its peak latency is thought to relate to stimulus evaluation speed (Sutton et al., 1965; Polich, 2007; Wronka et al., 2012; Verleger et al., 2018). In broad agreement with these interpretations, we found that the peak latency of the P3b was significantly correlated with subjects' reaction times (Polich, 2007) and was significantly shorter when the target was predictable compared to when it was unpredictable. Hence, the P2, CR, and P3b all appear to track subjects' degree of uncertainty about the statistical relationship between the cue and target and reflect interrelated neuronal processes underlying the generation of predictions (P2 and CR) and the top-down evaluation of task-relevant events (P3b) modulated by the strength of these predictions.

In summary, our results demonstrate implicit learning of cue-target contingencies and identify the CR as a potential neuronal signature of expectancies that are built up during the learning process. The CR occurred immediately prior to the onset of the target and its amplitude and latency were systematically modulated by the strength of the cue-target relationship: The more predictive the cue was of the target's appearance, the larger the amplitude and the earlier the latency of the CR. Behavioral and neurophysiological correlates of learning emerged through 'passive' exposure to the statistical regularities in the sensory input without explicit or prior knowledge of the 'rules' regarding the arbitrary relationships in the stimulus sequence. Thus, the CR may represent a neuronal correlate of prediction.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

Abla D, Katahira K, Okanoya K, 2008. On-line assessment of statistical learning by event-related potentials. J. Cogn. Neurosci. (United States) 20, 952–964.

Arjona A, Escudero M, Gomez CM, 2016. Cue validity probability influences neural processing of targets. Biol. Psychol. (Netherlands) 119, 171–183.

Arnal LH, Giraud AL, 2012. Cortical oscillations and sensory predictions. Trends Cogn. Sci. (England) 16, 390–398.

Auksztulewicz R, Barascud N, Cooray G, Nobre AC, Chait M, Friston K, 2017. The cumulative effects of predictability on synaptic gain in the auditory processing stream. J. Neurosci. (United States) 37, 6751–6760.
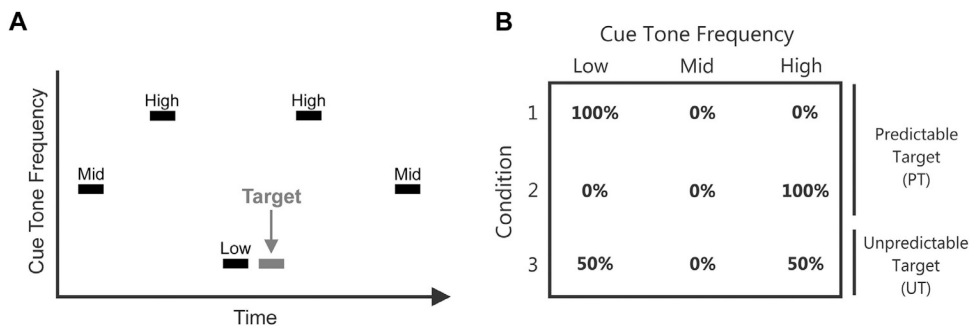
Auksztulewicz R, Schwiedrzik CM, Thesen T, Doyle W, Devinsky O, Nobre AC, Schroeder CE, Friston KJ, Melloni L, 2018. Not all predictions are equal: "What" and "when" predictions modulate activity in auditory cortex through different mechanisms. J. Neurosci. (United States) 38, 8680–8693.

Barascud N, Pearce MT, Griffiths TD, Friston KJ, Chait M, 2016. Brain respnoses in humans reveal ideal observer-like sensitivity to complex acoustic patterns. Proc. Natl. Acad. Sci. USA (United States) 113, E616–E625.

Bendixen A, SanMiguel I, Schröger E, 2012. Early electrophysiological indicators for predictive processing in audition: a review. Int. J. Psychophysiol. (Netherlands) 83, 120–131.

Bickel S, Dias EC, Epstein ML, Javitt DC, 2012. Expectancy-related modulations of neural oscillations in continuous performance tasks. Neuroimage (United States) 62, 1867–1876.

Bonte ML, Mitterer H, Zellagui N, Poelmans H, Blomert L, 2005. Auditory cortical tuning to statistical regularities in phonology. Clin. Neurophysiol. (Netherlands) 116, 2765–2774.

Brunia CH, van Boxtel GJ, 2001. Wait and see. Int. J. Psychophysiol. (Netherlands) 43, 59–75.

Chennu S, Noreika V, Gueorguiev D, Shtyrov Y, Bekinschtein TA, Henson R, 2016. Silent expectations: dynamic causal modeling of cortical prediction and attention to sounds that weren't. J. Neurosci. (United States) 36, 8305–8316.

