

Please cite this article in press as: Heilbron M, Chait M. Great expectations: Is there evidence for predictive coding in auditory cortex?. *Neuroscience* (2017), <http://dx.doi.org/10.1016/j.neuroscience.2017.07.061>

Neuroscience xxx (2017) xxx–xxx

GREAT EXPECTATIONS: IS THERE EVIDENCE FOR PREDICTIVE CODING IN AUDITORY CORTEX?

MICHA HEILBRON^{a,b,*} AND MARIA CHAIT^c

^a *Département de Biologie, École Normale Supérieure, Paris 75005, France*

^b *Université Pierre et Marie Curie P6, Paris 75005, France*

^c *Ear Institute, University College London, London WC1X 8EE, United Kingdom*

Abstract—Predictive coding is possibly one of the most influential, comprehensive, and controversial theories of neural function. While proponents praise its explanatory potential, critics object that key tenets of the theory are untested or even untestable. The present article critically examines existing evidence for predictive coding in the auditory modality. Specifically, we identify five key assumptions of the theory and evaluate each in the light of animal, human and modeling studies of auditory pattern processing. For the first two assumptions – that neural responses are shaped by expectations and that these expectations are hierarchically organized – animal and human studies provide compelling evidence. The anticipatory, predictive nature of these expectations also enjoys empirical support, especially from studies on unexpected stimulus omission. However, for the existence of separate error and prediction neurons, a key assumption of the theory, evidence is lacking. More work exists on the proposed oscillatory signatures of predictive coding, and on the relation between attention and precision. However, results on these latter two assumptions are mixed or contradictory. Looking to the future, more collaboration between human and animal studies, aided by model-based analyses will be needed to test specific assumptions and implementations of predictive coding – and, as such, help determine whether this popular grand theory can fulfill its expectations.

This article is part of a Special Issue entitled: Sequence Processing. © 2017 The Author(s). Published by Elsevier Ltd on behalf of IBRO. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Key words: predictive coding, auditory, MMN, SSA, bayesian brain.

*Correspondence to: M. Heilbron, Département de Biologie, École Normale Supérieure, 46 rue d'Ulm, 75005 Paris, France.

E-mail address: micha.heilbron@ens.fr (M. Heilbron).

Abbreviations: AB, alternations; AC, auditory cortex; IFG, inferior frontal gyrus; ISI, inter stimulus intervals; MMN, mismatch negativity; PC, predictive coding; rIFG, right inferior frontal gyrus; RS, repetition suppression; SSA, stimulus specific adaptation; STG, superior temporal gyrus; STRF, spectrotemporal patterns of intensities.

<http://dx.doi.org/10.1016/j.neuroscience.2017.07.061>

0306-4522/© 2017 The Author(s). Published by Elsevier Ltd on behalf of IBRO.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Contents

Introduction	00	12
Predictive coding in cortex – Foundations and assumptions	00	14
Animal studies	00	15
Prediction in auditory cortical neurons	00	16
The laminar profile of prediction	00	17
Discussion	00	18
Human imaging and electrophysiology	00	19
Predictive coding and MMN	00	20
Repetition suppression – Adaptation or expectation?	00	21
Expectation and surprise along the auditory hierarchy	00	22
Hearing silences: Omission as a window into prediction	00	23
Predictability and precision	00	24
Attention as precision	00	25
The rhythms of prediction	00	26
Effective connectivity – Clues from DCM	00	27
DCM for MEG and EEG	00	28
Dynamic Causal modeling of MMN	00	29
Temporal deviants and top-down predictions	00	30
Discussion	00	31
Conclusion	00	32
Uncited references	00	33
Acknowledgments	00	34
References	00	35

INTRODUCTION

How does the brain make sense of the world? A popular theory addressing this question is predictive coding (PC). Simply put, PC states that the brain infers what is 'out there' by constantly *predicting* what is out there, and then improving those predictions. More technically, PC proposes that the brain constructs a hierarchical, *generative* model of the world – a model capable of *generating* patterns of activity 'from the top-down' that external stimuli would elicit 'from the bottom-up'. The perceiving brain continuously tries to 'fit' such models by predicting the incoming sensory input. Bad fits signal *prediction errors* that leverage increasingly accurate estimates (recognition); and, over time, a modified model (perceptual learning).

As a biological basis for Bayesian theories of perception and cognition, PC offers compelling explanations for phenomena from psychology (Knill and Pouget, 2004) neuroanatomy (Friston, 2005) and electrophysiology (Rao and Ballard, 1999). Hailed by some as providing a 'grand unified theory of the brain' (Friston, 2010) the framework has drawn a considerable amount of attention (Hohwy, 2013; Clark, 2013, 2016). But predictive coding faces many challenges. By ascribing a central

62 role to top-down *expectations* of bottom-up inputs, PC
63 advocates a radical break with traditional feed-forward
64 accounts of perception. A break, some worry, too radical
65 since core tenets of the theory are, at best, untested
66 (Egner and Summerfield, 2013) or, at worst, untestable
67 (Kogo and Trengove, 2015).

68 Initially, PC was conceptualized in the context of
69 visual processing (Rao and Ballard, 1999; Lee and
70 Mumford, 2003). However, the auditory system quickly
71 became a popular test bed, with many studies capitalizing
72 on the auditory Mismatch Negativity (MMN; Näätänen
73 et al., 1978, 2007), perhaps the most well-studied neural
74 signature of surprise or error processing. The present
75 review critically evaluates the evidence for PC in auditory
76 cortex. In keeping with this Special Issue, we will limit our-
77 selves to relatively low-level auditory patterns (as
78 opposed to e.g. speech and language; but see Arnal
79 et al., 2011; Sohoglu et al., 2012; Gagnepain et al.,
80 2012). There exist several recent reviews of predictive
81 representation in audition (Winkler and Schröger, 2015;
82 Schröger et al., 2014, 2015; Winkler and Czigler, 2012;
83 Winkler et al., 2009). In contrast, the present analysis
84 specifically attempts to delineate key assumptions shared
85 by different PC models (cf. Rao and Ballard, 1999; Rao,
86 2005; Friston, 2005, 2010; Bastos et al., 2012;
87 Spratling, 2008a,b, 2010; see Spratling, 2015 for review)
88 and assess whether these assumptions are supported by
89 empirical evidence in the auditory modality.

90 In the next section we will briefly recap these basic
91 assumptions and their empirical ramifications, before
92 evaluating them in more detail in the light of recent
93 evidence.

94 **PREDICTIVE CODING IN CORTEX –**
95 **FOUNDATIONS AND ASSUMPTIONS**

96 Sensory cortex is organized hierarchically. At each
97 processing level, neurons integrate information from
98 multiple neurons at the level below, thus encoding
99 increasingly abstract information over ever larger
100 temporal and spatial scales. But cortex is reciprocally
101 connected, so neurons also receive input from the level
102 above (Felleman and van Essen, 1991).

103 Traditionally, higher levels were assumed only to
104 *modulate* lower levels, e.g. by prioritizing the processing of
105 certain inputs over others. But in PC, following the
106 proposal by Mumford (1992), the abstract information at
107 higher levels *informs* and potentially *drives* neurons at lower
108 levels by signaling a (prior) ‘best guess’ of their activity. At
109 the lower level, the difference between the predicted and
110 actual activity elicits a *prediction error* that is propagated
111 back to the level above, where it is used to generate a new
112 and improved (posterior) estimate. This routine is repeated,
113 simultaneously throughout the hierarchy, until the most likely
114 estimate is reached and the stimulus is perceived.

115 In this scheme – arguably the standard version of PC
116 (Rao and Ballard, 1999; Friston, 2005; Bastos et al.,
117 2012) – a strict cortical asymmetry exists between *back-*
118 *ward* connections (carrying predictions) and forward con-
119 *nections* (carrying prediction errors). Since forward
120 connections originate in superficial (II/III) pyramidal neu-
121 rons, and backward connections originate in deep (V/VI)
122 pyramidal neurons (Felleman and van Essen, 1991) this
123 asymmetry has a straightforward anatomical conse-
124 quence: prediction neurons reside in *deep layers*, and
125 error neurons in *superficial layers* (Fig. 1).

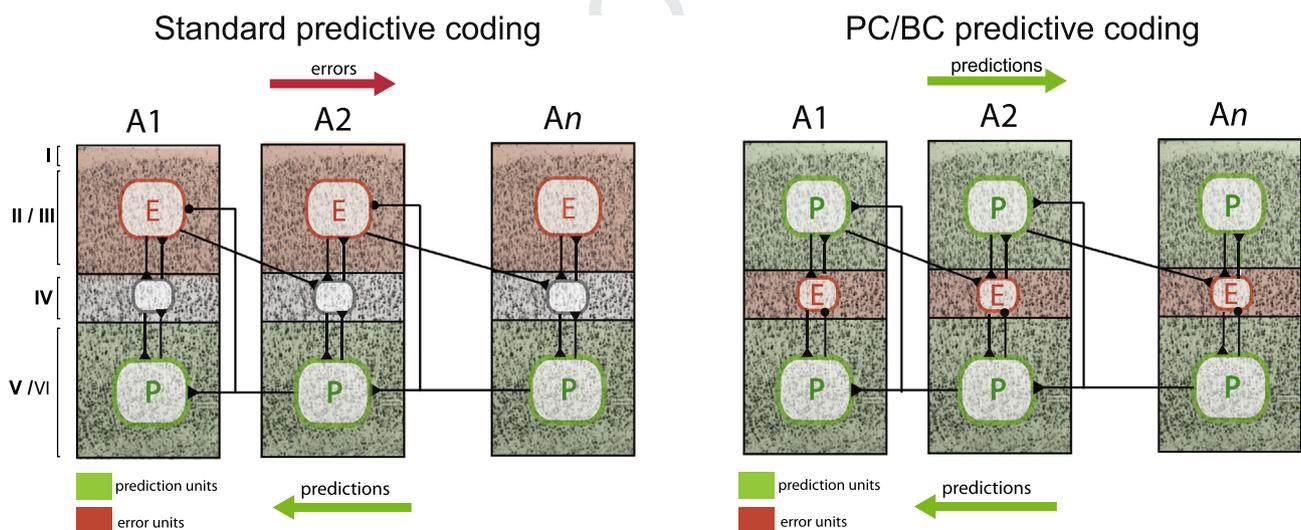


Fig. 1. Different arrangements of error and expectation neurons in auditory cortex implied by different formulations of Predictive Coding (PC). Columns denote hierarchically arranged cortical columns corresponding to primary (A1), secondary (A2) and higher order (An) auditory areas. In standard PC (left), errors flow upward and predictions downward; error units are therefore identified with superficial layers (II/III) and expectation units with deep layers (V/VI). Prediction units at higher levels can suppress error units at lower levels via (poly-synaptic) top-down inhibitory connections (black circles). In Biased Competition models of PC (Spratling, 2009; right), expectations flow upward and downward, error is computed at input layer IV, prediction units suppress error units only via intracolumnar inhibition, and top-down connections are fully excitatory (black arrows). Please note that this schematic is intended to illustrate differences in laminar profiles only. For simplicity, various details have been omitted, such as the distinction between excitatory and inhibitory populations, and between hidden causes and hidden states. For a more detailed exposition of the models, and possible physiological mappings, see Shipp (2016), Bastos et al. (2012), and Spratling (2009, 2017). Laminar image of auditory cortex was adapted from Winer (1985).

Note that this ‘standard model’ is not the only implementation of PC. Other models propose different arrangements, some dispensing with the functional asymmetry between forward and backward connections, and locating prediction and error neurons differently (Spratling, 2008a,b, 2010; See Fig. 1). However, all formulations assume that predictions and errors are computed by separate neurons in different cortical layers – as such, prediction and error responses are assumed to have distinct laminar profiles.

In PC, attention is formalized as a process that infers the level of predictability of sensory inputs. Again, models differ in details (Feldman and Friston, 2010; Rao, 2005; Spratling, 2008a,b, 2010) but all conceptualize attention as the weighting of sensory signals by their precision (inverse variance). The brain thus not only generates (first-order) predictions about the content of a signal, but also generates (second-order) estimates about its reliability. When this is low, deviations are down-weighted and may go unnoticed; when it is high, deviations are amplified and prioritized for further processing. Physiologically, this is thought to be implemented by the post-synaptic gain on superficial (error or prediction) neurons tuned to the attended dimension (e.g. feature-based or spatial attention).

Finally, different PC-variables are sometimes associated with different cortical rhythms: error cells are thought to propagate their messages forward via the gamma-band (> 30 Hz), while deep layers send downward predictions via lower, especially beta-band (12–30 Hz), frequencies (Arnal and Giraud, 2012; Bastos et al., 2012). Since this assumption is based on known oscillatory differences between forward and backward signals (e.g. van Kerkoerle et al., 2014) it only applies to standard PC, which postulates a strict functional asymmetry between backward connections (carrying predictions) and forward connections (carrying errors).

In sum, PC makes a number of key assumptions with clear empirical consequences:

- (1) Sensory cortex implements a hierarchical, generative model of the world: neurons at higher processing stages generate predictions that bias processing at lower levels.
- (2) Population responses (i.e. gross activity measured with MEG, EEG or BOLD) reflect (at least in part) ‘transient expressions of prediction error’ (Friston, 2005, p.829) – therefore, neural responses should be shaped by (hierarchically nested) expectations.
- (3) Prediction-generation and error-detection are implemented by separate neural subpopulations that reside in different cortical layers – as a consequence, prediction and error computations should have distinct laminar profiles.
- (4) Attention is the weighting of sensory input by its reliability – accordingly, the gain on upward projections should reflect (estimated) sensory precision.
- (5) In standard PC, top-down predictions and bottom-up errors have distinct oscillatory profiles: predictions are conveyed via lower frequencies (mostly beta) and (precision-weighted) prediction errors via higher frequencies (gamma).

In the next sections, we will evaluate each assumption in the light of recent evidence.

ANIMAL STUDIES

Prediction in auditory cortical neurons

Most animal research on auditory prediction and surprise focusses on Stimulus Specific Adaptation (SSA). SSA refers to the selective attenuation of responses to repeated (common) stimuli and can be seen as a single-cell analog of MMN. Although their exact relation remains debated, SSA is probably not a direct substrate of MMN, since the phenomena differ in latencies, NMDA-dependence, and sensitivity to certain regularities (Khouri and Nelken, 2015). There is a large literature on SSA, most of which is beyond the scope of this review as it does not address key features of PC such as prediction (but see Khouri and Nelken, 2015 for review). Interestingly, it is unclear whether SSA, despite what the name implies, is caused by simple adaptation. Ulanovsky et al. (2004) showed that SSA – here defined as the difference in responses to the same sound presented with different probabilities – depended not just on local context but also on a longer stimulus history, beyond the order of seconds at which habituation processes like synaptic depression are thought to occur. Moreover, SSA is observable for tones with frequency differences smaller than typical tuning curves, which also cannot be explained by models of synaptic habituation (Taaseh et al., 2011; Yaron et al., 2012).

Recently, Rubin et al. (2016) re-analyzed the data from Ulanovsky et al. (2004), in a first attempt to quantify the longer-term dependencies. Anesthetized cats were exposed to ‘Bernoulli sequences’ with two tones occurring independently with a fixed probability. The authors reasoned that some representation of (long-term) stimulus history influenced responses; moreover, this representation was not a one-to-one copy but a reduced representation. Assuming that only stimulus probability was represented, rather than transitional probability (but see Meyniel et al., 2016; Mittag et al., 2016) the authors computed the predictive power of representations reduced to a different degree. The key assumption here was that responses reflected prediction error, expressed as negative log probability. The prediction error account offered good fits, explaining up to 50% of observed variability. Interestingly, representations incorporating less than 10 preceding stimuli (7.3 s) were almost never in the top 10% with the most power. The authors concluded that neurons in A1 signal prediction errors, based on reduced representations incorporating long-term stimulus history ‘to generate predictions about the future’ (2016, p.2). Although the authors are agnostic about the underlying mechanism – which may or may not resemble schemes envisioned by PC – the interpretation forms a departure from earlier accounts of SSA, which (as the name suggests) tend to focus on stimulus-driven explanations such as synaptic depression.

More fundamental insights are presented by Gill et al. (2008) who explored surprise as a model for auditory receptive fields. At several levels in the Zebra Finch audi-

254 tory hierarchy, the authors compared three receptive field
255 models: first, a traditional approach modeling neurons as
256 responding to specific spectrotemporal patterns of intensi-
257 ties (STRF); secondly, a derivative approach, modeling
258 *changes* in intensities; finally, a model describing neurons
259 as responding to *surprise*, quantified as the inverse condi-
260 tional probability of a range of frequencies, given the pre-
261 ceding frequencies, based on naturalistic Zebra Finch
262 song. This ‘surprise model’ substantially outperformed
263 traditional models. Interestingly, its advantage depended
264 on hierarchical level: in area MLD (homolog of inferior col-
265 liculus) models did not differ significantly. In field L (homo-
266 log of thalamorecipient neurons in A1) surprise was 20%
267 better than STRF models on average. And in CLM (homo-
268 log of higher-order auditory cortex) the surprise model
269 performed a striking 67% better on average. The authors
270 concluded that expectations are increasingly important at
271 higher levels, leading to an ever sparser neural code that
272 eventually only propagates input not expected on the
273 basis of preceding input. The authors stress that these
274 ‘expectations’ were based on very short preceding time
275 windows (3–7 ms), and that the effect is thus not a direct
276 substrate of high-level (conscious) ‘surprise’. The paper is
277 also silent on how the expectations are implemented. But
278 it does show that, at a fundamental level, ‘*expectations . . .*
279 *form a key part of the [auditory] neural code*’ (2008 p.,
280 218).