Chennu S, Noreika V, Gueorguiev D, Blenkmann A, Kochen S, Ibanez A, Owen AM, Bekinschtein TA, 2013. Expectation and attention in hierarchical auditory prediction. J. Neurosci.J. Neurosci. 33, 11194–11205. [PubMed: 23825422]

Coull JT, 2009. Neural substrates of mounting temporal expectation. PLoS Biol (United States) 7, e1000166.

de Lange FP, Heilbron M, Kok P, 2018. How do expectations shape perception? Trends Cogn. Sci. (England) 22, 764–779.

Demarchi G, Sanchez G, Weisz N, 2019. Automatic and feature-specific prediction-related neural activity in the human auditory system. Nat. Commun. (England) 10:3440–019–11440–1.

den Ouden HE, Kok P, de Lange FP, 2012. How prediction errors shape perception, attention, and motivation. Front. Psychol. (Switzerland) 3, 548.

Donchin E, Gerbrandt LA, Leifer L, Tucker L, 1972. Is the contingent negative variation contingent on a motor response? Psychophysiology (United States) 9, 178–188.

Durschmid S, Reichert C, Hinrichs H, Heinze HJ, Kirsch HE, Knight RT, Deouell LY, 2019. Direct evidence for prediction signals in frontal cortex independent of prediction error. Cereb. Cortex (United States) 29 (11), 4530–4538.

Frings C, Merz S, Hommel B, 2019. The impact of stimulus uncertainty on attentional control. Cognition (Netherlands) 183, 208–212.

Friston K, 2005. A theory of cortical responses. Philos. Trans. R. Soc. Lond. B Biol. Sci. (England) 360, 815–836. [PubMed: 15937014]

Fritz JB, Elhilali M, David SV, Shamma SA, 2007. Auditory attention—focusing the searchlight on sound. Curr. Opin. Neurobiol. (England) 17, 437–455.

Garrido MI, Sahani M, Dolan RJ, 2013. Outlier responses reflect sensitivity to statistical structure in the human brain. PLoS Comput Biol (United States) 9, e1002999.

Giard MH, Perrin F, Echallier JF, Thevenet M, Froment JC, Pernier J, 1994. Dissociation of temporal and frontal components in the human auditory N1 wave: a scalp current density and dipole model analysis. Electroencephalogr. Clin. Neurophysiol. (Ireland) 92, 238–252.

Gomez CM, Arjona A, Donnarumma F, Maisto D, Rodriguez-Martinez EI, Pezzulo G, 2019. Tracking the time course of bayesian inference with event-related potentials:a study using the central cue posner paradigm. Front. Psychol. (Switzerland) 10, 1424.

Hamano T, Luders 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. (Ireland) 104, 257–268.

Han B, Mostert P, de Lange FP, 2019. Predictable tones elicit stimulus-specific suppression of evoked activity in auditory cortex. Neuroimage (United States) 200, 242–249.

Heilbron M, Chait M, 2018. Great expectations: is there evidence for predictive coding in auditory cortex? Neuroscience (United States) 389, 54–73.

Heimbauer LA, Conway CM, Christiansen MH, Beran MJ, Owren MJ, 2018. Visual artificial grammar learning by rhesus macaques (macaca mulatta): exploring the role of grammar complexity and sequence length. Anim. Cogn. (Germany) 21, 267–284.

Hillyard SA, 1969. Relationships between the contingent negative variation (CNV) and reaction time. Physiol. Behav. 4, 351–357.

Hillyard SA, Hink RF, Schwent VL, Picton TW, 1973. Electrical signs of selective attention in the human brain. Science (UNITED STATES) 182, 177–180.

Hsu YF, Le Bars S, Hamalainen JA, Waszak F, 2015. Distinctive representation of mispredicted and unpredicted prediction errors in human electroencephalography. J. Neurosci. (United States) 35, 14653–14660.

Hughes HC, Darcey TM, Barkan HI, Williamson PD, Roberts DW, Aslin CH, 2001. Responses of human auditory association cortex to the omission of an expected acoustic event. Neuroimage (United States) 13, 1073–1089.

Koelsch S, Vuust P, Friston K, 2019. Predictive processes and the peculiar case of music. Trends Cogn. Sci. (England) 23, 63–77.

Kok P, Failing MF, de Lange FP, 2014. Prior expectations evoke stimulus templates in the primary visual cortex. J. Cogn. Neurosci. (United States) 26, 1546–1554.

Lakens D, 2013. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Front. Psychol. (Switzerland) 4, 863.

Le Pelley ME, Mitchell CJ, Beesley T, George DN, Wills AJ, 2016. Attention and associative learning in humans: an integrative review. Psychol. Bull. (United States) 142, 1111–1140.

Lu K, Vicario DS, 2014. Statistical learning of recurring sound patterns encodes auditory objects in songbird forebrain. Proc. Natl. Acad. Sci. USA (United States) 111, 14553–14558.

Mackintosh NJ, 1975. A theory of attention: variations in the associability of stimuli with reinforcement. Psychol Rev 82, 276–298.

Mento G, 2017. The role of the P3 and CNV components in voluntary and automatic temporal orienting: a high spatial-resolution ERP study. Neuropsychologia (England) 107, 31–40.

Mento G, 2013. The passive CNV: carving out the contribution of task-related processes to expectancy. Front. Hum. Neurosci. (Switzerland) 7, 827.

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. (United States) 27, 425–439.

Mento G, Tarantino V, Sarlo M, Bisiacchi PS, 2013. Automatic temporal expectancy: a high-density event-related potential study. PLoS One (United States) 8, e62896.

Moldwin T, Schwartz O, Sussman ES, 2017. Statistical learning of melodic patterns influences the brain's response to wrong notes. J. Cogn. Neurosci. (United States) 29, 2114–2122.

Naatanen R, Picton T, 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. Psychophysiology (United States) 24, 375–425.

Naatanen R, Gaillard AW, 1974. The relationships between certain CNV and evoked-potential measures within and between vertex, frontal and temporal derivations. Biol. Psychol. (Netherlands) 2, 95–112.

Pelucchi B, Hay JF, Saffran JR, 2009. Statistical learning in a natural language by 8-month-old infants. Child Dev. (United States) 80, 674–685.

Perchet C, Garcia-Larrea L, 2005. Learning to react: anticipatory mechanisms in children and adults during a visuospatial attention task. Clin. Neurophysiol. (Netherlands) 116, 1906–1917.

Pernier J, Perrin F, Bertrand O, 1988. Scalp current density fields: concept and properties. Electroencephalogr. Clin. Neurophysiol. (Ireland) 69, 385–389.

Polich J, 2007. Updating P300: an integrative theory of P3a and P3b. Clin. Neurophysiol. (Netherlands) 118, 2128–2148.

Posner MI, 1980. Orienting of attention. Q. J. Exp. Psychol. (United States) 32, 3–25.

Press C, Kok P, Yon D, 2020a. The perceptual prediction paradox. Trends Cogn. Sci. (England) 24, 13–24.

Press C, Kok P, Yon D, 2020b. Learning to perceive and perceiving to learn. Trends Cogn. Sci. (England) 24, 260–261.

Rao A, Zhang Y, Miller S, 2010. Selective listening of concurrent auditory stimuli: an event-related potential study. Hear. Res. (Netherlands) 268, 123–132.

Ratcliff R, 1993. Methods for dealing with reaction time outliers. Psychol. Bull. (United States) 114, 510–532.

Richter D, de Lange FP, 2019. Statistical learning attenuates visual activity only for attended stimuli. Elife (England) 8. doi:10.7554/eLife.47869.

Rohrmeier MA, Koelsch S, 2012. Predictive information processing in music cognition. A critical review. Int. J. Psychophysiol. (Netherlands) 83, 164–175.

Rosahl SK, Knight RT, 1995. Role of prefrontal cortex in generation of the contingent negative variation. Cereb Cortex (United States) 5, 123–134.

Saffran JR, Aslin RN, Newport EL, 1996. Statistical learning by 8-month-old infants. Science (United States) 274, 1926–1928.

Saffran JR, Johnson EK, Aslin RN, Newport EL, 1999. Statistical learning of tone sequences by human infants and adults. Cognition (Netherlands) 70, 27–52.

SanMiguel I, Saupe K, Schröger E, 2013a. I know what is missing here: electrophysiological prediction error signals elicited by omissions of predicted "what" but not "when". Front. Hum. Neurosci. (Switzerland) 7, 407.

SanMiguel I, Widmann A, Bendixen A, Trujillo-Barreto N, Schröger E, 2013b. Hearing silences: human auditory processing relies on preactivation of sound-specific brain activity patterns. J. Neurosci. (United States) 33, 8633–8639.

Schiavo JK, Froemke RC, 2019. Capacities and neural mechanisms for auditory statistical learning across species. Hear. Res. (Netherlands) 376, 97–110.

Schröger E, Marzecova A, SanMiguel I, 2015. Attention and prediction in human audition: a lesson from cognitive psychophysiology. Eur J. Neurosci. (France) 41, 641–664.

Sherman MT, Seth AK, Kanai R, 2016. Predictions shape confidence in right inferior frontal gyrus. J. Neurosci. (United States) 36, 10323–10336.