281 The laminar profile of prediction

282 Laminar differences between prediction and error signals
283 (*Assumption 3*) are a key tenet of predictive coding
284 theories. This important notion remains woefully under-
285 studied; we have only found two papers relevant to this
286 assumption. The first is by [Szymanski et al. \(2009\)](#), who
287 studied the laminar profile of SSA in rat auditory cortex
288 using an oddball paradigm. To the best of our knowledge,
289 this is the first auditory study comparing laminar profiles
290 for expected and unexpected stimuli. Remarkably, their
291 results seemed opposite to what PC predicts: the authors
292 found no clear differences between layers. In all layers,
293 neurons responded more strongly to deviants than to
294 the same stimulus when it was a standard (SSA). If any-
295 thing, this difference seemed to be *stronger* for deeper
296 layers. If SSA is an expression of prediction error, this
297 effect seems to contradict PC schemes that identify deep
298 layers as ‘prediction layers’. It is worth noting, however,
299 that the rats were anesthetized with ketamine, an
300 NMDA-antagonist that, probably by blocking NMDA-
301 dependent plasticity, impairs MMN ([Umbricht et al.,](#)
302 [2002](#)) and abolishes global mismatch responses ([Uhrig](#)
303 [et al., 2016](#)) leading to a pattern of effects that has been
304 interpreted ([Strauss et al., 2015](#)) as a ‘disruption of pre-
305 dictive coding’ (see Section Expectation and surprise
306 along the auditory hierarchy).

307 A recent study in awake animals did find strong
308 laminar-specific effects. The authors ([Rummell et al.,](#)
309 [2016](#)) trained mice to press a lever that generated noise
310 bursts. Every 5–10 s the bursts were also generated ran-
311 domly, allowing the researchers to compare responses to
312 the same sound when it was self-generated and when it
313 occurred randomly ([Fig. 2A](#)). Clear attenuation of

314 responses to self-generated sounds was observed in pri-
315 mary auditory cortex, replicating non-invasive work in
316 humans ([Baess et al., 2009](#)). The predictive, stimulus-
317 specific nature of this effect was confirmed when the
318 authors replaced the noise bursts by pure tones that
319 had one frequency in 75% of trials and another in 25%.
320 Responses to common (‘expected’) self-generated
321 sounds were attenuated more strongly – suggesting that
322 auditory cortex embodied a stimulus-specific expectation
323 of the generated sound. Multi-site recordings from audi-
324 tory thalamus, auditory cortex and hippocampus revealed
325 increasing attenuation, often resulting in near-silencing in
326 hippocampus (see [Fig. 3](#)). This again points to a familiar
327 picture of the auditory hierarchy in which predictable com-
328 ponents are progressively ‘filtered out’.

329 Intriguingly however, the authors found that
330 attenuation was much *stronger* in deep layers. This
331 again seems to contradict the idea that deep layers
332 encode predictions, since it should be the *error* that is
333 suppressed. Conclusions should be drawn with care,
334 however, since predicting the sensory consequences of
335 motor commands may be very different from sensory
336 prediction in general, so that evidence for the one (see
337 [Eliades and Wang, 2008](#); [Keller and Hahnloser, 2009](#)
338 for earlier evidence for auditory efference copy) is not
339 necessarily evidence for the other.

340 Finally, [Jaramillo and Zador \(2011\)](#) studied expecta-
341 tion in rat auditory cortex. Rats were presented a train
342 of short pure tones containing a frequency-modulated tar-
343 get. The target, which appeared either ‘early’ (450 ms) or
344 ‘late’ (1500 ms), signaled if the correct (rewarded)
345 response was right or left. Expectation was manipulated
346 over blocks in which the target appeared early in 85% of
347 trials and late in 15%, or *vice versa*. Behaviorally, rats
348 responded faster and more accurately to targets appear-
349 ing at the expected time-window. Expectation also modu-
350 lated single-unit and LFP responses, and this modulation
351 correlated with performance. For both the preceding stimu-
352 lus and the target itself, expectation *increased* rather
353 than attenuated the neural response, which apparently
354 contradicts the notion that neurons signal surprise. How-
355 ever, the stimulus of which expectancy was manipulated,
356 was also the *target*. As such, *prediction* (what is likely?)
357 and *attention* (what is relevant?) are confounded. This
358 confound characterizes many common paradigms,
359 including the classic Posner task ([Posner, 1980](#)), where
360 attention is controlled by manipulating probability
361 ([Fig. 2B](#)). In such situations, PC makes similar predictions
362 as conventional accounts of attention: enhanced gain on
363 the relevant (informative) feature, which is prioritized for
364 processing. Hence, to distinguish assumptions of PC,
365 attention and prediction must be manipulated indepen-
366 dently (see Section Expectation and surprise along the
367 auditory hierarchy).

368 Discussion

369 In sum, animal-model studies relevant to the assumptions
370 of predictive coding are scarce and show mixed results.
371 None of the discussed studies explicitly tested PC,
372 which may contribute to the inconclusiveness of the
373 results. Nevertheless, they report some remarkable

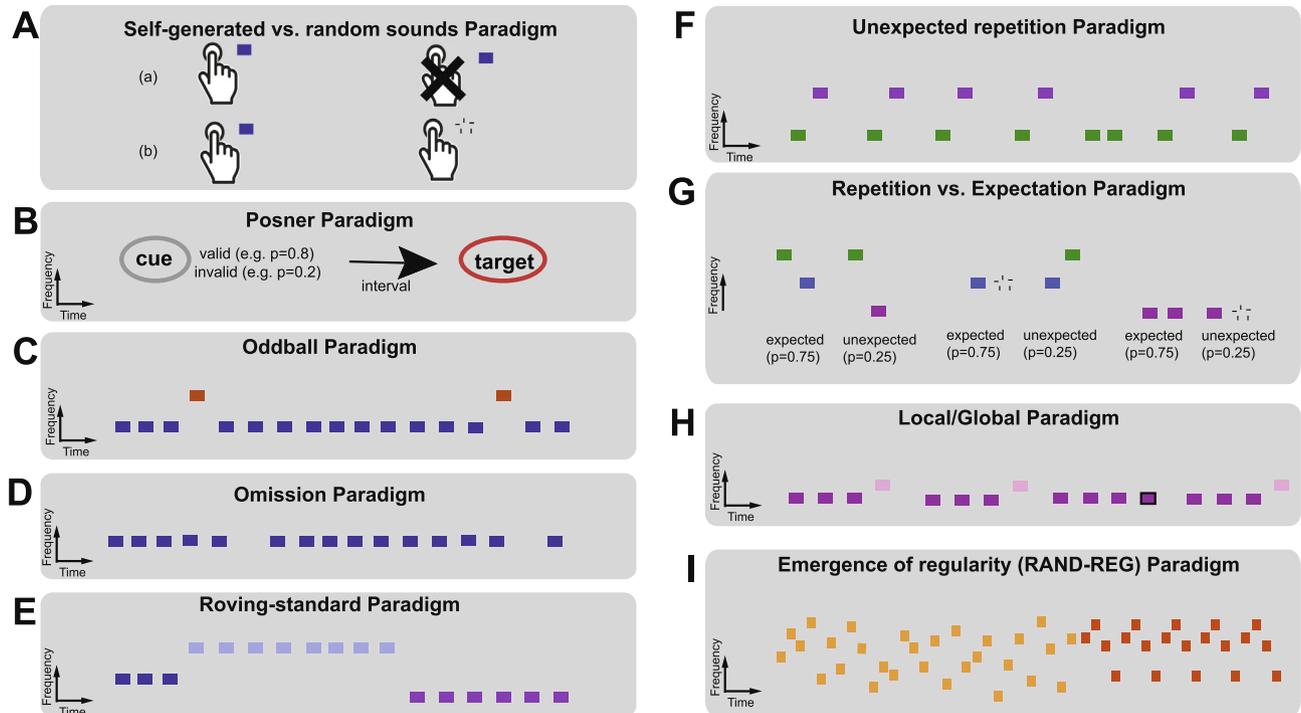
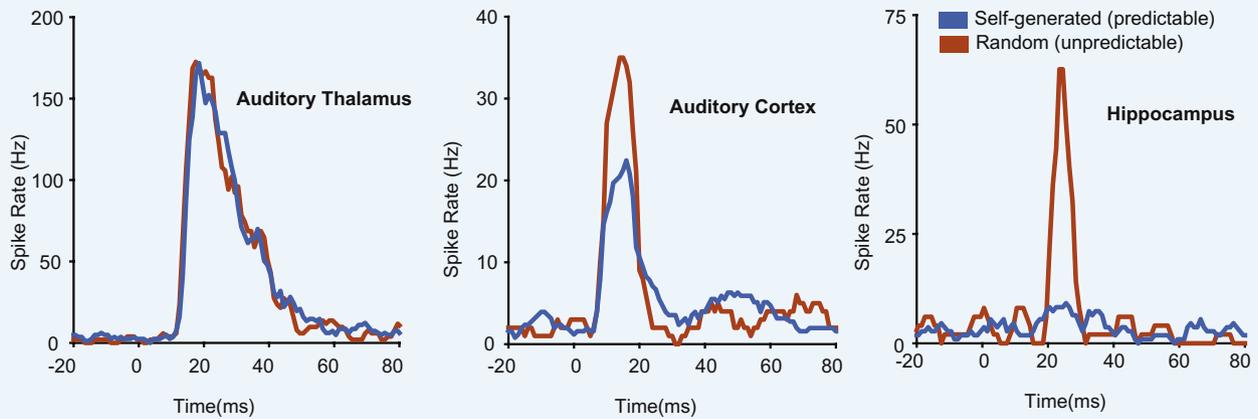


Fig. 2. Paradigms often used in the literature to study the effects of context and predictability on behavioral and brain responses. In the schematic representations below colored squares represent sounds. (A) The ‘Self-Generated vs. Random sounds’ paradigm compares responses to sounds when they are self-generated (triggered by a button press; and therefore predictable) or randomly generated by a computer (and therefore unpredictable; a). Another version of the paradigm (b) compares self-generated sounds (triggered by a button press) to omissions (when the participant pressed the button but no sound was presented). (B) The ‘Posner paradigm’ is a class of experimental designs where a ‘cue’, which can be implemented as specific stimulus or a context which is induced during the experimental session, that predicts the target with a certain probability. The paradigm therefore allows to measure responses to the target as a function of its predictability. (C) The standard MMN Oddball paradigm involves the presentation of a repeating standard tone, occasionally replaced by a deviant tone. (D) The standard MMN omission paradigm is similar to the Oddball paradigm except the deviant tone is replaced by silence. (E) The Roving standard paradigm is a variation of the oddball paradigm that replaces the deviant stimulus with a variable standard. After a number of repetitions, the standard changes, creating a ‘deviant’ that *becomes* a ‘standard’ – while remaining physically identical. (F) The unexpected repetition paradigm consists of pairs of sounds that are infrequently replaced by a repetition. The schematic here shows a simple version of the paradigm where the tone pairs consist of the same sounds, but instances where different pairs are presented are also used. (G) The ‘Repetition vs. Expectation’ paradigm is used to dissociate the effects of prediction from simple effects of repetition. The paradigm depicted here was used in [Todorovic and de Lange \(2012\)](#). The stimulus set consisted of 3 different tones (illustrated here by the use of different colors) arranged in pairs but such that the first tone in a pair was predictive of the second one. For example tone1 (green) was predictive of tone2 (blue) in 75% of the trials but was occasionally (in 25% of the trials) followed by tone2 (purple). Tone2 (blue) was predictive of an omission but which was replaced in 25% of the trials by tone 3 (green), etc. (H) The Local/Global paradigm is designed to dissociate responses to local deviants from responses to global deviants. In the example depicted here the stimulus consists of ‘standard’ (commonly occurring) and ‘oddball’ (rarely occurring) sequences. The last tone in each ‘standard’ sequence is a local deviant; In contrast, ‘global deviance’ is manifested here by the *absence* of change. A similar approach with expected and unexpected tone omissions is also commonly used. (I) The ‘Emergence of regularity’ (RAND-to-REG) paradigm introduced by [Barascud et al. \(2016\)](#) is based on rapid tone-pip sequences which contain transitions from a random (RAND) frequency pattern (in yellow) to a regularly repeating (REG; predictable) frequency pattern (in orange). In this example the REG pattern consists of a cycled sequence of 4 different tones.

374 findings. Firstly, in support of *Assumption 2*, expectation
 375 appears to shape neural responses in auditory cortex.
 376 Surprise – both task-based at timescales of several
 377 seconds ([Rubin et al., 2016](#)), and species-based at time-
 378 scales of milliseconds ([Gill et al., 2008](#)) – offers a good
 379 model for neural responsiveness. The methodological differ-
 380 ences between these studies, and the fact that both did
 381 not address the mechanisms of prediction, unfortunately
 382 limit their conclusiveness with respect to PC. However,
 383 both studies make the conceptual shift from charac-
 384 terizing neurons as encoding bottom-up data features,
 385 to encoding hypotheses or predictions, and propagating
 386 only the divergence from these predictions. In support of
 387 *Assumption 1*, there is also evidence for the idea that
 388 the effects of expectations are hierarchical, in the sense

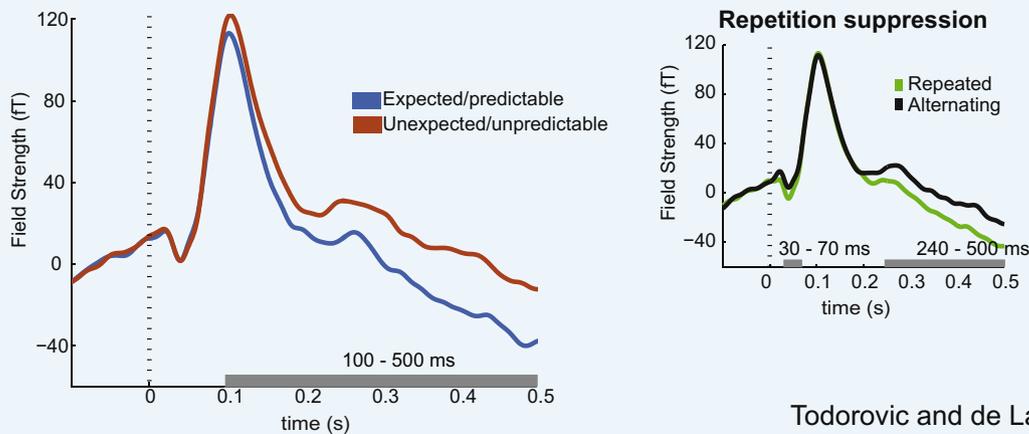
that expected components seem to be progressively fil- 389
 390 tered out ([Rummell et al., 2016](#); [Gill et al., 2008](#)). Attention,
 391 as shown by [Jaramillo and Zador \(2011\)](#), can
 392 influence processing in A1 in an anticipatory way – how-
 393 ever, it remains unclear whether this form of modulation
 394 is in line with attentional modulation as described by
 395 PC. Finally, the two studies that investigated laminar differ-
 396 ences in processing of expected versus unexpected
 397 stimuli – a signature characteristic of PC – found (under
 398 ketamine anesthesia) no distinct laminar profiles and
 399 (using self-generated sounds) strong expectation sup-
 400 pression in the deep rather than superficial layers of cor-
 401 tex. Although methodological issues prevent strong
 402 conclusions from being drawn, the animal-model literature
 403 contains fascinating results that call for more experiments

Expectation suppression along the auditory hierarchy in an animal model



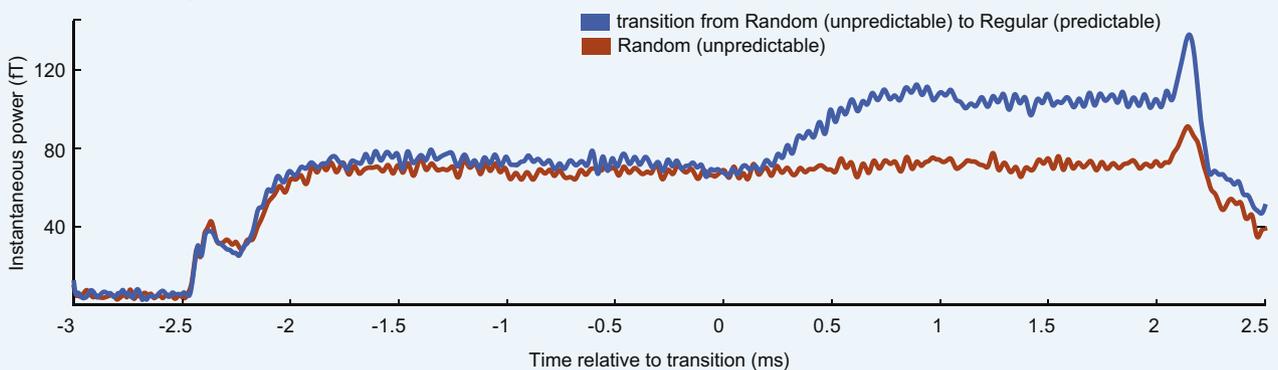
Rummell et al., (2016)

MEG responses in humans reflect expectation suppression



Todorovic and de Lange (2012)

MEG responses in humans to emerging predictable structure in rapid sound sequences



Barascud et al., (2016)

404 in awake animals, since only studies of this type can ultimately
405 confirm or falsify key assumptions of predictive
406 coding.