Southwell R, Baumann A, Gal C, Barascud N, Friston K, Chait M, 2017. Is predictability salient? A study of attentional capture by auditory patterns. Philos. Trans. R. Soc. Lond. B Biol. Sci. (England) 372. doi:10.1098/rstb.2016.0105, Epub 2017 Jan 2.

Summerfield C, de Lange FP, 2014. Expectation in perceptual decision making: neural and computational mechanisms. Nat. Rev. Neurosci. (England) 15, 745–756.

Summerfield C, Egner T, 2009. Expectation (and attention) in visual cognition. Trends Cogn. Sci. (England) 13, 403–409.

Sutton S, Braren M, Zubin J, John ER, 1965. Evoked-potential correlates of stimulus uncertainty. Science (United States) 150, 1187–1188.

Symonds RM, Lee WW, Kohn A, Schwartz O, Witkowski S, Sussman ES, 2017a. Distinguishing neural adaptation and predictive coding hypotheses in auditory change detection. Brain Topogr. (United States) 30, 136–148.

Tecce JJ, 1972. Contingent negative variation (CNV) and psychological processes in man. Psychol. Bull. (United States) 77, 73–108.

Teinonen T, Huotilainen M, 2012. Implicit segmentation of a stream of syllables based on transitional probabilities: an MEG study. J Psycholinguist. Res. (United States) 41, 71–82.

Tenke CE, Kayser J, 2012. Generator localization by current source density (CSD): implications of volume conduction and field closure at intracranial and scalp resolutions. Clin. Neurophysiol. (Netherlands) 123, 2328–2345.

Todorovic A, van Ede F, Maris E, de Lange FP, 2011. Prior expectation mediates neural adaptation to repeated sounds in the auditory cortex: an MEG study. J. Neurosci. (United States) 31, 9118–9123.

van Boxtel GJ, Brunia CH, 1994a. Motor and non-motor components of the contingent negative variation. Int. J. Psychophysiol. (Netherlands) 17, 269–279.

van Boxtel GJ, Brunia CH, 1994b. Motor and non-motor aspects of slow brain potentials. Biol. Psychol. (Netherlands) 38, 37–51.

Verleger R, 2020. Effects of relevance and response frequency on P3b amplitudes: review of findings and comparison of hypotheses about the process reflected by P3b. Psychophysiology (United States) 57, e13542.

Verleger R, Keppeler M, Sassenhagen J, Smigasiewicz K, 2018. The oddball effect on P3 disappears when feature relevance or feature-response mappings are unknown. Exp. Brain Res. (Germany) 236, 2781–2796.

Volpe U, Mucci A, Bucci P, Merlotti E, Galderisi S, Maj M, 2007. The cortical generators of P3a and P3b: a LORETA study. Brain Res. Bull. (United States) 73, 220–230.

Wacongne C, Labyt E, van Wassenhove V, Bekinschtein T, Naccache L, Dehaene S, 2011. Evidence for a hierarchy of predictions and prediction errors in human cortex. Proc. Natl. Acad. Sci. USA (United States) 108, 20754–20759.

Walter WG, Cooper R, Aldridge VJ, Mccallum WC, Winter AL, 1964. Contingent negative variation: an electric sign of sensorimotor association and expectancy in the human brain. Nature (England) 203, 380–384.

Warker JA, Dell GS, 2006. Speech errors reflect newly learned phonotactic constraints. J Exp. Psychol. Learn Mem. Cogn. (United States) 32, 387–398.

Wiener M, Turkeltaub P, Coslett HB, 2010. The image of time: a voxel-wise meta–analysis. Neuroimage (United States) 49, 1728–1740.

Woldorff MG, Gallen CC, Hampson SA, Hillyard SA, Pantev C, Sobel D, Bloom FE, 1993. Modulation of early sensory processing in human auditory cortex during auditory selective attention. Proc. Natl. Acad. Sci. USA (United States) 90, 8722–8726.

Wronka E, Kaiser J, Coenen AM, 2012. Neural generators of the auditory evoked potential components P3a and P3b. Acta Neurobiol. Exp. (Wars) (Poland) 72, 51–64.

A

B

Cue Tone Frequency

|	Low	Mid	High	
1	**100%**	**0%**	**0%**	Predictable Target (PT)
2	**0%**	**0%**	**100%**	
3	**50%**	**0%**	**50%**	Unpredictable Target (UT)

Fig. 1.
Schematic diagram of the experimental paradigm. A. Stimuli consisted of randomly presented pure tones differing in frequency (low, mid, high). Pure tones served as cues to a white noise burst target that followed them. B. Two predictable conditions (PT), in which the target followed the low tone or the high tone with 100% probability, and one unpredictable condition (UT), in which the target followed either the low or the high tone randomly within the block. The target never followed the middle tone, which served as a control.
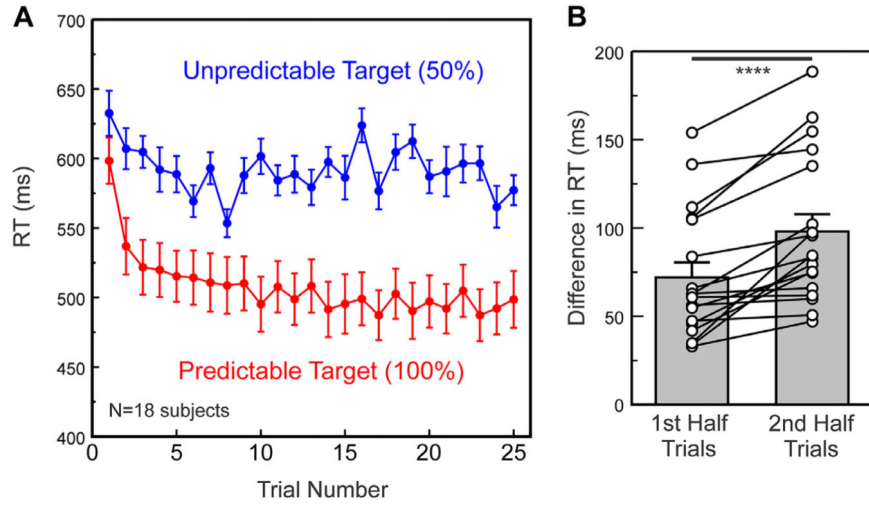
Fig. 2.
Behavioral results. A. Mean reaction times (RTs) are plotted for the first 25 trials within a stimulus block separately for the Predictable (red line) and Unpredictable (blue line) target conditions. Error bars represent the standard error of the mean. RTs are relative to cue-tone onset. B. Mean difference in average RT between Predictable and Unpredictable target conditions are separately plotted, along with individual subject data, for the first half (1–13) and second half (14–25) of trials within a stimulus block. Error bars represent standard error of the mean difference between conditions. Asterisks indicate significant difference between first and second half trials ($p < 0.0001$).
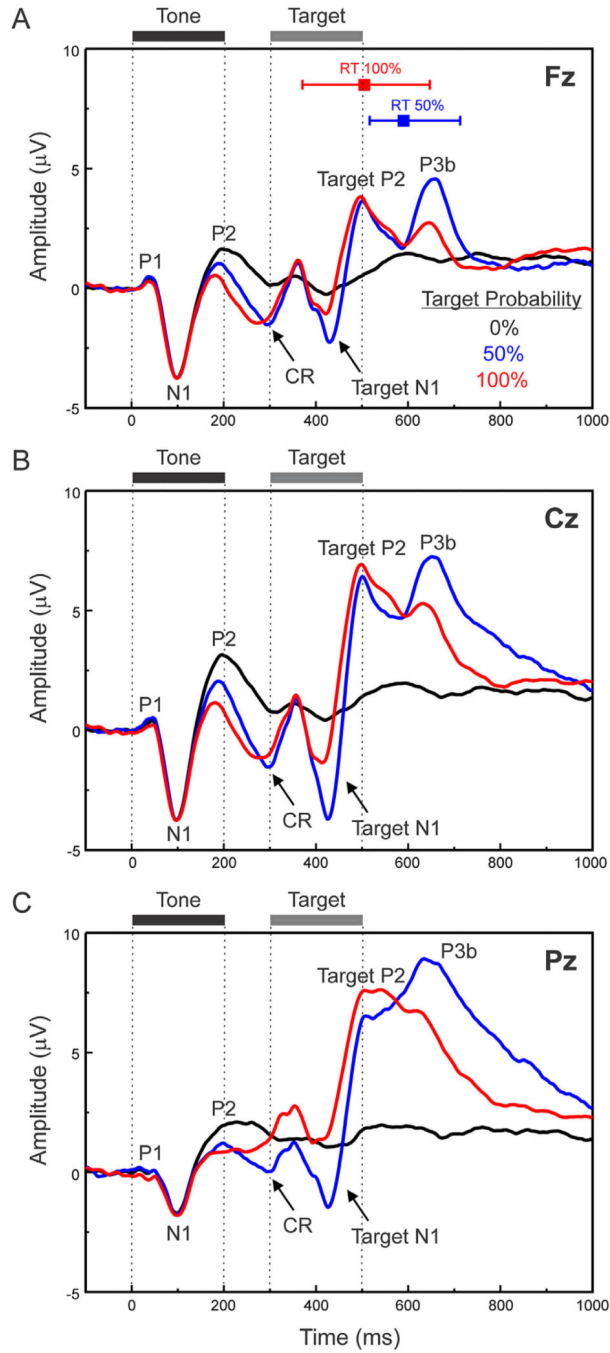
Fig. 3.
Grand mean ERPs elicited by the low and high cue tones and targets for each target probability condition. ERPs elicited in conditions where the target appeared with 0% (black trace), 50% (blue trace), and 100% (red trace) probability following the cue tone are shown. Data from Fz, Cz, and Pz electrodes are displayed in panels A, B, and C, respectively. Time frames of cue tones and targets are represented by the horizontal black and gray bar above the panels. Major response components are indicated in the plots. Mean reaction times (RTs)