407 HUMAN IMAGING AND ELECTROPHYSIOLOGY

408 Predictive coding and MMN

409 Human auditory studies on predictive coding often use
410 some variation of the Mismatch Negativity or ‘MMN’
411 paradigm. MMN is measured using a method in which a
412 sequence of stimuli (typically a repeated tone)
413 establishes a regularity that is violated by a ‘deviant’
414 stimulus (‘oddball paradigm’; Fig. 2C). MMN is the
415 negative component of a *difference wave*, obtained by
416 subtracting the ‘standard’ from the ‘deviant’ response,
417 and is found at 100–250 ms.

418 Traditionally, two main hypotheses on MMN exist.
419 According to the memory-based hypothesis (Näätänen
420 et al., 1978; Winkler and Czigler, 1998), MMN is gener-
421 ated by a system comparing auditory inputs with a mem-
422 ory template. When a difference is detected, the system
423 signals an error, and adjusts the template. According to
424 the adaptation hypothesis (May et al., 1999;
425 Jääskeläinen et al., 2004; May and Tiitinen, 2010) cells
426 tuned to repeated ‘standard’ tones simply adapt – due
427 to passive processes such as synaptic depression – while
428 neighboring inputs tuned to ‘deviants’ remain unadapted
429 and elicit stronger responses. By implication, the
430 oddball-evoked MMN is not a separate evoked potential
431 but rather a delayed and attenuated N1, that *appears* sep-
432 arate only in the difference wave.

433 In this dispute, PC takes a middle ground position.
434 Like all memory-based accounts, PC interprets MMN as
435 a mismatch signal – a mismatch, however, between the
436 input and a *prospective* prediction, rather than a
437 *retrospective* template. But like the adaptation
438 hypothesis, PC considers MMN not as a separate
439 evoked response, but simply as an amplified contrast
440 between an expected (standard) response and a
441 surprising (novel) response. In the adaptation
442 hypothesis, however, both the response and its
443 suppression are stimulus-driven: there is no error signal.
444 Under PC, *every* response is an expression of error and
445 can be larger or smaller depending on predictions. This
446 last point – the dependence on predictions – is also
447 what makes PC considerably less parsimonious than
448 the adaptation hypothesis. Applying Ockham’s razor, we

449 can only consider evidence in favor of PC if it cannot be
450 explained by a simpler process – which, in addition, is
451 often simple adaptation. Beyond the simple oddball
452 paradigm, the MMN literature has shown that listeners
453 are sensitive to the violation of potentially very complex
454 patterns (see Paavilainen et al., 2007 for review), which
455 is usually interpreted as evidence for the exquisite sensi-
456 tivity of auditory cortex to patterns in sound. Unfortu-
457 nately, most studies with an explicit focus on predictive
458 coding in the auditory modality (see reviewed below) have
459 used the simple oddball paradigm or its variations in
460 which predictability, or regularity, is manipulated by repe-
461 tition, which makes adaptation all the more difficult to
462 exclude.

463 Repetition suppression – Adaptation or expectation?

464 The neuroimaging analog of the physiological
465 phenomenon of adaptation is repetition suppression
466 (RS). As reviewed by Grill-Spector et al. (2006), multiple
467 mechanisms for RS have been proposed. We can distin-
468 guish between mechanisms that explain RS via passive
469 adaptation effects, sometimes called ‘neural fatigue’,
470 and accounts that interpret it as a signature of increased
471 processing *efficiency*. PC belongs to the second type: it
472 ascribes the suppression not only to the repetition itself,
473 but also to the expectations it induces. Interesting support
474 for this account comes from Costa-Faidella et al. (2011)
475 who recorded EEG responses in a roving standard para-
476 digm (Fig. 2E). This is a variation of the oddball paradigm
477 that replaces the deviant stimulus with a variable stan-
478 dard. After a number of repetitions, the standard changes,
479 creating a ‘deviant’ that *becomes* a ‘standard’ – while
480 remaining physically identical. The authors used two con-
481 ditions, with predictable and unpredictable timing. In the
482 predictable condition, Inter Stimulus Intervals (ISI) were
483 fixed. In the unpredictable condition, ISIs varied ran-
484 domly. The suppressive effect of repetition (reduction of
485 the response to the standard tone) was reduced in the
486 condition with unpredictable timing. Because the average
487 ISI and number of stimulations were identical between
488 conditions, this suggests that repetition suppression is
489 (at least in part) modulated by predictability.

490 Also in a roving paradigm, Lieder et al. (2013) used
491 computational modeling to compare prediction and adap-
492 tion. For each stimulus presentation they calculated the
493 ‘MMN amplitude’, by subtracting the final (‘standard’) pre-

←
Fig. 3. Brain responses to predictable and unpredictable sounds. Top: Progressive attenuation of responses to self-generated sounds at different cortical regions of a mouse model (Rummell et al., 2016). The progressive weakening often resulted in near-silencing in hippocampus, and suggests the existence of an increasingly sparse code, in which eventually only non-predicted components are propagated for further processing. Middle: Repetition suppression and expectation dissociated in time (Todorovic and de Lange, 2012): Repetition (but not expectation) attenuated the early MEG responses (40–60 ms). Expectation (but not repetition) attenuated the response at an intermediate latency (100–200 ms). Both repetition and expectation affected the late response (200–500 ms). Gray horizontal bars under the figure indicate the time intervals with a significant difference between conditions. Bottom: In contrast to the oft reported attenuation of brain responses to predictable sounds, Barascud et al. (2016) found the opposite effect: Brain responses to rapid tone sequences that transitioned from a random to repeating pattern are manifest as a substantial increase in the MEG response. This finding demonstrates that the brain appears to encode the state (RAND vs REG) rather than the transition (as in e.g. MMN). As discussed in Barascud et al. (2016) the DC shift appears to vary consistently with the predictability (negentropy) of the ongoing stimulus pattern.

494 sentation from the earlier ('deviant') presentations. The
495 authors then compared different models to explain trial-
496 by-trial fluctuations in this MMN amplitude. The first model
497 was an adaptation model. This model was 'phenomenolo-
498 gical' in the sense that it made no assumptions on the
499 *mechanism* behind adaptation, but simply embodied fluctu-
500 ating responsiveness of populations tuned to different
501 frequencies. This 'phenomenological' approach was con-
502 trasted to a computational approach in which MMN ampli-
503 tudes were compared to several parameters in a hidden
504 Markov model which tracked transition probabilities.
505 Overall, parameters of the computational model explained
506 the fluctuations better than adaptation. Interestingly, most
507 variance was explained not by prediction error, but by
508 model-adjustment. Together, the authors write, this sug-
509 gests that attenuation observed in a roving paradigm is
510 best explained as a form of *learning*, rather than as
511 adaptation.

512 More modeling results are found in [Wacongne et al.](#)
513 [\(2012\)](#) who present a PC model of A1. Contrary to Lieder
514 et al. (2011) and the DCM studies (see Section Effective
515 connectivity – clues from DCM) [Wacongne et al. \(2012\)](#)
516 specified their model at the level of individual spiking neu-
517 rons, thus committing to a much more detailed implemen-
518 tation of PC. The model comprised two cortical columns,
519 each selectively responsive to a different tone (A or B).
520 Crucially (and unlike standard PC) error units are located
521 in the thalamorecipient granular layer. In that same error
522 layer, GABA-ergic neurons receive excitatory input from
523 predictive units in layer II/III, effectively subtracting the
524 prediction from the incoming input, resulting in an error
525 term. This error term is sent to the predictive layers,
526 where it forms a memory trace used to adapt the internal
527 model via spike-timing dependent plasticity at NMDA-
528 weights. Using the sum of postsynaptic currents in each
529 layer as a proxy for the ERP, [Wacongne et al. \(2012\)](#)
530 show that this set-up – intentionally lacking synaptic habi-
531 tuation mechanisms – can account for an array of phe-
532 nomena from the MMN literature, such as the
533 parametric modulation of MMN amplitude by stimulus
534 probability (e.g. [Sams et al., 1983](#)); MMN to unexpected
535 *repetition* (Fig. 2F; e.g. [Saarinen et al., 1992](#); [Horváth](#)
536 [and Winkler, 2004](#)); MMN to omission (Fig. 2D; e.g.
537 [Yabe et al., 1997](#); [Rajj et al., 1997](#)); and blindness to con-
538 text ([Wacongne et al., 2011](#), see below).

539 To further dissociate adaptation and PC based
540 accounts, [Wacongne et al. \(2012\)](#) performed a MEG
541 experiment on the 'repetition MMN'. Participants listened
542 to tone-pairs that were overwhelmingly alternations
543 (AB), and only rarely repetitions (AA; Fig. 2F). PC predicts
544 that an unexpected repetition (AA) should evoke a stron-
545 ger response – an *inversion* of the standard MMN. The
546 adaptation-based explanation ([May and Tiitinen, 2010](#))
547 suggests this reflects adaptation at higher order neurons,
548 tuned to the AB tone-pair. To exclude this possibility,
549 [Wacongne et al. \(2012\)](#) inserted an interval of 10 s
550 between each pair – much longer than the recovery time
551 of synaptic depression. In every individual participant, AA
552 indeed elicited an MMN, while no difference between BA
553 and BB was observed. Although this result seems highly

suggestive, a replication with a larger number of
participants is needed, since [Wacongne and colleagues](#)
tested only 5.

Using a similar paradigm, [Todorovic et al. \(2011\)](#) mea-
sured RS for expected and unexpected repetitions.
Expectancy was manipulated in blocks where either
75% of stimuli were tone-pairs and, 25% single tones
(repetition expected) or *vice versa* (repetition unex-
pected). Clear RS was observed in the 100–500 ms
range, that was strongly reduced in the unexpected con-
dition, suggesting that RS itself might comprise an expec-
tancy effect. However, since the blockwise manipulation
affected the overall occurrence of the tones – and the
authors used an inter-trial interval of 4–6 s – the effect
could, theoretically at least, be explained by passive
adaptation.

Expectation and surprise along the auditory hierarchy

In a follow-up study, [Todorovic and de Lange \(2012\)](#)
addressed this issue by adding an extra hierarchical level
of expectations, that allowed them to manipulate repeti-
tion and expectation orthogonally (Fig. 2G). Each trial
consisted of either an identical or non-identical tone-
pair, or a single tone. Orthogonally to this, the frequency
of the first tone predicted that of the second tone with a
high validity. Using MEG, the authors observed a dissoci-
ation: repetition (but not expectation) attenuated the early
response (40–60 ms) and expectation (but not repetition)
attenuated the intermediate response (100–200 ms). By
contrast, both repetition and expectation affected the late
response (200–500 ms; Fig. 3). This suggests that RS
might be non-unitary, consisting of a late stage, which
reflects the effects of expectation, and an early stage,
which does not (cf. [Grotheer and Kovács, 2015](#)). How-
ever, the results may also be compatible with PC, if one
casts repetition as a special, low-level form of expectation
(cf. [Aukstulewicz and Friston, 2015b](#)).

Similar findings were reported by [Wacongne et al.](#)
[\(2011\)](#), who recorded MEG and EEG responses to viola-
tions of local and global regularities (Fig. 2H). Participants
passively listened to stimuli consisting of five tones, of
which the first four were always identical and the last
one varied. Within each block, a particular variation (e.g.
'xxxxY') was dominant, occurring 75% of trials. In the
remaining 25% of trials, the last tone was replaced by
either a deviation (e.g. 'xxxxX' within 'xxxxY' blocks) or
an omission ('xxxx_'). The authors found that local devi-
ants (i.e. 'xxxxY' even when it was the common stimulus)
were always accompanied by a measurable MMN (at 80–
150 ms), but that in xxxX blocks (where the *local* deviant
'xxxxY' was also *globally* unexpected) this deflection was
larger. By contrast, global deviant responses were found
at later latencies (150–600 ms), although no interaction
was reported for this dissociation. Note that, for xxxX
blocks, a global deviant amounts to an unexpected repe-
tition eliciting a *stronger* response than an expected alter-
nation – an inversion of the MMN.

Recently, [Dürschmid et al. \(2016\)](#) provided more evi-
dence for a hierarchical organization of mismatch signals,

613 dissociating not time-scales but brain regions, using high-
614 gamma (> 60 Hz) activity as an index of local spiking.
615 Dürschmid et al. (2016) were able to measure high-
616 gamma components using ECoG recordings from
617 patients with frontal and temporal electrodes, who listened
618 to predictable and unpredictable deviants embedded in an
619 uninterrupted train of tones. In the predictable
620 condition, the deviant tone (550 Hz) always occurred after
621 four consecutive standard tones (500 Hz), rendering the
622 deviant ‘globally’ predictable. In the unpredictable condition,
623 the deviant tone occurred randomly after at least
624 three presentations of the standard tone, rendering the
625 deviant fully unpredictable. The authors found no main
626 effect of block type, but they did find an interaction: high
627 gamma was found for unpredictable, but not predictable
628 deviants at frontal electrodes, while at temporal electrodes
629 both deviant-types elicited high-gamma responses. The
630 authors interpreted this as demonstrating that frontal cortex
631 monitors ‘the bigger picture’. This interpretation is compatible
632 with the source reconstruction results by Wacongne et al. (2011),
633 who also found that global (but not local) deviants activated
634 a broad fronto-parietal network. However, this claim could
635 have been stronger had Dürschmid et al. (2016) manipulated
636 local and global regularities independently, instead of comparing
637 repeating versus random patterns.

639 Strauss et al. (2015) did present such an independent
640 manipulation. In the same MEG-EEG paradigm as used
641 by Wacongne et al. (2011), the authors showed that
642 late-latency responses to global deviants disappeared
643 categorically in all stages of sleep, concluding that predictive
644 coding was ‘disrupted’. This is remarkable, because the
645 MMN persists during sleep (Scutthorpe et al., 2009)
646 and even coma (Fischer et al., 2000; but see Dykstra
647 and Gutschalk, 2015). However, Strauss et al. (2015)
648 demonstrate that the persisting ‘sleep MMN’ is strongly
649 reduced and lacks sustained fronto-parietal activity. Over
650 and above these differences in degree, sleep-MMN also
651 seemed to be *qualitatively* different. Strauss et al.
652 (2015) showed this by training a classifier to distinguish
653 local standards and deviants. When trained and tested
654 on responses recorded during wakefulness, the algorithm
655 reliably distinguished signals from early (76 ms) to late
656 (620 ms) latencies. However, when the classifier – trained
657 on wakefulness data – was tested on sleep data, it only
658 generalized to early (76–100 ms) and late (212–588 ms)
659 signals. For signals from the MMN latency (100–
660 200 ms) it did not generalize at all, and failed to perform
661 better than chance. The authors interpret this result as
662 new evidence for the idea that MMN might be a consequence
663 of several independent processes: an automatic process
664 arising from passive adaptation (May and Tiitinen, 2010)
665 and therefore persistent under sleep, as well as an independent
666 process arising from predictive coding.

668 Interestingly, the effects of sleep were corroborated in
669 the realm of anesthesia. Uhrig et al. (2014) had earlier
670 reported the first neural signature hierarchical novelty
671 responses (potentially an index of PC) in non-human
672 animals, using primate fMRI in macaque. They found that
673 only globally deviant sequences recruited a large

674 fronto-parietal network known in humans as the neuronal
675 workspace (Dehaene et al., 1998). Recently, Uhrig et al.
676 (2016) repeated the experiment under varying degrees
677 of anesthesia. Both anesthetics (propofol and ketamine)
678 weakened local and distorted global mismatch responses.
679 Ketamine was especially powerful, effectively abolishing
680 the global mismatch effect. Since both plasticity
681 (Collingridge and Bliss, 1987) and intra-regional feedback
682 (Self et al., 2012) are thought to be NMDA-dependent,
683 and ketamine impairs MMN even at light dosages
684 (Umbricht et al., 2002), this is perhaps unsurprising. However,
685 ketamine is a popular anesthetic, used by three of
686 the five animal studies here reviewed (Szymanski et al.,
687 2009; Jaramillo and Zador, 2011; Rubin et al., 2016).
688 Since it abolishes global mismatch responses, and the
689 persisting responses may be qualitatively different
690 (Strauss et al., 2015), these findings underline that future
691 studies of PC should avoid the use of ketamine – and, ideally,
692 of anesthesia altogether.