under the 50% and 100% probability conditions are represented by the square symbols at the top of Panel A. Error bars represent the upper and lower ranges of RT.
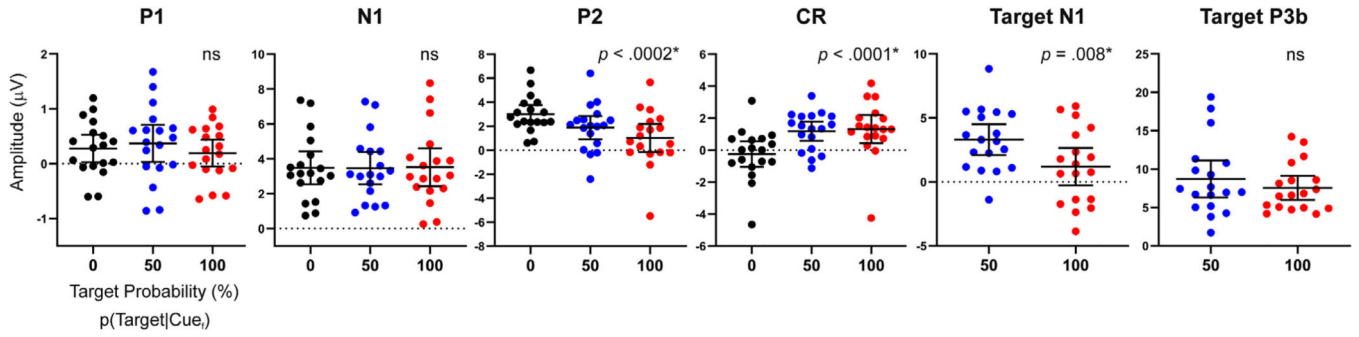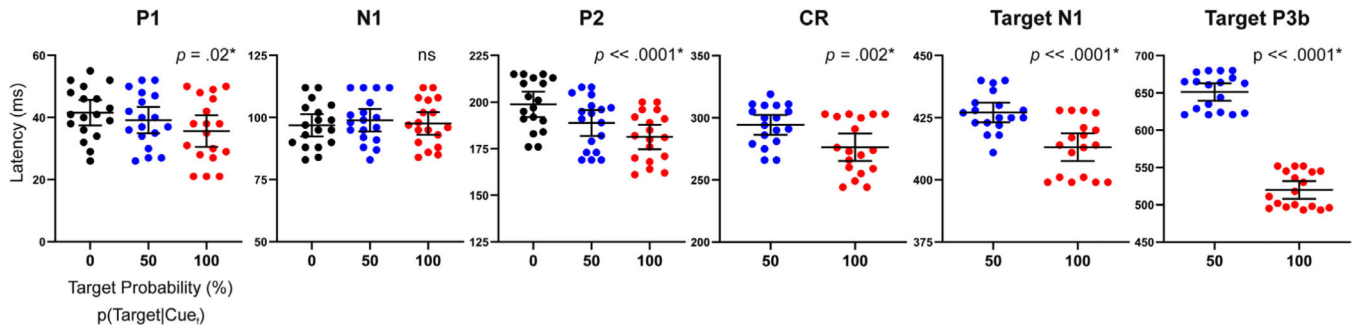
Fig. 4.
Amplitudes of the P1, N1, P2, CR, Target N1, and Target P3b components elicited by the low and high cue tones and by the target as a function of the conditional probability of the target. Data from individual subjects are plotted, with large horizontal bars representing the mean across subjects and smaller error bars representing 95% confidence intervals. The sign of negative-going ERP component (N1 and CR) amplitudes is reversed for display purposes. Asterisks represent significant differences across conditions (repeated measures ANOVA for 3-condition comparisons, and repeated measures *t*-test for 2-condition comparisons), with corresponding *p*-values indicated in the plots; 'ns': not statistically significant. Data for each of the components are plotted on different ordinate scales to facilitate visual comparison of differences across conditions.

Fig. 5.

Peak latencies of the P1, N1, P2, CR, Target N1, and Target P3b components elicited by the low and high cue tones and by the target as a function of the conditional probability of the target. Data from individual subjects are plotted, with large horizontal bars representing the mean across subjects and smaller error bars representing 95% confidence intervals. Asterisks represent significant differences across conditions (repeated measures ANOVA for 3-condition comparisons, and repeated measures *t*-test for 2-condition comparisons), with corresponding *p*-values indicated in the plots; 'ns': not statistically significant. Mean data for each of the components are plotted on different ordinate scales to facilitate visual comparison of differences across conditions. Note that latency data are not reported for the CR at a conditional probability of 0% because there was no significant CR elicited in this condition.
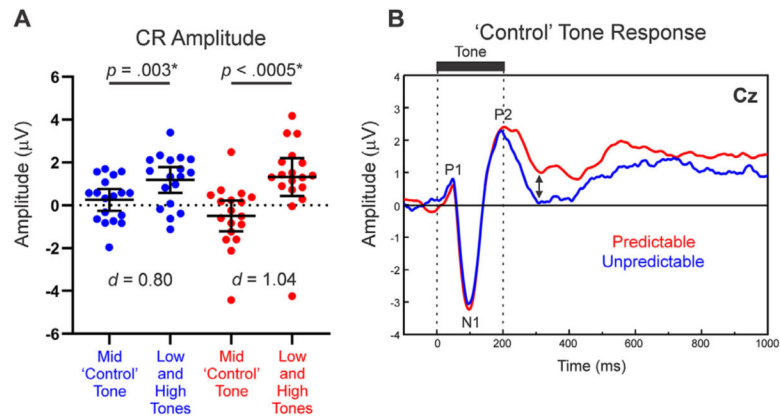
Fig. 6.

A. Amplitudes of the CR elicited by the low and high frequency tones and the middle frequency 'control' tone under the Unpredictable (blue symbols) and Predictable (red symbols) target conditions. Same conventions as in Figs. 4 and 5. Effect sizes (Cohen's *d*) corresponding to each comparison are indicated in the plot. B. Grand mean ERPs at Cz elicited by the control tone under the Unpredictable (blue trace) and Predictable (red trace) target conditions.