693 Finally, Lecaigard et al. (2015) manipulated (global)
694 predictability of auditory deviants, but found no hierarchical
695 effects. Deviant predictability affected ERP amplitudes
696 at early (< 70 ms), MMN (100–250 ms) and late
697 (> 300 ms) latencies. Puzzlingly, however, the biggest
698 effect of global predictability was found at the earliest
699 time-window (< 70 ms), where the MMN was completely
700 abolished *only* in the *globally* predictable condition; an
701 effect which stands in contrast to other studies on hierarchical
702 PC and hierarchical deviance detection more generally
703 (Grimm and Escera, 2012; Cornella et al., 2012;
704 Escera and Malmierca, 2014).

705 Altogether, hierarchy is central to PC and converging
706 evidence now demonstrates that effects of prediction
707 depend on hierarchical processing level. Nevertheless,
708 some ambiguities remain. A first issue is whether
709 hierarchically ‘high’ and ‘low’ effects reflect a single
710 mechanism. Some human electrophysiology studies
711 claim they do not (Todorovic and de Lange, 2012;
712 Strauss et al., 2015; López-Caballero et al., 2016). These
713 studies associate effects on early processing with passive
714 adaptation (e.g. the early, sleep-persistent MMN in
715 Strauss et al., 2015) and effects on later, ‘higher’ processing
716 with prediction (see also Grotheer and Kovács, 2015).
717 However, this hard dichotomy seems at odds with results
718 from animal electrophysiology which report prediction
719 effects already at A1 (Rubin et al., 2016; Rummell et al.,
720 2016; Gill et al., 2008; Ulanovsky et al., 2004). What
721 adds to the ambiguity is that most studies used repetitions
722 or Bernoulli sequences to manipulate prediction, causing
723 expectation and adaptation to be confounded. A second,
724 more subtle ambiguity is whether the discussed interactions
725 between hierarchy and prediction constitute evidence
726 for *hierarchical prediction*, in the sense of hierarchical
727 Bayesian inference. Theoretically, interactions as those
728 in the animal literature – showing that prediction effects
729 become stronger at higher hierarchical levels (Rummell
730 et al., 2016; Gill et al., 2008) – do not necessarily
731 support the notion of hierarchically nested predictions,
732 which would require a task which manipulates multiple,
733 nested (or hierarchically dependent) regularities – as
734 only few studies have done so far. Accordingly, while

735 there is clear evidence for the effect of expectations on
736 responses (Assumption 2) and suggestive clues for hier-
737 archical organization of expectations (Assumption 1) pro-
738 gress will now depend on studies using stimuli with
739 multiple nested regularities, and which manipulate expect-
740 ation in a way not confounded by adaptation.

741 **Hearing silences: Omission as a window into** 742 **prediction**

743 When omitting a highly expected sound such as a tone in
744 a beat, listeners can ‘hear’ the absence. In such
745 circumstances, neural responses time-locked to the
746 omitted sound have been observed (Yabe et al., 1997;
747 Raji et al., 1997; Fig. 2D). These ‘omission responses’
748 offer an appealing vantage point to study top-down pre-
749 diction decoupled from bottom-up input, and have
750 become a popular method for studies on predictive
751 coding.

752 Theoretically, detecting silences could happen either
753 retrospectively (by comparing perceptual input and
754 memory template *after* the input is processed) or
755 prospectively (by directly matching *predictions* to input,
756 as proposed by PC). Bendixen et al. (2009) attempted
757 to dissociate these possibilities. Participants listened to
758 isofrequent tone-pairs of which either the first or the sec-
759 ond tone was occasionally omitted. If the second tone
760 was omitted, it could nonetheless be predicted by the first
761 tone (‘predictable’ condition). But if the first tone was omit-
762 ted, its identity could only be ‘restored’ after hearing the
763 second tone (restorable condition). The authors com-
764 pared evoked responses to a control condition in which
765 the tones were neither predictable nor restorable. When
766 comparing the amplitudes of the early component (up to
767 50 ms post tone/omission onset) the authors found omis-
768 sion responses in the predictable condition which were
769 essentially identical to responses evoked by actually pre-
770 sented tones, but not in the restorable condition. This was
771 interpreted as pre-activation of the sensory representa-
772 tions of the predicted tones. The authors concluded that
773 auditory expectation works prospectively *and not retro-*
774 *spectively*. However, since they looked for main effects
775 at very short latencies (< 50 ms post onset, identical to
776 the duration of the tone), and focused exclusively on
777 evoked (as opposed to not time locked) responses the
778 analysis may have been biased to finding prospective
779 pre-activations, and not retrospective memory effects.

780 Hughes et al. (2001) took a similar approach to test
781 whether change-detection involves prediction. Patients,
782 undergoing intracranial recordings from temporal cortex,
783 performed an oddball paradigm with tones or tone-pairs
784 as standards and silences as oddballs. Strikingly, in all
785 patients, channels firing to tones *also* fired to omissions,
786 often more strongly. Furthermore, 5 of 10 patients exhib-
787 ited ‘omission selective’ channels that *only* responded to
788 unexpected omissions, and to other unexpected stimuli
789 like bird-chirps. Finally, and contrary to other demonstra-
790 tions of omission responses (Raji et al., 1997; Chennu
791 et al., 2016) the effects seemed wholly independent of
792 attention. The omission-selective channels may have
793 been the first recordings of error-units. Unfortunately,
794 Hughes et al. (2001) did not reconstruct the exact location

795 or depth of their electrodes, other than being associative
796 (non-primary) auditory cortex, which makes the striking
797 findings somewhat anecdotal.

798 A different approach is described in SanMiguel et al.
799 (2013a,b), who used self-generated sounds to elicit omis-
800 sion responses (Fig. 2A). Participants were asked to
801 press a button every 600–1200 ms, after which a sound
802 was generated in 88%, 50% or 0% of trials. To control
803 for motor activity, the response after button presses that
804 were never followed by a sound (the 0% block) was sub-
805 tracted from the omission AEP evoked by the unexpected
806 ‘silence’. After subtraction, significant omission responses
807 were present in the 88% block, but not in the 50% (ran-
808 dom) block. In a follow-up experiment, SanMiguel et al.
809 (2013b) showed that omission responses to self-
810 generated sounds were only elicited if a button press
811 was predictive of both the identity and timing of the eli-
812 cited sound, rather than just the timing, which suggests
813 that timing alone is not enough to form an accurate pre-
814 diction of a stimulus.

815 Chennu et al. (2016) compared omission responses
816 recorded with EEG and MEG. Using a local–global para-
817 digm (Fig. 2H), the fifth tone was a global standard in 74%
818 of trials, and a global deviant or omission in 13% of trials.
819 To confirm that omission responses reflected *expectation*
820 effects and not passive carry-over effects such as oscilla-
821 tory entrainment (May and Tiitinen, 2010), unexpected
822 omissions of a fifth tone (occurring 14% of trials) were
823 compared to ‘expected omissions’ from sequences in
824 which the fifth tone was always omitted. In the EEG
825 recordings, this revealed clear omission responses that
826 were modulated by attention. Surprisingly, in the MEG
827 data the omission response was absent. This divergence
828 between MEG and EEG is interesting but difficult to inter-
829 pret, and most likely arises from the orientation of the neu-
830 ral sources or measurement noise. However, it might also
831 be consistent with a specific interpretation of omission
832 responses as reflecting prediction units only, which may
833 reside in deeper layers and should therefore be more dif-
834 ficult to detect with MEG.

835 Fujioka et al. (2009), who also used MEG, induced
836 (and violated) expectations by using a regular musical
837 beat, from which tones were occasionally omitted. Each
838 tone elicited a short gamma (> 40 Hz) burst, as is typical
839 for external stimuli. However, the authors also observed a
840 slow, oscillatory modulation of the beta band that was
841 phase locked with the occurrence of the tone. This slow
842 power-modulation steadily decreased after each beat,
843 reaching its peak just *before* the occurrence of a new
844 tone, thus indicative of a potential internal oscillatory or
845 rhythmic anticipation signal (see also Fujioka et al.,
846 2012). Intriguingly, when a tone was unexpectedly omit-
847 ted from the beat, the decrease in beta power was not
848 observed, but a (stimulus-like) sudden peak in gamma
849 was observed. This observation not only supports sen-
850 sory prediction during beat perception, but also, indirectly,
851 the notion that beta (‘prediction’) and gamma (‘error’)
852 bands signal different computational variables (see Sec-
853 tion The rhythms of prediction).

854 Finally, a number of earlier discussed studies also
855 reported omission responses of a varying extent.

856 Todorovic et al. (2011) and Todorovic and de Lange
857 (2012) reported higher field strengths after unexpected
858 than expected silences. However, their effects were
859 rather small and in Todorovic and de Lange (2012) limited
860 to late latencies (200–500 ms). More akin to ‘real’ evoked
861 responses are the omission responses in Wacongne et al.
862 (2011), who also compared expected with unexpected
863 omissions and found (contra May and Tiitinen, 2010) sig-
864 nificant responses only for unexpected omissions, in both
865 MEG and EEG.

866 Altogether, evidence from EEG (Bendixen et al., 2009;
867 SanMiguel et al., 2013a,b; Chennu et al., 2016), MEG
868 (Wacongne et al., 2011; Todorovic et al., 2011;
869 Todorovic and de Lange, 2012; Andreou et al., 2015;
870 Chennu et al., 2016) and ECoG (Hughes et al., 2001)
871 shows that omissions can evoke responses that are
872 time-locked to the omitted stimulus and appear to be gen-
873 erated in auditory cortex and superior temporal gyrus.
874 Crucially, omission responses seem to occur only after
875 unexpected omissions (Wacongne et al., 2011; Chennu
876 et al., 2016) – challenging the suggestion that they could
877 reflect passive carry-over effects – and if the omitted
878 sounds are prospectively predictable (Bendixen et al.,
879 2009) – suggesting a predictive mechanism (cf. Assump-
880 tion 1). However, the literature also shows some remark-
881 able variability. For instance, using MEG, Todorovic et al.
882 (2011) and Todorovic and de Lange (2012) find small and
883 late deflections, unlike ‘real’ auditory-evoked fields, and
884 Chennu et al. (2016) find no omission responses at all.
885 Using EEG, Chennu et al. (2016) and Bendixen et al.
886 (2009) find clearer omission responses. However, they
887 are still quite different from ‘real’ AEPs, or from the spec-
888 tacular responses in SanMiguel et al. (2013a,b) or the
889 ECoG omission responses in Hughes et al. (2001). More-
890 over, while the MEG/EEG omission responses in Rajj
891 et al. (1997) and Chennu et al. (2016) are strongly
892 affected by attention, attention had no effect on the ECoG
893 omission responses in Hughes et al. (2001).

894 Beyond the empirical variability, there is some
895 theoretical variability in how omission responses should
896 be interpreted. For some authors (e.g. SanMiguel et al.,
897 2013a,b; Schröger et al., 2015) they are simply expres-
898 sions of prediction error. This would render omission
899 responses as perhaps *the* signature finding of PC, by
900 showing that evoked responses fundamentally reflect *sur-*
901 *prise* – even in the absence of sensory input. However, as
902 Wacongne et al. (2012) point out, this interpretation criti-
903 cally depends on how prediction error is calculated. If
904 one uses subtraction, performed by a focussed inhibitory
905 pulse that ‘subtracts’ predictions from sensory input, it is
906 difficult to see how omissions could elicit prediction error
907 without allowing negative firing rates. In that case, omis-
908 sion responses are perhaps better interpreted as reflect-
909 ing purely prediction (or pre-activation) signals, which
910 speaks to its relative weakness and variability. Due to
911 these ambiguities, it is difficult to directly interpret the
912 implications of omission responses to (specific formula-
913 tions of) predictive coding. Nevertheless, collectively,
914 these studies present highly suggestive, converging evi-
915 dence of anticipatory mechanisms, operating without con-
916 scious expectation, in auditory cortex.

Predictability and precision

917 Results with the MMN paradigm demonstrate that
918 listeners are sensitive to the violation of a variety of
919 sound patterns, including very complex regularities. This
920 has been interpreted as (indirect) evidence for the
921 brain’s remarkable sensitivity to acoustic patterning.
922 However, a crucial missing link is an understanding of
923 the process by which the brain acquires an internal
924 model of regularities in the environment.

925 Recently, Barascud et al. (2016); see below for repli-
926 cation by Southwell et al. (2017) presented direct evi-
927 dence of the discovery and representation of acoustic
928 patterns, using rapid, statistically structured sequences
929 of tone-pips that transitioned from random to regular,
930 and *vice versa* (Fig. 2I). Methodologically, this paradigm
931 constitutes a departure from previous paradigms in two
932 ways: firstly, the use of very rapid sequences precludes
933 conscious discovery of regularity, instead mostly tapping
934 bottom-up-driven processes. Secondly, regularity was
935 manipulated independently from repetition, thus decou-
936 pling the effects of predictability from low-level adaptation.

937 Behaviorally, Barascud et al. (2016) first observed that
938 listeners were extremely quick at detecting the emer-
939 gence of regular patterns, performing on par with an ideal
940 observer model. Brain responses measured from naïve
941 listeners were equally rapid. Remarkably, the onset of
942 regularity manifested as a large-scale increase in sus-
943 tained amplitude (Fig. 3). Offsets of regularity (transitions
944 toward randomness), by contrast, were associated with a
945 large *drop* in sustained power. Source reconstruction
946 identified a network of sources in auditory cortex (AC),
947 inferior frontal gyrus (IFG) and the hippocampus. AC
948 and IFG sources are commonly reported in the context
949 of the MMN and interpreted as supporting the re-entrant
950 error-minimizing process underlying it (Molholm et al.,
951 2005; Opitz et al., 2002; Garrido et al., 2009b; see Sec-
952 tion Dynamic Causal Modeling of MMN).

953 The finding that emergence (and disappearance) of
954 regularity in unfolding sound sequences is associated
955 with large-scale sustained responses is interesting for
956 various reasons: firstly, it suggests the brain encodes the
957 *state* (‘regular’ vs ‘random’) rather than just the
958 transition (as in e.g. MMN). Secondly, the amplitude
959 pattern [(regular) > (random)] is not easily interpretable
960 in terms of simple physical attributes of the signal –
961 adaptation, for example, would result in the opposite
962 pattern. Finally, the neural signature of complex
963 regularity detection (i.e. enhanced responses) is
964 *opposite* to that of simpler regularity detection (i.e.
965 attenuated responses) observed in many previous
966 experiments, for example using the roving standard
967 paradigm.

968 Interestingly, the effect is also opposite to all PC
969 effects we have been considering so far, in which
970 predictability is associated with *weaker* responses.
971 Barascud et al. (2016) suggested precision-weighting
972 could underlie this inversion: if regularity is an index of
973 reliability or precision, PC predicts that regular signals
974 are up-weighted and prioritized for further processing
975 (see Section Attention as precision). As many biological
976 stimuli unfold as regularities over time, it also seems bio-
977

logically useful to prioritize such signals, for instance for subsequent auditory object formation or scene analysis.

Evidence in line with this interpretation was subsequently presented by [Sohoglu and Chait \(2016b\)](#) who used artificial ‘scenes’ consisting of concurrent tone-pip streams (modeling acoustic sources) which were temporally regular or random. Participants were quicker and more accurate to detect an object appearing in a temporally regular scene, and enjoyed an additional slight benefit if the object itself was regular. MEG responses in both passive listeners and listeners actively engaged in detecting the occasional appearance of a new source within the scene revealed increased sustained activity in scenes comprised of regular sources. Over and above this ‘scene effect’, new source appearance in regular scenes was also associated with increased responses relative to random scenes – an effect interpreted as evidence for a mechanism that infers the precision of sensory input and uses this information to up-regulate neural processing toward more reliable sensory signals.

More clues on the amplifying effect of regularity are found in [Hsu et al. \(2015\)](#). Subjects listened to sequences of tones with ascending frequencies in which the final tone varied. In 75% of trials, the tone complied with the local regularity (predicted condition). In 12.5% of trials, the last tone was unexpectedly lower than the first tone, violating the expectation induced by the ascending sequence (‘mispredicted’ condition). Finally, in 12.5% of trials, the sequence was jumbled altogether. The authors found that while predicted tones elicited a weaker N1 deflection than mispredicted ones (a well-documented expectation effect), wholly *unpredicted* tones elicited an even weaker N1 still. According to [Hsu et al. \(2015\)](#), this is because predicted and mispredicted responses express both a prediction and a (small or large) prediction error, but unpredicted responses reflect *only prediction error* and are therefore weakest. However, as remarked by [Ross and Hansen \(2016\)](#), it seems at odds with the probabilistic nature of PC to assume predictions are *absent* in the unpredicted condition: rather, what distinguishes the unpredicted condition is the low predictability of the signal. The attenuated N1 to wholly unpredictable stimuli might be understood as *inversion* of the enhanced response to predictable stimuli in [Barascud et al. \(2016\)](#) and [Sohoglu and Chait \(2016b\)](#): the brain might consider the jumbled tone ladder as noisy and uninformative, hence down-weighting the response.