Voltage Distribution

Scalp Current Density
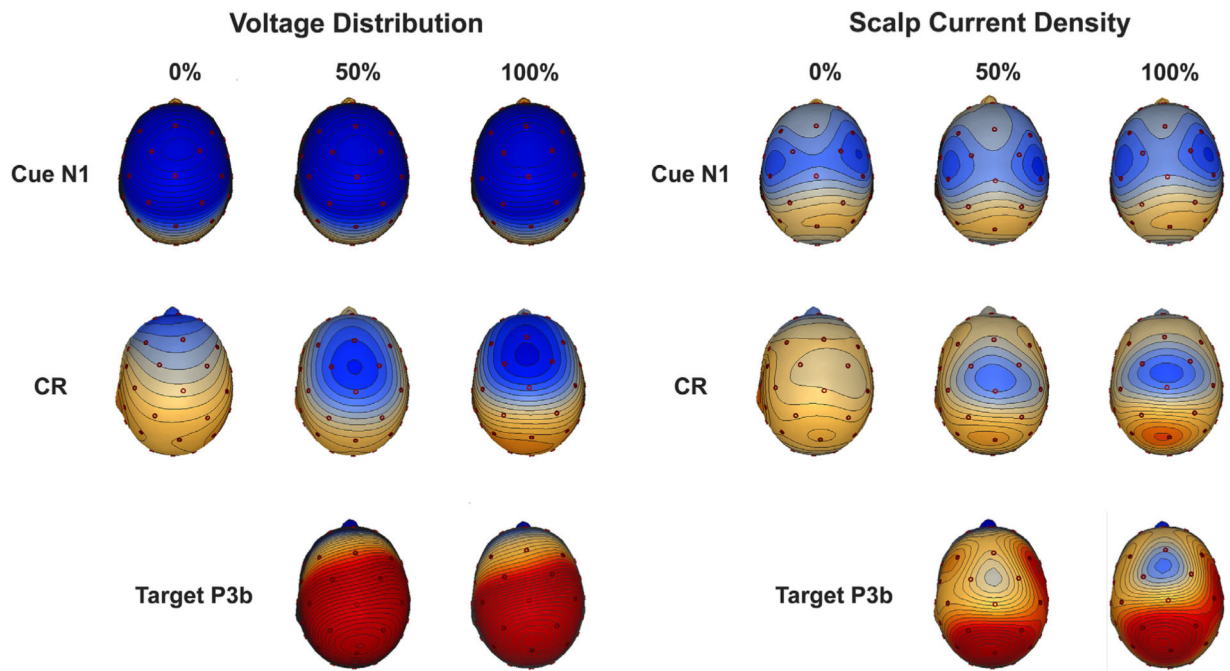
Fig. 7.
Grand-mean ERP voltage and corresponding derived scalp current density (SCD) distributions under the three conditional probability conditions for the N1 and CR elicited by the low and high tone cues and the P3b elicited by the target. Distributions for the N1, CR, and Target P3b components are shown in the top, middle, and bottom rows, as indicated. Distributions under the 0, 50, and 100% conditional probability conditions are shown in the left, middle, and right columns, as indicated. Warmer (redder) and cooler (bluer) colors represent progressively more positive and negative values, respectively, of voltage and SCD. Red circles indicate the location of the electrode contacts. Contour lines demarcate values of voltage and SCD differing by steps of 0.20 μV and 0.02 μV/cm^2, respectively. Voltage and SCD values have been normalized to the maximum value across all of the components and conditions shown.
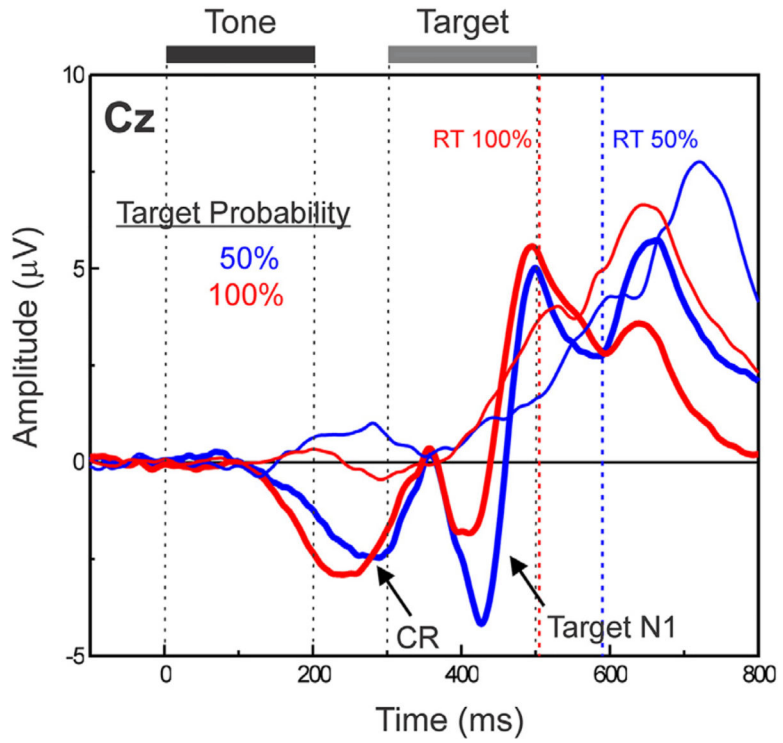
Fig. 8.
Grand mean ERPs elicited at the Cz electrode in the 50% and 100% conditional probability conditions are shown re-averaged with respect to the latency of subjects' button press responses on individual trials (blue and red thin lines, respectively) to assess motor-related activity. Included for comparison are the differences in grand mean ERPs, averaged with respect to cue tone onset, elicited in the 50% and 100% conditional probability conditions and the 0% conditional probability condition, represented by the blue and red thick lines, respectively. Mean reaction times in the 50% and 100% conditions are denoted by the blue and red vertical dashed lines. Negative deflections corresponding to the CR and target N1 components are indicated by arrows. Note the absence of negative deflections in the ERP waveforms that were re-averaged relative to the latency of the button press.

Table 1

Pearson correlations (*r*) between the amplitudes and peak latencies of ERP components and individual subjects' reaction times. *P*-values are included below the correlation coefficients, with statistically significant correlations indicated by asterisks.

	P1	N1	P2	CR	Target N1	Target P3b
Amplitude	*r* = 0.22	*r* = 0.24	*r* = 0.45	*r* = 0.42	*r* = −0.30	*r* = −0.08
	p = .19	*p* = .15	*p* = .006 *	*p* = .01 *	*p* = .07	*p* = .63
Latency	*r* = 0.20	*r* = −0.26	*r* = 0.20	*r* = 0.45	*r* = 0.51	*r* = 0.61
	p = .13	*p* = .13	*p* = .15	*p* = .006 *	*p* = .002 *	*p* = .0001 *