In sum, accumulating evidence suggests that, at least under certain conditions, predictability may enhance, rather than suppress, neural responses. This result fits into the PC framework if one considers effects of precision: sequences of random stimuli may be ‘flagged’ as uninteresting noise (low precision) and hence down-weighted, while streams containing a regularity are considered informative and are hence up-weighted. Since precision can explain effects that are opposite to ‘traditional’ PC effects, invoking it begs the question when, exactly, predictability is supposed to suppress neural responses and when it should enhance them. As we will see, this need for a ‘principled account’ will be a

recurring theme in studies that examine the main manifestation of precision-weighting — i.e. attention.

Attention as precision

Because the world is variable and the brain noisy, a degree of prediction error is inevitable. Distinguishing such ‘residue error’ (related to noise) from relevant error (related to incorrect beliefs or changes in the world) requires that not all prediction error is treated equally. A Bayes-optimal approach, successfully applied in engineering ([Kalman, 1960](#)) as well as neuroscience ([Yu, 2014](#)) is to weight errors by their *reliability*, typically quantified as the uncertainty of predictions relative to the that of observations, a coefficient known as *Kalman gain* ([Kalman, 1960](#); [Anderson and Moore, 1979](#)). When the gain (precision) is high, inputs are up-weighted and will dominate inference; when it is low, inputs are down-weighted and predictions dominate inference. Several authors in the predictive coding field ([Rao, 2005](#); [Spratling, 2008a,b, 2010](#); [Feldman and Friston, 2010](#); see also [Dayan and Zemel, 1999](#); [Yu and Dayan, 2003](#); [Yu and Dayan, 2005a,b](#)) have used such optimal handling of uncertainty as a framework for attention, since it offers normative principles that can explain selective processing by motivating *why* some signals are computationally more relevant than others.

Uncertainty-weighting affects inference and learning differently; here, we will focus on perceptual *inference* (but see [Yu, 2014](#), for a treatment of Bayesian approaches to attention which also covers learning). During inference, reliable inputs are weighted more strongly, and PC proposes that attending to a feature amounts to *expecting* that signals with this feature will be *reliable* or *informative*, and should thus be prioritized for processing ([Feldman and Friston, 2010](#)). Response strength should therefore always be a function of both the size of the error and its precision. In other words, every brain response should be sensitive to attentional modulation. This proposal implies a departure from accounts of MMN that describe MMN as pre-attentive ([Garrido et al., 2009b](#); [Jääskeläinen et al., 2004](#); [Winkler and Czigler, 1998](#)).

Preliminary support is found in [Chennu et al. \(2013\)](#). The authors recorded EEG while presenting blocks of tone sequences to one ear, occasionally replacing the fifth tone by either a different tone in the same ear (monaural deviant) or by the same tone in the opposite ear (interaural deviant). Additionally, participants counted deviant tones (attend tones) or deviant sequences (attend sequences) or performed a visual task (distraction). Focussing here on local deviants at MMN latencies, both monaural and interaural MMN were reduced during distraction compared to attending sequences. Attending tones, however, instead of amplifying the MMN (by increasing gain of error-neurons), attenuated it. The authors suggest their manipulation may have been confounded: counting deviant tones did not only focus attention on tones (just as counting sequences did) but might

1095 have also increased conscious expectation of unexpected
1096 tones, thus decreasing overall surprise.

1097 To circumvent this confound, other studies
1098 manipulated attention and prediction orthogonally.
1099 [Aukstulewicz and Friston \(2015a\)](#) used a roving stan-
1100 dard paradigm in which participants attended to one of
1101 two time windows (early or late), after which the roving
1102 standard was presented at each window with an inde-
1103 pendent probability of 50%. Participants reported if there
1104 was *no* stimulus at the attended latency. Only trials where
1105 the tone was presented at *both* latencies were included,
1106 thus rejecting all possible motor artifacts. A significant
1107 interaction effect was found; specifically, MMN was
1108 observed in attended, but not in unattended time-
1109 windows. This attentional enhancement of MMN is com-
1110 patible with precision weighting. Note, however, that the
1111 non-significance of MMN outside the scope of attention
1112 seems to contradict earlier findings that MMN is not
1113 dependent on attention; note, too, that the effects
1114 reported by [Aukstulewicz and Friston \(2015a\)](#) are rela-
1115 tively late and relatively short – for instance, the MMN
1116 only reached significance between, 190 and 210 ms and
1117 the deviance-attention interaction only between 193 and
1118 197 ms. Both may have been related to a lack of power
1119 after rejecting so many trials.

1120 Another independent manipulation was reported by
1121 [Hsu et al. \(2014\)](#). The authors presented two streams of
1122 tone pairs: in one stream, the frequency of the second
1123 tone in a pair was always two natural keys higher than
1124 the first; in the other stream, the relationship between
1125 the first and second tone was random. Attention was
1126 manipulated by asking participants to report occasional
1127 tones with decreased loudness in one of the streams.
1128 The authors found an interaction of attention and predic-
1129 tion on N1 amplitudes. Specifically, attended/predictable
1130 tones elicited a stronger response than all other tones,
1131 between which differences were non-significant. This
1132 includes attended *versus* unattended unpredictable
1133 tones, hence the authors concluded that attentional
1134 enhancement of N1 depends on prediction. Note, how-
1135 ever, that this interacting effect between attention and
1136 prediction (attention reversing the effect of prediction) is
1137 at odds with [Aukstulewicz and Friston \(2015a\)](#), who
1138 found the opposite (attention enhancing prediction
1139 effects) 100 ms later.

1140 A recent EEG study by [Garrido et al. \(2017\)](#) compared
1141 the two accounts explicitly. Participants were presented
1142 Gaussian white noise to both ears and instructed to detect
1143 silent gaps in one or both ears. Embedded in the noise,
1144 task-irrelevant oddball sequences were presented. The
1145 authors formulated two models of the interplay between
1146 attention and prediction: in the first, attention could
1147 reverse the effect of prediction ([Hsu et al., 2014](#); see also
1148 [Kok et al., 2012](#) in vision). In the second model, attention
1149 enhanced responses, predicted and unpredicted alike.
1150 The authors observed an MMN, and found that attention
1151 enhanced it, but contrary to [Hsu et al. \(2014\)](#) they found
1152 no interaction. In line with this observation, Bayesian
1153 model comparison favored the opposition model. Con-
1154 trary to [Aukstulewicz and Friston \(2015a\)](#), but in line with

the MMN literature, MMN was also found in the absence
of attention.

1155 Rather than deliberately directed, attention is
1156 sometimes automatically attracted to a stimulus. Stimuli
1157 with this capacity are called *salient* ([Itti et al., 1998](#)). Pre-
1158 dictive coding accounts for salience by appealing to the
1159 *intrinsic precision* of stimuli. Intense stimuli, for instance,
1160 can be seen as having a high signal-to-noise ratio due
1161 to sheer signal strength; inversely, regular stimuli would
1162 enjoy high precision by virtue of *low variance*. Indeed, this
1163 latter effect was proposed by [Barascud et al. \(2016\)](#) to
1164 explain large increases in MEG signals induced by audi-
1165 tory regularities ([Fig. 3](#)). Such up-weighting of regular
1166 sounds seems ethologically sensible, as regular patterns
1167 often carry stable, behaviorally relevant information about
1168 the world. The account also has a straightforward empiri-
1169 cal consequence – regular stimuli should attract atten-
1170 tion. In vision, a recent study indeed demonstrates this
1171 effect ([Zhao et al., 2013](#)).

1172 [Southwell et al. \(2017\)](#) tested this idea in the auditory
1173 domain. Using EEG, the authors first replicated the MEG-
1174 effects observed by [Barascud et al. \(2016\)](#): task-irrelevant
1175 regular sequences (as used by Barascud and colleagues)
1176 induced large increases in sustained EEG amplitude.
1177 Next, the authors tested behaviorally whether the same
1178 regular patterns would capture attention more strongly,
1179 measured as the interference with concurrent tasks.
1180 Remarkably, their results suggested that regularity was
1181 not more distracting (if task-irrelevant) or more salient (if
1182 task-relevant) than random patterns. The fact that *neu-
1183 rally*, regularity induces marked sustained amplitude
1184 increases, but *behaviorally* the same patterns are not
1185 more salient, contradicts the attentional gain explanation
1186 proposed by [Barascud et al. \(2016\)](#). [Southwell et al.
1187 \(2017\)](#) suggest that this leaves us with three alternative
1188 hypotheses: Either the gain in amplitude reflects an
1189 upsurge of (poly-synaptic) *inhibition* or explaining away
1190 by higher regions, which is not dissociable from excitation
1191 using M/EEG. Alternatively, it may reflect a number of
1192 quite distinct processes. Or finally, it could reflect some
1193 form of precision-weighting which does not manifest as
1194 high-level attentional capture. This last possibility, how-
1195 ever, would imply that under PC attention is (by definition)
1196 adaptive precision-weighting, but adaptive precision-
1197 weighting is not (always) attention. While logically possi-
1198 ble, this creates an awkward disconnect between neural
1199 responses and cognition, and calls for a more principled
1200 approach to decide when precision weighting is ‘high-
1201 level attentional’ or not.

1202 Altogether, the depiction of attention as the weighting
1203 of sensory signals by their (expected) *precision* ([Feldman
1204 and Friston, 2010](#); see also [Rao, 2005](#); [Spratling, 2008a,
1205 b](#)) elegantly integrates many known attentional effects
1206 into the realm of prediction. However, the increased
1207 opportunities this creates for post-hoc explanations are
1208 – at least in the auditory domain – not yet met by a propor-
1209 tional increase in rigorous confirmatory results. For
1210 endogenous attention, studies explicitly testing the
1211 account report small and sometimes conflicting effects
1212 ([Chennu et al., 2013](#); [Hsu et al., 2014](#); [Aukstulewicz
1213 1214](#)

and Friston, 2015a; Garrido et al., 2017). For exogenous attention, precision-weighting offers a compelling explanation for the enhancing effect of regularity (Barascud et al., 2016; Sohoglu and Chait, 2016b; Hsu et al., 2015; Southwell et al., 2017); however, the direct consequence of this claim (that regularity should be salient) was consistently not found (Southwell et al., 2017). More research is needed to test and potentially revise the notion of auditory precision-weighting, and to explore differences with vision where it may apply more readily (e.g. Kok et al., 2012; Zhao et al., 2013).

1226 The rhythms of prediction

1227 In systems neuroscience, distinct oscillatory signatures
1228 for feedforward processing (operating mainly via the
1229 gamma band) and feedback processing (using alpha
1230 and mainly beta bands) have been demonstrated in
1231 considerable detail (van Kerkoerle et al., 2014;
1232 Buschman and Miller, 2007). In standard PC, this oscillatory
1233 asymmetry is hypothesized to be linked to the functional
1234 asymmetry between (upward) errors and
1235 (backward) predictions. In other words, predictions and
1236 errors should have distinct oscillatory signatures (Arnal
1237 and Giraud, 2012; Bastos et al., 2012). However, evidence
1238 for this claim has remained indirect (see Arnal
1239 et al., 2011 for a demonstration in speech perception;
1240 van Pelt et al., 2016 in causal cognition).

1241 Recently, Sedley et al. (2016) provided more direct
1242 evidence, using a simple parametric task to generate
1243 auditory stimuli while recording local field potentials using
1244 ECoG. Three human subjects listened to short (300-ms)
1245 sequences of harmonic complexes of which only the fundamental
1246 frequency varied. In any given trial there was a
1247 7/8 chance that f_0 would be sampled from the same
1248 Gaussian population, and a 1/8 chance that it would be
1249 sampled from a new one. Assuming that subjects unconsciously
1250 tracked the statistics, the authors used a Bayes-optimal
1251 inversion of their generative algorithm to calculate
1252 trial-by-trial estimates of four key inferential variables:
1253 *prediction error*, *surprise*, *prediction change* and *prediction
1254 precision* (where surprise is the precision-weighted
1255 variant of prediction error). The authors correlated these
1256 estimates with a time–frequency analysis of the LFP
1257 trace. As expected, the authors found that gamma was
1258 correlated with surprise (more than prediction error).
1259 Moreover, *prediction change* correlated with an increase
1260 in beta-band around 400 ms. Finally, and not explicitly
1261 predicted by PC, the authors found significant correlations
1262 between the alpha band and *precision of predictions*,
1263 although this effect was less pronounced than that in
1264 the beta and gamma band.

1265 Among the earlier discussed studies, only Fujioka
1266 et al. (2009) reported effects similarly compatible with
1267 PC. There, an oscillatory stimulus (a beat) induced an
1268 oscillatory modulation of the beta band that was time-
1269 locked to the beat. When a tone was omitted, the immediate
1270 decrease in beta-power was not observed, suggesting
1271 that the beta-power may have been an oscillatory expectation.
1272 Moreover, omissions did induce short gamma
1273 bursts, characteristic of stimuli (or surprise). Other studies,
1274 however, did not report clear oscillatory dissociations.

1275 Signatures of prediction in the beta-band, for instance,
1276 were absent in Dürschmid et al. (2016) who reported
1277 ECoG recordings to predictable and unpredictable deviants.
1278 The authors made sure they compared electrodes
1279 with similar sensitivity for different frequency bands, and
1280 nevertheless only found effects seemed in the high-
1281 gamma band (> 60 Hz) and at low frequencies related
1282 to evoked potentials, but hardly in between.

1283 El Karoui et al. (2015) presented ECoG recordings of
1284 patients performing a local–global paradigm and found a
1285 decrease in sustained beta power after global mismatches
1286 (which would arguably involve more prediction-change).
1287 However, the global deviants were also the behavioral
1288 target, confounding attention and prediction, and making
1289 interpretation difficult. Finally, Todorovic et al. (2015)
1290 found effects of attention and expectation only in the
1291 beta-band, which decreased in power after unexpected
1292 tones, but only if attention was directed to another,
1293 earlier time window.

1294 To summarize, evidence for distinct oscillatory
1295 signatures of prediction and error processing is limited,
1296 indirect and mixed: only two of six studies revealed
1297 spectral patterns compatible with the predictions of PC.
1298 Methodological differences make it difficult to draw an
1299 unequivocal conclusion on the existence of oscillatory
1300 differences between prediction and error processing.
1301 Given the increasing evidence for laminar differences
1302 between alpha/beta and gamma band dominance (e.g.
1303 Scheeringa et al., 2016), oscillatory differences are a
1304 potential tool to test the standard implementation of PC,
1305 and future studies using parametric methods like Sedley
1306 et al. (2016) may offer much needed confirmatory
1307 evidence. However, simply interpreting different bands as
1308 reflecting different variables without employing a parametric
1309 approach to calculate the relevant variables on a trial-
1310 by-trial basis seems empirically unwarranted given the
1311 highly variable results of studies without such a model-
1312 based approach.

1313 EFFECTIVE CONNECTIVITY – CLUES FROM 1314 DCM

1315 Measurement in neuroscience typically allows for high
1316 spatial or high temporal resolution. Accordingly, many
1317 studies probe the ‘when’ or ‘where’ of neural responses.
1318 However, this provides little insight in how responses
1319 emerge. Causal modeling techniques attempt to
1320 overcome this by estimating changes in causal
1321 influences between sources underlying effects of
1322 interest. One of these techniques – Dynamic Causal
1323 Modelling (Friston et al., 2003) – has been extensively
1324 used to test predictive coding, especially in relation to
1325 the MMN. Since DCM is a theory-driven method which
1326 makes several enabling assumptions, we will briefly
1327 recapitulate the ideas behind DCM before discussing the
1328 studies that used it.

1329 DCM for MEG and EEG

1330 DCM is a hypothesis testing framework, which works by
1331 predicting neural responses based on several
1332 hypotheses, and then comparing these predictions to
1333

1333 the data. Predictions are generated by combining a
 1334 neuronal and an observational model. In DCM for M/
 1335 EEG (Kiebel et al., 2006, 2008), the observational model
 1336 is a *lead field* as used in source reconstruction, which
 1337 maps hidden dipoles in the skull to observable deflections
 1338 at the scalp. DCM goes beyond this ‘common’ reconstruction
 1339 method by using a *neuronal model* to explicitly model
 1340 intracranial current flow. Neuronal models in DCM for M/
 1341 EEG (see Moran et al., 2013 for review) are mostly *mass*
 1342 *models*, which do not capture the complex dynamics
 1343 between large numbers of individual neurons (as found
 1344 in the skull) but rather the simpler dynamics between
 1345 massively synchronized *populations* of neurons (as mea-
 1346 surable at the scalp). Typically, a region is described with
 1347 three or four sub-populations of inhibitory and excitatory
 1348 neurons (each modeled using an ordinary second-order
 1349 differential equation) that operate as a dampened linear
 1350 oscillator (David and Friston, 2003; David et al., 2006).

1351 In DCM, hypotheses are embodied as *architectures*:
 1352 cortical sources connected in a specific, directional way.
 1353 Responses can be generated by injecting a Gaussian
 1354 impulse into one source (e.g. A1), after which the
 1355 current flow ensuing from the network is passed through
 1356 the lead field to generate observational patterns for the
 1357 modality in use (EEG or MEG). Between-trial effects are
 1358 modeled as changes in *extrinsic* or *intrinsic* connectivity.
 1359 *Extrinsic connectivity* refers to coupling strength
 1360 between regions, is modeled by directional coupling
 1361 parameters, and can be thought of as inter-regional
 1362 synaptic modulation (c.f. learning). *Intrinsic connectivity*
 1363 refers to the strength with which a signal is propagated

1364 *within* a region. It is adjusted by changing the maximum
 1365 firing rate of excitatory populations, and can be thought
 1366 of as changing the *excitability* of a region (c.f.
 1367 adaptation). Ultimately, the architecture that can most
 1368 readily explain the effect – yielding the best fit with the
 1369 least complexity – is deemed *most likely*.

1370 By virtue of these assumptions, DCM aims to provide
 1371 an *in silico* environment for testing hypotheses about both
 1372 the neural architecture underlying experimental data, and
 1373 the changes *within* this architecture that best explain
 1374 between-trial effects of interest.

Dynamic Causal modeling of MMN

1375 The first application of DCM to MEG and EEG is
 1376 described in Garrido et al. (2007a,b) who modeled the dif-
 1377 ference between standard and deviant ERPs from an
 1378 oddball paradigm. Garrido et al. (2007a,b) found that the
 1379 difference between standard and deviant responses was
 1380 best explained by bidirectional connectivity changes
 1381 between Heschl’s Gyrus (A1), superior temporal gyrus
 1382 (STG) and right inferior frontal gyrus (rIFG). Garrido
 1383 et al. (2007b) replicated this basic result at the group-
 1384 level and verified that backward modulations were espe-
 1385 cially important for explaining ERP differences at later
 1386 latencies (200–400 ms).

1387 Having established these foundational results,
 1388 Garrido et al. (2008) used DCM to compare theoretical
 1389 accounts of MMN. In the study, the authors modeled a
 1390 series of responses from the roving standard paradigm,
 1391 from deviant (first tone) to standard (last tone). They then
 1392

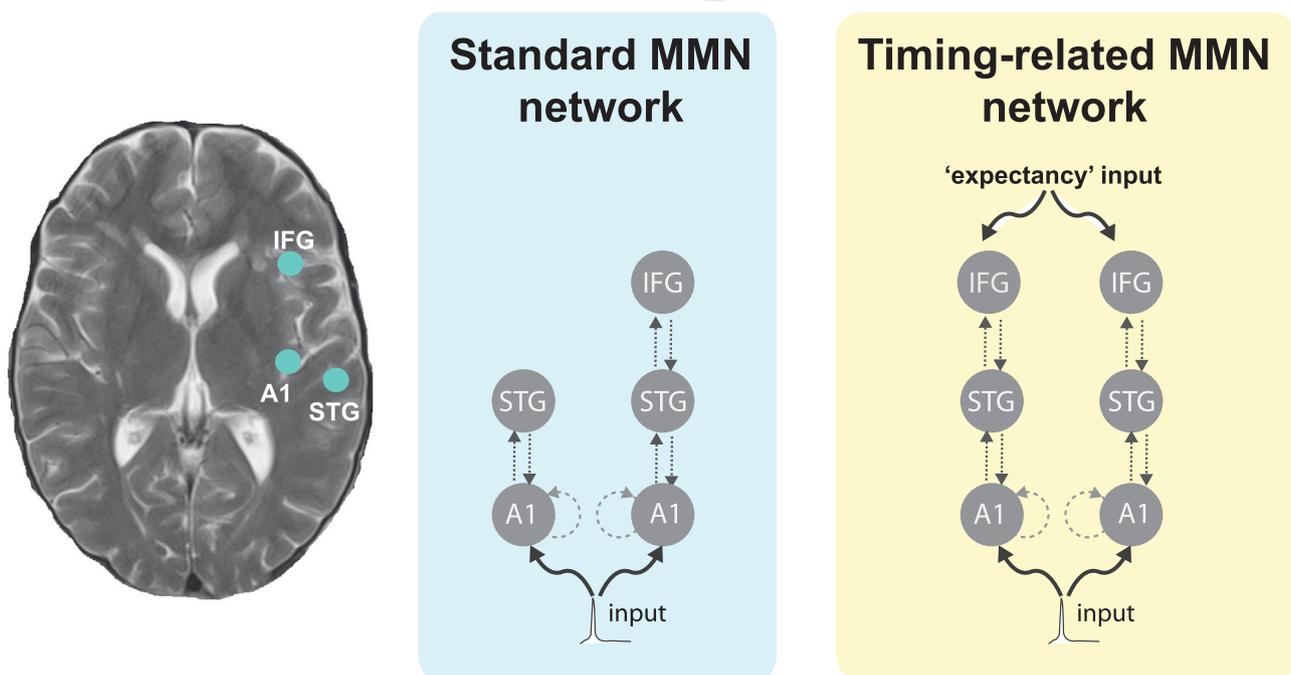


Fig. 4. Graphical specification of connectivity models underlying the MMN as suggested by DCM. Left: connectivity modulations in an asymmetric frontotemporal network, combined with neuronal excitability modulations in A1, was shown to best explain the MMN across a variety of paradigms and modalities (Garrido et al., 2008, 2009a; Phillips et al., 2015, 2016; Chennu et al., 2016; Barascud et al., 2016). Right: connectivity model including left IFG and frontal ‘expectancy inputs’ which was found to best explain MMN responses to temporal irregularities (duration and silent gap) or omissions (Phillips et al., 2015, 2016; Chennu et al., 2016).

1393 compared which MMN–hypothesis could explain the
1394 associated ERP differences – and thus the differential
1395 MMN – best. Each MMN hypothesis was embodied as a
1396 different variation of the frontotemporal architecture out-
1397 lined above (see Fig. 4). The adaptation hypothesis was
1398 modeled as a network in which only the excitability of
1399 A1 varied over trials. The model-adjustment hypothesis
1400 (which explains the MMN as a fronto-temporal memory-
1401 adjustment; cf. Näätänen et al., 1978, 2007) was modeled
1402 as a network in which only the between-region connectiv-
1403 ity varied between trials. Finally, predictive coding was
1404 embodied in a model in which both the excitability of A1
1405 and inter-regional connectivity varied. The idea was that
1406 PC incorporates both adaptation and model adjustment
1407 (see also Section Predictive coding and MMN) – in this
1408 view, changes in excitability of A1 and fronto-temporal
1409 coupling are expressions of belief-updating at different
1410 hierarchical levels (intra-regional microcircuitry versus
1411 inter-regional network connectivity). Model comparison
1412 showed that the hybrid PC model explained the ERP dif-
1413 ferences best. The superiority of hybrid model was later
1414 replicated in a study using the ‘classic’ frequency oddball
1415 (Garrido et al., 2009a).

1416 Temporal deviants and top-down predictions

1417 Within the same model space, Phillips et al. (2015) repli-
1418 cated this result using MEG and stimuli that deviated
1419 across various dimensions, such as frequency, intensity,
1420 or duration. To study all these deviant dimensions, the
1421 authors used an optimized oddball paradigm (Näätänen
1422 et al., 2004), in which each block starts with several stan-
1423 dard tones, after which standards start alternating with dif-
1424 ferent deviants – e.g. standard, frequency-deviant,
1425 standard, duration-deviant, standard, etc. First, within
1426 the model space of Garrido et al. (2008, 2009a), the
1427 model with forward, backward and intrinsic modulations
1428 in A1 was confirmed to ‘win’ for all dimensions. Subse-
1429 quently, the model space was extended to include archi-
1430 tectures with left IFG and models with an additional,
1431 frontal input. Usually, Gaussian impulse functions are
1432 models of sensory inputs, and are only ‘injected’ at thala-
1433 morecipient regions. By contrast, Phillips et al. (2015)
1434 located a second input at IFG (‘expectancy inputs’; see
1435 Fig. 4). Interestingly, models that included a prefrontal
1436 ‘expectation’ input only provided a better fit for temporal
1437 deviants – that is, either tones containing a short silent
1438 gap in the center, or tones that deviated in duration. Mod-
1439 els with an additional IFG were more likely across all stim-
1440 ulus dimensions.

1441 In a follow-up study, Phillips et al. (2016) first repli-
1442 cated these findings by performing the same analysis on
1443 a new MEG recording of 50 subjects. They then extended
1444 the analyses to ECoG data. As explained above, in DCM
1445 for MEG and EEG current flow ensuing from the network
1446 is passed through a *lead field* to generate observational
1447 patterns specific to M/EEG. As this additional model
1448 may introduce uncertainty, it is important to verify whether
1449 inverting a DCM *without* observation model (i.e. using sig-
1450 nals directly from cortex) yields similar results. The
1451 authors recruited two patients: one with electrodes over
1452 *right* IFG and STG, and one with electrodes over left

1453 IFG and STG. The ECoG DCM results matched earlier
1454 DCM results with respect to the relative importance of for-
1455 ward/backward interactions. However, the frontal expect-
1456 tancy input ‘won’ only in the patient with left-lateralized
1457 electrodes. Strikingly, this asymmetry was also found in
1458 the MEG results: temporal deviants were best explained
1459 by models with a *left*, rather than bilateral, IFG input. This
1460 apparent lateralization is remarkable and calls for a repli-
1461 cation, since earlier studies did not consider left IFG a
1462 ‘main MMN generator’ (Opitz et al., 2002; Garrido et al.,
1463 2008, 2009a,b; Chennu et al., 2016). Alternatively, the
1464 effect may be related to differences in electrode locations
1465 of left versus right IFG. This artifact would be propagated
1466 to the MEG results because the coordinates from the
1467 ECoG electrodes were used as source coordinates in the
1468 observation model.

1469 Finally, Chennu et al. (2016) performed a DCM analy-
1470 sis on MEG and EEG data from a local–global paradigm
1471 that included omissions. In two conditions, participants
1472 either counted uncommon sequences (attend-auditory)
1473 or performed an unrelated visual task (attend-visual).
1474 For deviant tones, the ‘classic’ architecture used by
1475 (Garrido et al., 2007a,b, 2008, 2009a) best explained
1476 the data both in the attended and unattended condition.
1477 For the omission responses, by contrast, an architecture
1478 that included bilateral IFG and a frontal expectancy input
1479 (which replaced the thalamic sensory input) best
1480 explained the data, which is compatible with the idea that
1481 omission responses reflect top-down prediction (rather
1482 than prediction error).

1483 Discussion

1484 To summarize, DCM studies show that models which
1485 modulate both A1 excitability and fronto-temporal
1486 connectivity explain deviant responses in oddball
1487 paradigms (Garrido et al., 2007a,b, 2009a) and variations
1488 thereof (Garrido et al., 2008; Phillips et al., 2015, 2016;
1489 Chennu et al., 2016) better than models that modulate
1490 only A1 excitability or fronto-temporal connectivity. More-
1491 over, responses to tones that deviate temporally, or are
1492 omitted altogether, are best explained by models with
1493 frontal ‘expectation inputs’ which replace (Chennu et al.,
1494 2016) or augment (Phillips et al., 2015, 2016) the thalamic
1495 sensory input.

1496 These patterns of effects are in line with PC by
1497 describing MMN not only via A1 adaptation or long-
1498 range connectivity, but via a mechanism that combines
1499 both. Moreover, an interesting analogy might be drawn
1500 between the need for frontal inputs to explain temporal
1501 deviants in DCM (Phillips et al., 2015, 2016) and the fact
1502 that temporal deviants constitute a key difference
1503 between network-level MMN, which is sensitive to tempo-
1504 ral deviants, and neuron-level SSA, which is not (Khoury
1505 and Nelken, 2015). Although this post-hoc analogy would
1506 require further investigation, the fact that only intracellular
1507 recordings and DCM appear to consistently distinguish
1508 temporal deviants from other deviants illustrates the
1509 potential of the technique to extend beyond traditional
1510 analysis of non-invasive data.

1511 However, the DCM studies have several limitations.
1512 The first issue is that DCM relies on assumptions and

simplifications which are not fully validated. The neural mass models used in most DCM for M/EEG studies are even abstracted to such degree that some parameters don't have obvious physiological substrates. One response to this is to develop more complex models with more biologically meaningful parameters (Moran et al., 2013); an approach that is showing promising results (Gilbert et al., 2016). However, this does not yet address the issue of validation. Although initial studies have established the *face validity* of DCM for M/EEG (Garrido et al., 2007a,b, 2009a,b) and the extensively replicated MMN results demonstrate *predictive validity* (Phillips et al., 2015, 2016; Chennu et al., 2016) much needs to be done before DCM can be said to have *construct validity*. Combining different techniques, such as in Phillips et al. (2016), will be critical in this process. Note however that Phillips et al. (2016) only partially validated the *observation model*, which was arguably the least controversial.

A second issue is to what extent these results support predictive coding. Even if we fully accept the network modulations suggested by DCM, this doesn't mean that these changes necessarily reflect predictive coding, or even a single underlying mechanism. Indeed, it is difficult to see why changes in A1 excitability and STG-IFG connectivity should be uniquely characteristic of predictive coding. This problem is reinforced by the fact that the discussed studies have mostly used designs in which expectation and adaptation are confounded, which makes arbitrating between predictive and non-predictive interpretations even more difficult. As such, while the discussed studies constitute exciting methodological developments in the analysis of non-invasive electrophysiological data, their strength as empirical support for predictive coding theory seems rather limited.

CONCLUSION

In this review we aimed to provide a comprehensive empirical evaluation of five key assumptions of predictive coding theory in the context of auditory pattern processing. Findings from animal, human and computational neuroscience provide converging evidence for the fundamental influence of expectations on neural responses and specifically the notion of *prediction error* as a model of sensory responsiveness (Assumption 2). Studies on unexpectedly omitted stimuli provide support for the anticipatory, *predictive* nature of these expectancy effects (Assumption 1). Moreover, the dissociation of expectancy effects at different hierarchical levels in both animal and human literature seems suggestive of the hierarchical nesting of predictions, as postulated by predictive coding theory and implied by Dynamic Causal Modeling results (Assumption 1), although more experiments are needed that explicitly manipulate multiple, nested regularities. As to the remaining three assumptions, the picture is less clear. Critically, for the existence of separate prediction and error neurons residing in distinct cortical layers (Assumption 3), there is currently no evidence in

the auditory domain in line with this idea (but see Bell et al., 2016; Kok et al., 2016, for recent studies in vision). The recent development to conceptualize attention as the weighting of sensory input by sensory precision (Assumption 4), has provided elegant post-hoc explanations for a broad range of phenomena, but has yet to provide rigorous a priori confirmatory results. Finally, the dissociation between different frequency bands and computational variables in PC (Assumption 5) has been demonstrated by one study which explicitly estimated the variables on a trial-by-trial basis; studies that did not use such a model-based approach however mostly failed to find similar associations. Looking to the future, progress in the field will critically depend on investigating these assumptions in order to test and revise or falsify specific implementations of PC. Doing so will require closer collaboration between sub-disciplines, in particular between animal and human research, where methodological and conceptual differences currently create interpretational difficulties. Finally, to test crucial theoretical distinctions (e.g. prediction error versus precision-weighted prediction error) there is an ongoing need for computationally explicit analyses in both human and animal neuroscience.

In short, over the past decade a broad range of findings in auditory neuroscience have pointed to a fundamental role of expectations and prediction errors in sensory processing. Going from these findings to the alternative, overarching framework envisioned by PC, however, requires a number of theoretical steps between which the empirical links are currently missing. Uncovering, revising or potentially refuting these 'missing links' is difficult but feasible, and provides an exciting neuroscientific challenge for the years to come.

UNCITED REFERENCES

Barascud et al. (2014), Chen et al. (2015), Denham and Winkler (2006), Körding and Wolpert (2004), Malmierca et al. (2009), Nelken et al. (2003, 2013), Paavilainen (2013), Sohoglu and Chait (2016a), Sussman et al. (1998), Ulanovsky et al. (2003), Wallace et al. (2008) and Angela (2014).

Acknowledgments—We are grateful to Jesse Geerts, Florent Meyniel, Maxime Maheu, James Kilner and Karl Friston for insightful discussions and helpful comments, and to Roman Strijbos for the *Illustrator lessons*. This work was funded by a BBSRC project grant to MC, and Prins Bernhard Cultuurfonds and Institut français des Pays-Bas fellowships to MH.

REFERENCES

Anderson BDO, Moore JB (1979) Optimal filtering. Eaglewood Cliffs, NJ: Prentice-Hall.
Arnal LH, Wyart V, Giraud A-L (2011) Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nat Neurosci* 14:797–801. <http://dx.doi.org/10.1038/nn.2810>.
Arnal LH, Giraud AL (2012) Cortical oscillations and sensory predictions. *Trends Cogn Sci*. <http://dx.doi.org/10.1016/j.tics.2012.05.003>.

- Auksztulewicz R, Friston K (2015a) Attentional enhancement of auditory mismatch responses: a DCM/MEG study. *Cereb Cortex* 1–11. <http://dx.doi.org/10.1093/cercor/bhu323>.
- Auksztulewicz R, Friston K (2015b) Repetition suppression and its contextual determinants in predictive coding. *Cortex*. <http://dx.doi.org/10.1016/j.cortex.2015.11.024>.
- Baess P, Widmann A, Roye A, Schröger E, Jacobsen T (2009) Attenuated human auditory middle latency response and evoked 40-Hz response to self-initiated sounds. *Eur J Neurosci* 29:1514–1521. <http://dx.doi.org/10.1111/j.1460-9568.2009.06683>.
- Barascud N, Griffiths TD, McAlpine D, Chait M (2014) “Change deafness” arising from inter-feature masking within a single auditory object. *J Cogn Neurosci* 26:514–528. <http://dx.doi.org/10.1162/jocn.a.00481>.
- Barascud N, Pearce MT, Griffiths TD, Friston KJ, Chait M (2016) Brain responses in humans reveal ideal observer-like sensitivity to complex acoustic patterns. *Proc Natl Acad Sci USA* 113: E616–E625. <http://dx.doi.org/10.1073/pnas.1508523113>.
- Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ (2012) Canonical microcircuits for predictive coding. *Neuron* 76:695–711. <http://dx.doi.org/10.1016/j.neuron.2012.10.038>.
- Bell AH, Summerfield C, Morin EL, Malecek NJ, Ungerleider LG (2016) Encoding of stimulus probability in macaque inferior temporal cortex. *Curr Biol*. <http://dx.doi.org/10.1016/j.cub.2016.07.007>.
- Bendixen A, Schröger E, Winkler I (2009) I heard that coming: event-related potential evidence for stimulus-driven prediction in the auditory system. *J Neurosci* 29:8447–8451. <http://dx.doi.org/10.1523/JNEUROSCI.1493-09.2009>.
- Buschman TJ, Miller EK (2007) Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315:1860–1862. <http://dx.doi.org/10.1126/science.1138071>.
- Chen I-W, Helmchen F, Lütcke H (2015) Specific early and late oddball-evoked responses in excitatory and inhibitory neurons of mouse auditory cortex. *J Neurosci* 35:12560–12573. <http://dx.doi.org/10.1523/JNEUROSCI.2240-15.2015>.
- 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. <http://dx.doi.org/10.1523/JNEUROSCI.0114-13.2013>.
- 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* 36:8305–8316. <http://dx.doi.org/10.1523/JNEUROSCI.1125-16.2016>.
- Clark A (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav Brain Sci* 36:181–204. <http://dx.doi.org/10.1017/S0140525X12000477>.
- Clark A (2016) *Surfing uncertainty: Prediction, action, and the embodied mind*. Oxford: Oxford University Press.
- Collingridge GL, Bliss TYP (1987) NMDA receptors – Their role in long-term potentiation. *Trends Neurosci* 10:288–293. [http://dx.doi.org/10.1016/0166-2236\(87\)90175-5](http://dx.doi.org/10.1016/0166-2236(87)90175-5).
- Cornella M, Leung S, Grimm S, Escera C (2012) Detection of simple and pattern regularity violations occurs at different levels of the auditory hierarchy. *PLoS One* 7:e43604. <http://dx.doi.org/10.1371/journal.pone.0043604>.
- Costa-Faidella J, Baldeweg T, Grimm S, Escera C (2011) Interactions between “what” and “when” in the auditory system: temporal predictability enhances repetition suppression. *J Neurosci* 31:18590–18597. <http://dx.doi.org/10.1523/JNEUROSCI.2599-11.2011>.
- David O, Friston KJ (2003) A neural mass model for MEG/EEG: Coupling and neuronal dynamics. *Neuroimage* 20:1743–1755. <http://dx.doi.org/10.1016/j.neuroimage.2003.07.015>.
- David O, Kilner JM, Friston KJ (2006) Mechanisms of evoked and induced responses in MEG/EEG. *Neuroimage* 31:1580–1591. <http://dx.doi.org/10.1016/j.neuroimage.2006.02.034>.
- Dayan P., Zemel, R.S.R., 1999. Statistical models and sensory attention. 9th Int. Conf. Artif. Neural Networks ICANN '99 1999, 1017–1022, doi: 10.1049/cp:19991246.
- Dehaene S, Kerszberg M, Changeux JP (1998) A neuronal model of a global workspace in effortful cognitive tasks. *Proc Natl Acad Sci USA* 95:14529–14534. <http://dx.doi.org/10.1073/pnas.95.24.14529>.
- Denham SL, Winkler I (2006) The role of predictive models in the formation of auditory streams. *J Physiol Paris* 100:154–170. <http://dx.doi.org/10.1016/j.jphysparis.2006.09.012>.
- Dürschmid S, Edwards E, Reichert C, Dewar C, Hinrichs H, Heinze H-J, 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 USA* 113:6755–6760. <http://dx.doi.org/10.1073/pnas.1525030113>.
- Dykstra AR, Gutschalk A (2015) Does the mismatch negativity operate on a consciously accessible memory trace? *Sci Adv* 1: e1500677. <http://dx.doi.org/10.1126/sciadv.1500677>.
- Egner T, Summerfield C (2013) Grounding predictive coding models in empirical neuroscience research. *Behav Brain Sci* 36:210–211. <http://dx.doi.org/10.1017/S0140525X1200218X>.
- El Karoui I, King JR, Sitt J, Meyniel F, Van Gaal S, Hasboun D, Adam C, Navarro V, Baulac M, Dehaene S, Cohen L, Naccache L (2015) Event-related potential, time-frequency, and functional connectivity facets of local and global auditory novelty processing: An intracranial study in humans. *Cereb Cortex* 25:4203–4212. <http://dx.doi.org/10.1093/cercor/bhu143>.
- Eliades SJ, Wang X (2008) Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature* 453:1102–1106. <http://dx.doi.org/10.1038/nature06910>.
- Escera C, Malmierca MS (2014) The auditory novelty system: an attempt to integrate human and animal research. *Psychophysiology* 51:111–123. <http://dx.doi.org/10.1111/psyp.12156>.
- Feldman H, Friston KJ (2010) Attention, uncertainty, and free-energy. *Front Hum Neurosci* 4:215. <http://dx.doi.org/10.3389/fnhum.2010.00215>.
- Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1:1–47. <http://dx.doi.org/10.1093/cercor/1.1.1>.
- Fischer C, Morlet D, Giard M (2000) Mismatch negativity and N100 in comatose patients. *Audiol Neurootol* 5:192–197. <http://dx.doi.org/10.1159/000013880>.
- Friston KJ, Harrison L, Penny W (2003) Dynamic causal modelling. *Neuroimage* 19:1273–1302. <http://dx.doi.org/10.1016/B978-012264841-0/50054-8>.
- Friston K (2005) A theory of cortical responses. *Philos Trans R Soc Lond B Biol Sci* 360:815–836. <http://dx.doi.org/10.1098/rstb.2005.1622>.
- Friston K (2010) The free-energy principle: a unified brain theory? *Nat Rev Neurosci* 11:127–138. <http://dx.doi.org/10.1038/nrn2787>.
- Fujioka T, Trainor LJ, Large EW, Ross B (2012) Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. *J Neurosci* 32:1791–1802. <http://dx.doi.org/10.1523/JNEUROSCI.4107-11.2012>.
- Fujioka T, Trainor LJ, Large EW, Ross B (2009) Beta and gamma rhythms in human auditory cortex during musical beat processing. *Ann N Y Acad Sci*:89–92. <http://dx.doi.org/10.1111/j.1749-6632.2009.04779.x>.
- Gagnepain P, Henson RN, Davis MH (2012) Temporal predictive codes for spoken words in auditory cortex. *Curr Biol* 22:615–621. <http://dx.doi.org/10.1016/j.cub.2012.02.015>.
- Garrido MI, Kilner JM, Kiebel SJ, Friston KJ (2007a) Evoked brain responses are generated by feedback loops. *Proc Natl Acad Sci USA* 104:20961–20966. <http://dx.doi.org/10.1073/pnas.0706274105>.
- Garrido MI, Kilner JM, Kiebel SJ, Stephan KE, Friston KJ (2007b) Dynamic causal modelling of evoked potentials: A reproducibility study. *Neuroimage* 36:571–580. <http://dx.doi.org/10.1016/j.neuroimage.2007.03.014>.

- Garrido MI, Friston KJ, Kiebel SJ, Stephan KE, Baldeweg T, Kilner JM (2008) The functional anatomy of the MMN: a DCM study of the roving paradigm. *Neuroimage* 42:936–944. <http://dx.doi.org/10.1016/j.neuroimage.2008.05.018>.
- Garrido MI, Kilner JM, Kiebel SJ, Friston KJ (2009a) Dynamic causal modeling of the response to frequency deviants. *J Neurophysiol* 101:2620–2631. <http://dx.doi.org/10.1152/jn.90291.2008>.
- Garrido MI, Kilner JM, Stephan KE, Friston KJ (2009b) The mismatch negativity: A review of underlying mechanisms. *Clin Neurophysiol*. <http://dx.doi.org/10.1016/j.clinph.2008.11.029>.
- Garrido Marta I, Rowe Elise G, Veronika Halász JBM (2017) Bayesian mapping reveals that attention boosts neural responses to predicted and unpredicted stimulus. *Cereb Cortex* 1–12. <http://dx.doi.org/10.1093/cercor/bhx087>.
- Gilbert JR, Symmonds M, Hanna MG, Dolan RJ, Friston KJ, Moran RJ (2016) Profiling neuronal ion channelopathies with non-invasive brain imaging and dynamic causal models: Case studies of single gene mutations. *Neuroimage* 124:43–53. <http://dx.doi.org/10.1016/j.neuroimage.2015.08.057>.
- Gill P, Woolley SMN, Fremouw T, Theunissen FE (2008) What's that sound? Auditory area CLM encodes stimulus surprise, not intensity or intensity changes. *J Neurophysiol* 99:2809–2820. <http://dx.doi.org/10.1152/jn.01270.2007>.
- Grill-Spector K, Henson R, Martin A (2006) Repetition and the brain: neural models of stimulus-specific effects. *Trends Cogn Sci* 10:14–23. <http://dx.doi.org/10.1016/j.tics.2005.11.006>.
- Grimm S, Escera C (2012) Auditory deviance detection revisited: evidence for a hierarchical novelty system. *Int J Psychophysiol* 85:88–92. <http://dx.doi.org/10.1016/j.ijpsycho.2011.05.012>.
- Grotheer M, Kovács G (2015) Can predictive coding explain repetition suppression? *Cortex*. <http://dx.doi.org/10.1016/j.cortex.2015.11.027>.
- Hohwy J (2013) *The predictive mind*. Oxford: Oxford University Press.
- Horváth J, Winkler I (2004) How the human auditory system treats repetition amongst change. *Neurosci Lett* 368:157–161. <http://dx.doi.org/10.1016/j.neulet.2004.07.004>.
- Hsu Y, Le Bars S, Ha JA (2015) Distinctive representation of mispredicted and unpredicted prediction errors in human electroencephalography. *J Neurosci* 35:14653–14660. <http://dx.doi.org/10.1523/JNEUROSCI.2204-15.2015>.
- Hsu Y-F, Hamalainen J, Waszak F (2014) Both attention and prediction are necessary for adaptive neuronal tuning in sensory processing. *Front Hum Neurosci* 8:152.
- 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* 13:1073–1089. <http://dx.doi.org/10.1006/nimg.2001.0766>.
- Itti L, Koch C, Niebur E (1998) A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans Pattern Anal Mach Intell* 20:1254–1259. <http://dx.doi.org/10.1109/34.730558>.
- Jääskeläinen IP, Ahveninen J, Bonmassar G, Dale AM, Ilmoniemi RJ, Levänen S, Lin F-H, May P, Melcher J, Stufflebeam S, Tiitinen H, Belliveau JW (2004) Human posterior auditory cortex gates novel sounds to consciousness. *Proc Natl Acad Sci USA* 101:6809–6814. <http://dx.doi.org/10.1073/pnas.0303760101>.
- Jaramillo S, Zador AM (2011) The auditory cortex mediates the perceptual effects of acoustic temporal expectation. *Nat Neurosci* 14:246–251. <http://dx.doi.org/10.1038/nn.2688>.
- Keller GB, Hahnloser RHR (2009) Neural processing of auditory feedback during vocal practice in a songbird. *Nature* 457:187–190. <http://dx.doi.org/10.1038/nature07467>.
- Kalman RE (1960) A new approach to linear filtering and prediction problems. *J Basic Eng* 82:35. <http://dx.doi.org/10.1115/1.3662552>.
- Khouri L, Nelken I (2015) Detecting the unexpected. *Curr Opin Neurobiol*. <http://dx.doi.org/10.1016/j.conb.2015.08.003>.
- Kiebel SJ, David O, Friston KJ (2006) Dynamic causal modelling of evoked responses in EEG/MEG with lead field parameterization. *Neuroimage* 30:1273–1284. <http://dx.doi.org/10.1016/j.neuroimage.2005.12.055>.
- Kiebel SJ, Garrido MI, Moran RJ, Friston KJ (2008) Dynamic causal modelling for EEG and MEG. *Cogn Neurodyn* 2:121–136. <http://dx.doi.org/10.1007/s11571-008-9038-0>.
- Knill DC, Pouget A (2004) The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends Neurosci* 27:712–719. <http://dx.doi.org/10.1016/j.tins.2004.10.007>.
- Kogo N, Trengove C (2015) Is predictive coding theory articulated enough to be testable? *Front Comput Neurosci* 9:111. <http://dx.doi.org/10.3389/fncom.2015.00111>.
- Kok P, Rahnev D, Jehee JFM, Lau HC, De Lange FP (2012) Attention reverses the effect of prediction in silencing sensory signals. *Cereb Cortex* 22:2197–2206. <http://dx.doi.org/10.1093/cercor/bhr310>.
- Kok P, Bains LJ, Van Mourik T, Norris DG, De Lange FP (2016) Selective activation of the deep layers of the human primary visual cortex by top-down feedback. *Curr Biol* 26:371–376. <http://dx.doi.org/10.1016/j.cub.2015.12.038>.
- Körding KP, Wolpert DM (2004) Bayesian integration in sensorimotor learning. *Nature* 427:244–247. <http://dx.doi.org/10.1038/nature02169>.
- Lecaignard F, Bertrand O, Gimenez G, Mattout J, Caclin A (2015) Implicit learning of predictable sound sequences modulates human brain responses at different levels of the auditory hierarchy. *Front Hum Neurosci* 9. <http://dx.doi.org/10.3389/fnhum.2015.00505>.
- Lee TS, Mumford D (2003) Hierarchical Bayesian inference in the visual cortex. *J Opt Soc Am A Opt Image Sci Vis* 20:1434–1448. <http://dx.doi.org/10.1364/JOSAA.20.001434>.
- Lieder F, Daunizeau J, Garrido MI, Friston KJ, Stephan KE (2013) Modelling trial-by-trial changes in the mismatch negativity. *PLoS Comput Biol* 9. <http://dx.doi.org/10.1371/journal.pcbi.1002911>.
- López-Caballero F, Zarnowicz K, Escera C (2016) Differential deviant probability effects on two hierarchical levels of the auditory novelty system. *Biol Psychol* 120:1–9. <http://dx.doi.org/10.1016/j.biopsycho.2016.08.001>.
- Malmierca MS, Cristaudo S, Pérez-González D, Covey E (2009) Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat. *J Neurosci* 29:5483–5493. <http://dx.doi.org/10.1523/JNEUROSCI.4153-08.2009>.
- May P, Tiitinen H, Ilmoniemi RJ, Nyman G, Taylor JG, Näätänen R (1999) Frequency change detection in human auditory cortex. *J Comput Neurosci* 6:99–120. <http://dx.doi.org/10.1023/A:1008896417606>.
- May P, Tiitinen H (2010) Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology* 47:66–122. <http://dx.doi.org/10.1111/j.1469-8986.2009.00856.x>.
- Meyniel F, Maheu M, Dehaene S (2016) Human inferences about sequences: A minimal transition probability model. *PLoS Comput Biol*:66–122. <http://dx.doi.org/10.1371/journal.pcbi.1005260>.
- Mittag M, Takegata R, Winkler I (2016) Transitional probabilities are prioritized over stimulus/pattern probabilities in auditory deviance detection: memory basis for predictive sound processing. *J Neurosci* 36:9572–9579. <http://dx.doi.org/10.1523/JNEUROSCI.1041-16.2016>.
- Moran R, Pinotsis DA, Friston K (2013) Neural masses and fields in dynamic causal modeling. *Front Comput Neurosci* 7:57. <http://dx.doi.org/10.3389/fncom.2013.00057>.
- Mumford D (1992) On the computational architecture of the neocortex – II The role of cortico-cortical loops. *Biol Cybern* 66:241–251. <http://dx.doi.org/10.1007/BF00198477>.
- Näätänen R, Gaillard AW, Mäntysalo S (1978) Early selective-attention effect on evoked potential reinterpreted. *Acta Psychol* 42:313–329. [http://dx.doi.org/10.1016/0001-6918\(78\)90006-9](http://dx.doi.org/10.1016/0001-6918(78)90006-9).
- Näätänen R, Pakarinen S, Rinne T, Takegata R (2004) The mismatch negativity (MMN): Towards the optimal paradigm. *Clin Neurophysiol* 115:140–144. <http://dx.doi.org/10.1016/j.clinph.2003.04.001>.
- Näätänen R, Paavilainen P, Rinne T, Alho K (2007) The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin Neurophysiol* 118:2544–2590. <http://dx.doi.org/10.1016/j.clinph.2007.04.026>.

- Nelken I, Fishbach A, Las L, Ulanovsky N, Farkas D (2003) Primary auditory cortex of cats: feature detection or something else? *Biol Cybern* 89:397–406. <http://dx.doi.org/10.1007/s00422-003-0445-3>.
- Nelken I, Yaron A, Polterovich A, Hershenhoren I (2013) Stimulus-specific adaptation beyond pure tones. *Adv Exp Med Biol* 787:411–418. http://dx.doi.org/10.1007/978-1-4614-1590-9_45.
- Opitz B, Rinne T, Mecklinger A, von Cramon DY, Schröger E (2002) Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. *Neuroimage* 15:167–174. <http://dx.doi.org/10.1006/nimg.2001.0970>.
- Paavilainen P, Aarjälvi P, Takegata R (2007) Preattentive detection of nonsalient contingencies between auditory features. *Neuroreport* 18:159–163. <http://dx.doi.org/10.1097/WNR.0b013e328010e2ac>.
- Paavilainen P (2013) The mismatch-negativity (MMN) component of the auditory event-related potential to violations of abstract regularities: A review. *Int J Psychophysiol*. <http://dx.doi.org/10.1016/j.ijpsycho.2013.03.015>.
- 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. <http://dx.doi.org/10.1523/JNEUROSCI.5095-14.2015>.
- 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. <http://dx.doi.org/10.1016/j.cortex.2016.05.001>.
- Posner MI (1980) Orienting of attention. *Q J Exp Psychol* 32:3–25. <http://dx.doi.org/10.1080/00335558008248231>.
- Raij T, McEvoy L, Mäkelä JP, Hari R (1997) Human auditory cortex is activated by omissions of auditory stimuli. *Brain Res* 745:134–143. [http://dx.doi.org/10.1016/S0167-8760\(97\)85548-1](http://dx.doi.org/10.1016/S0167-8760(97)85548-1).
- Rao RPN (2005) Bayesian inference and attentional modulation in the visual cortex. *Neuroreport* 16:1843–1848. <http://dx.doi.org/10.1097/01.wnr.0000183900.92901.fc>.
- Rao RPN, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci* 2:79–87. <http://dx.doi.org/10.1038/4580>.
- Ross S, Hansen NC (2016) Dissociating prediction failure: considerations from music perception. *J Neurosci* 36:3103–3105. <http://dx.doi.org/10.1523/JNEUROSCI.0053-16.2016>.
- Rubin J, Ulanovsky N, Nelken I, Tishby N (2016) The representation of prediction error in auditory cortex. *PLoS Comput Biol* 12:e1005058. <http://dx.doi.org/10.1371/journal.pcbi.1005058>.
- Rummell BP, Klee JL, Sigurdsson XT (2016) Attenuation of responses to self-generated sounds in auditory cortical neurons. *J Neurosci* 36:12010–12026. <http://dx.doi.org/10.1523/JNEUROSCI.1564-16.2016>.
- Saarinen J, Paavilainen P, Schöger E, Tervaniemi M, Näätänen R (1992) Representation of abstract attributes of auditory-stimuli in the human brain. *Neuroreport* 3:1149–1151.
- Sams M, Alho K, Näätänen R (1983) Sequential effects on the ERP in discriminating two stimuli. *Biol Psychol* 17:41–58. [http://dx.doi.org/10.1016/0301-0511\(83\)90065-0](http://dx.doi.org/10.1016/0301-0511(83)90065-0).
- SanMiguel I, Saupé K, Schröger E (2013b) I know what is missing here: electrophysiological prediction error signals elicited by omissions of predicted "what" but not "when". *Front Hum Neurosci* 7. <http://dx.doi.org/10.3389/fnhum.2013.00407>.
- SanMiguel I, Widmann A, Bendixen A, Trujillo-Barreto N, Schröger E (2013a) Hearing silences: human auditory processing relies on preactivation of sound-specific brain activity patterns. *J Neurosci* 33:8633–8639. <http://dx.doi.org/10.1523/JNEUROSCI.5821-12.2013>.
- Scheeringa R, Koopmans PJ, van Mourik T, Jensen O, Norris DG (2016) The relationship between oscillatory EEG activity and the laminar-specific BOLD signal. *Proc Natl Acad Sci USA* 113:6761–6766. <http://dx.doi.org/10.1073/pnas.1522577113>.
- Schröger E, Bendixen A, Denham SL, Mill RW, Bohm TM, Winkler I (2014) Predictive regularity representations in violation detection and auditory stream segregation: From conceptual to computational models. *Brain Topogr*. <http://dx.doi.org/10.1007/s10548-013-0334-6>.
- Schröger E, Marzecová A, SanMiguel I (2015) Attention and prediction in human audition: a lesson from cognitive psychophysiology. *Eur J Neurosci* 41:641–664. <http://dx.doi.org/10.1111/ejn.12816>.
- Sculthorpe LD, Ouellet DR, Campbell KB (2009) MMN elicitation during natural sleep to violations of an auditory pattern. *Brain Res* 1290:52–62. <http://dx.doi.org/10.1016/j.brainres.2009.06.013>.
- Sedley W, Gander PE, Kumar S, Kovach CK, Oya H, Kawasaki H, Howard MA, Griffiths TD (2016) Neural signatures of perceptual inference. *Elife* 5. <http://dx.doi.org/10.7554/eLife.11476>.
- Self MW, Kooijmans RN, Supèr H, Lamme VA, Roelfsema PR (2012) Different glutamate receptors convey feedforward and recurrent processing in macaque V1. *Proc Natl Acad Sci USA* 109:11031–11036. <http://dx.doi.org/10.1073/pnas.1119527109>.
- Shipp S (2016) Neural elements for predictive coding. *Front Psychol* 7:1792. <http://dx.doi.org/10.3389/FPSYG.2016.01792>.
- Sohoglu E, Chait M (2016a) Neural dynamics of change detection in crowded acoustic scenes. *Neuroimage* 126:164–172. <http://dx.doi.org/10.1016/j.neuroimage.2015.11.050>.
- Sohoglu E, Chait M (2016b) Detecting and representing predictable structure during auditory scene analysis. *Elife* 5:1–17. <http://dx.doi.org/10.7554/eLife.19113>.
- Sohoglu E, Peelle JE, Carlyon RP, Davis MH (2012) Predictive top-down integration of prior knowledge during speech perception. *J Neurosci* 32:8443–8453. <http://dx.doi.org/10.1523/JNEUROSCI.5069-11.2012>.
- 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*.
- Spratling MW (2008a) Predictive coding as a model of biased competition in visual attention. *Vision Res* 48:1391–1408. <http://dx.doi.org/10.1016/j.visres.2008.03.009>.
- Spratling MW (2008b) Reconciling predictive coding and biased competition models of cortical function. *Front Comput Neurosci* 2:4. <http://dx.doi.org/10.3389/neuro.10.004.2008>.
- Spratling MW (2010) Predictive coding as a model of response properties in cortical area V1. *J Neurosci* 30:3531–3543. <http://dx.doi.org/10.1523/JNEUROSCI.4911-09.2010>.
- Spratling MW (2015) A review of predictive coding algorithms. *Brain Cogn*. <http://dx.doi.org/10.1016/j.bandc.2015.11.003>.
- Strauss M, Sitt JD, King J-R, Elbaz M, Azizi L, Buiatti M, Naccache L, van Wassenhove V, Dehaene S (2015) Disruption of hierarchical predictive coding during sleep. *Proc Natl Acad Sci USA* 112:E1353–E1362. <http://dx.doi.org/10.1073/pnas.1501026112>.
- Sussman E, Ritter W, Vaughan HG (1998) Predictability of stimulus deviance and the mismatch negativity. *Neuroreport* 9:4167–4170. <http://dx.doi.org/10.1097/00001756-199812210-00031>.
- Szymanski FD, Garcia-Lazaro JA, Schnupp JWH (2009) Current source density profiles of stimulus-specific adaptation in rat auditory cortex. *J Neurophysiol* 102:1483–1490. <http://dx.doi.org/10.1152/jn.00240.2009>.
- Taaseh N, Yaron A, Nelken I (2011) Stimulus-specific adaptation and deviance detection in the rat auditory cortex. *PLoS One* 6:e23369. <http://dx.doi.org/10.1371/journal.pone.0023369>.
- Todorovic A, de Lange FP (2012) Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. *J Neurosci* 32:13389–13395. <http://dx.doi.org/10.1523/JNEUROSCI.2227-12.2012>.
- 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* 31:9118–9123. <http://dx.doi.org/10.1523/JNEUROSCI.1425-11.2011>.
- Todorovic A, Schoffelen JM, Van Ede F, Maris E, De Lange FP (2015) Temporal expectation and attention jointly modulate auditory oscillatory activity in the beta band. *PLoS One* 10. <http://dx.doi.org/10.1371/journal.pone.0120288>.

- 2053 Uhrig L, Dehaene S, Jarraya B (2014) A hierarchy of responses to
2054 auditory regularities in the macaque brain. *J Neurosci*
2055 34:1127–1132. [http://dx.doi.org/10.1523/JNEUROSCI.3165-](http://dx.doi.org/10.1523/JNEUROSCI.3165-13.2014)
2056 [13.2014](http://dx.doi.org/10.1523/JNEUROSCI.3165-13.2014).
- 2057 Uhrig L, Janssen D, Dehaene S, Jarraya B (2016) Cerebral
2058 responses to local and global auditory novelty under general
2059 anesthesia. *Neuroimage* 141:326–340. [http://dx.doi.org/10.1016/](http://dx.doi.org/10.1016/j.neuroimage.2016.08.004)
2060 [j.neuroimage.2016.08.004](http://dx.doi.org/10.1016/j.neuroimage.2016.08.004).
- 2061 Ulanovsky N, Las L, Farkas D, Nelken I (2004) Multiple time scales of
2062 adaptation in auditory cortex neurons. *J Neurosci*
2063 24:10440–10453. [http://dx.doi.org/10.1523/JNEUROSCI.1905-](http://dx.doi.org/10.1523/JNEUROSCI.1905-04.2004)
2064 [04.2004](http://dx.doi.org/10.1523/JNEUROSCI.1905-04.2004).
- 2065 Ulanovsky N, Las L, Nelken I (2003) Processing of low-probability
2066 sounds by cortical neurons. *Nat Neurosci* 6:391–398. [http://dx.](http://dx.doi.org/10.1038/nn1032)
2067 [doi.org/10.1038/nn1032](http://dx.doi.org/10.1038/nn1032).
- 2068 Umbricht D, Koller R, Vollenweider FX, Schmid L (2002) Mismatch
2069 negativity predicts psychotic experiences induced by NMDA
2070 receptor antagonist in healthy volunteers. *Biol Psychiatry*
2071 51:400–406. [http://dx.doi.org/10.1016/S0006-3223\(01\)01242-2](http://dx.doi.org/10.1016/S0006-3223(01)01242-2).
- 2072 van Kerkoerle T, Self MW, Dagnino B, Gariel-Mathis M-A, Poort J,
2073 van der Togt C, Roelfsema PR (2014) Alpha and gamma
2074 oscillations characterize feedback and feedforward processing
2075 in monkey visual cortex. *Proc Natl Acad Sci USA*
2076 111:14332–14341. <http://dx.doi.org/10.1073/pnas.1402773111>.
- 2077 van Pelt S, Heil L, Kwisthout J, Ondobaka S, van Rooij I, Bekkering H
2078 (2016) Beta- and gamma-band activity reflect predictive coding in
2079 the processing of causal events. *Soc Cogn Affect Neurosci*
2080 11:973–980. <http://dx.doi.org/10.1093/scan/nsw017>.
- 2081 Wacongne C, Labyt E, van Wassenhove V, Bekinschtein T,
2082 Naccache L, Dehaene S (2011) Evidence for a hierarchy of
2083 predictions and prediction errors in human cortex. *Proc Natl Acad Sci USA*
2084 108:20754–20759. [http://dx.doi.org/10.1073/](http://dx.doi.org/10.1073/pnas.1117807108)
2085 [pnas.1117807108](http://dx.doi.org/10.1073/pnas.1117807108).
- 2086 Wacongne C, Changeux J-P, Dehaene S (2012) A neuronal model of
2087 predictive coding accounting for the mismatch negativity. *J Neurosci*
2088 32:3665–3678. [http://dx.doi.org/10.1523/](http://dx.doi.org/10.1523/JNEUROSCI.5003-11.2012)
2089 [JNEUROSCI.5003-11.2012](http://dx.doi.org/10.1523/JNEUROSCI.5003-11.2012).
- Wallace MN, Palmer AR (2008) Laminar differences in the response
2090 properties of cells in the primary auditory cortex. *Exp Brain Res*
2091 184:179–191. <http://dx.doi.org/10.1007/s00221-007-1092-z>.
2092
- Winer JA (1985) Structure of layer II in cat primary auditory cortex
2093 (AI). *J Comp Neurol* 238:10–37. [http://dx.doi.org/10.1002/](http://dx.doi.org/10.1002/cne.902380103)
2094 [cne.902380103](http://dx.doi.org/10.1002/cne.902380103).
- Winkler I, Czigler I (1998) Mismatch negativity: deviance detection or
2095 the maintenance of the “standard”. *Neuroreport* 9:3809–3813.
2096
- Winkler I, Czigler I (2012) Evidence from auditory and visual event-
2097 related potential (ERP) studies of deviance detection (MMN and
2098 vMMN) linking predictive coding theories and perceptual object
2099 representations. *Int J Psychophysiol*. [http://dx.doi.org/10.1016/j.](http://dx.doi.org/10.1016/j.ijpsycho.2011.10.001)
2100 [ijpsycho.2011.10.001](http://dx.doi.org/10.1016/j.ijpsycho.2011.10.001).
- Winkler I, Denham SL, Nelken I (2009) Modeling the auditory scene:
2101 predictive regularity representations and perceptual objects.
2102 *Trends Cogn Sci*. <http://dx.doi.org/10.1016/j.tics.2009.09.003>.
2103
- Winkler I, Schröger E (2015) Auditory perceptual objects as
2104 generative models: Setting the stage for communication by
2105 sound. *Brain Lang* 148:1–22. [http://dx.doi.org/10.1016/j.](http://dx.doi.org/10.1016/j.bandl.2015.05.003)
2106 [bandl.2015.05.003](http://dx.doi.org/10.1016/j.bandl.2015.05.003).
- Yabe H, Tervaniemi M, Reinikainen K, Näätänen RN (1997)
2107 Temporal window of integration revealed by MMN to sound
2108 omission. *Neuroreport* 8:1971–1974. [http://dx.doi.org/10.1097/](http://dx.doi.org/10.1097/00001756-199705260-00035)
2109 [00001756-199705260-00035](http://dx.doi.org/10.1097/00001756-199705260-00035).
- Yaron A, Hershenhoren I, Nelken I (2012) Sensitivity to complex
2110 statistical regularities in rat auditory cortex. *Neuron* 76:603–615.
2111 <http://dx.doi.org/10.1016/j.neuron.2012.08.025>.
2112
- Angela JY (2014) Bayesian Models of Attention. In: *The Oxford Handbook of Attention*. Oxford: Oxford University Press.
2113
- Yu AJ, Dayan P (2005a) Inference, attention, and decision in a
2114 Bayesian neural architecture. *Adv Neural Inf Process Syst*:1577–1584.
2115
- Yu AJ, Dayan P (2005b) Uncertainty, neuromodulation, and attention.
2116 *Neuron* 46:681–692. [http://dx.doi.org/10.1016/j.](http://dx.doi.org/10.1016/j.neuron.2005.04.026)
2117 [neuron.2005.04.026](http://dx.doi.org/10.1016/j.neuron.2005.04.026).
- Zhao J, Al-Aidroos N, Turk-Browne NB (2013) Attention is
2118 spontaneously biased toward regularities. *Psychol Sci*
2119 24:667–677. <http://dx.doi.org/10.1177/0956797612460407>.
2120
2121
2122
2123
2124
2125
2126
2127

(Received 29 May 2017, Accepted 26 July 2017)
(Available online xxxx)

2128
2129
2130