Journal Club

Editor's Note: These short, critical reviews of recent papers in the *Journal*, written exclusively by graduate students or postdoctoral fellows, are intended to summarize the important findings of the paper and provide additional insight and commentary. For more information on the format and purpose of the Journal Club, please see http://www.jneurosci.org/misc/ifa_features.shtml.

Dissociating Prediction Failure: Considerations from Music Perception

Suzi Ross* and Niels Chr. Hansen*

Center for Music in the Brain, Department of Clinical Medicine, Aarhus University and The Royal Academy of Music Aarhus/Aalborg, 8000 Aarhus C, Denmark

Review of Hsu et al.

Recent years have seen an abundance of theoretical and empirical work framing human cognition and neuroscience in the context of predictive processing (Friston, 2009; Vuust et al., 2009; Clark, 2013; Hansen and Pearce, 2014). One key tenet of predictive coding theory is the proposal that perception is a hierarchical process in which prediction errors, which occur when the activity in a given level cannot be explained by descending predictions from the next higher level, are passed upwards in the processing hierarchy (Friston, 2009). This gives rise to neurophysiological activity, which can be measured noninvasively with electroencephalography (EEG) and magnetoencephalography (MEG), as an early negative deflection called N1 and a mismatch negativity (MMN) response.

A recent EEG study by Hsu et al. (2015) aimed to address a key question in current predictive coding research regarding whether two types of prediction failure, where people either fail to predict correctly (mispredicted) or fail to predict at all (unpredicted), are neurophysiologically dissociable. To this end, the authors

devised an experiment in the auditory domain, comparing event-related potentials for these two types of events to each other and to predicted events where predictions were fulfilled. Specifically, electrophysiological recordings were time-locked to the final tone of five-tone sequences drawn from the C major scale, i.e., C-D-E-F-G-A-B (Hsu et al., 2015, their Fig. 2). Stimuli comprised four stepwise sinusoidal tones followed by either a predicted continuation of the ascending pattern (e.g., D-E-F-G-A) or a mispredicted presentation of a tone four notes below the preceding tone (e.g., D-E-F-G-C). In the unpredicted condition, five tones were presented in random order (e.g., G-D-E-A-F).

Hsu et al. (2015) obtained evidence for differential processing, specifically N1 attenuation for unpredicted stimuli and N1 enhancement for mispredicted stimuli, compared with predicted stimuli. Because this difference was absent in a control experiment with ascending (e.g., G-A) and descending (e.g., G-C) tone pairs, the main findings could not be ascribed to physical differences in stimulus frequency.

Importantly, Hsu et al.'s (2015) findings of medium N1 responses for predicted stimuli are inconsistent with Arnal and Giraud's (2012) theory that these contexts would generate minimal prediction error because input would be fully accounted for. As noted by Hsu et al. (2015), this previous model assumed an all-or-none representation of prediction,

in that prediction was regarded as either entirely fulfilled or unfulfilled (Hsu et al., 2015, their Fig. 1A). Crucially, this view differs from the distributional encoding of prediction posited by traditional predictive coding models (Friston, 2009). In this conception, prediction is imprecise because its underlying probability function attains either the shape of a continuous unimodal density function (Hsu et al., 2015, their Fig. 1B), or that of a discrete distribution with multiple peaks where several event categories are anticipated to varying degrees due to internalized environmental regularities (Hsu et al., 2015, their Fig. 1C). In both cases, while a correctly predicted event would explain away sensory input, prediction error would still be triggered by weaker predictions remaining unfulfilled. This would induce medium prediction error relative to enhanced error from mispredicted events and attenuated error from unpredicted events (Hsu et al., 2015, their Fig. 1B,C).

The study by Hsu et al. (2015) makes two important contributions, with implications for the design of future predictive coding research. These will be discussed in turn, specifically in the light of theoretical and empirical work on uncertainty and tonal hierarchies.

First, the neurophysiological dissociation between distinct types of prediction failure raises interesting questions regarding how prediction operates under states of uncertainty. Predictive coding theory posits that humans are always engaged in
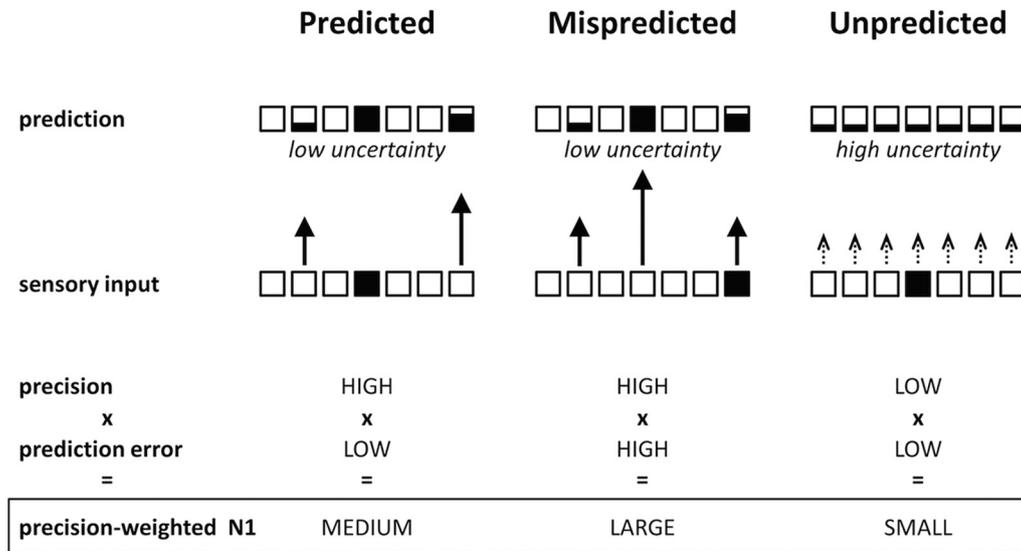
Figure 1. Revised hypothesis concerning precision-weighted prediction error as reflected in the N1 response. Boxes represent stimulus dimensions colored according to their intensity, and lengths of the arrows reflect the intensity of prediction error. This revised model is similar to Hsu et al.'s (2015, their Figure 1C) with the important exception that unpredicted stimuli are governed by predictions with high levels of uncertainty (rather than no predictions at all). This is hypothesized to cause low precision-weighting of prediction error, ultimately resulting in attenuated N1 responses for stimuli in this condition.

predictive processing simply because it would be maladaptive for the brain not to adopt an anticipatory state (Friston, 2009). In other words, without prediction, perception would not be possible. The interpretation that prediction was entirely absent in the unpredicted condition, leaving only forward-propagating input, therefore seems inconsistent with predictive coding theory itself. An alternative interpretation is that these stimuli indeed generate expectations in the listener, but that these expectations are characterized by high degrees of uncertainty representative of a weak predictive model. This can be quantified with Shannon entropy, which has been shown to relate to both explicit and more indirect measures of predictive uncertainty (Hansen and Pearce, 2014). Predictive coding theory sometimes formalizes it as precision-weighting of the prediction error and proposes that this may be encoded in terms of neuromodulation of synaptic gain (Friston, 2009). Attenuated N1 responses to unpredicted stimuli may thus be due to weak and imprecise expectations causing low synaptic gain rather than to a complete lack of top-down predictions. Because this possibility was merely acknowledged in passing, we explore it further in our Figure 1, complementing Hsu et al.'s (2015) Figure 1, A–C.

Second, Hsu et al. (2015) provide significant empirical weight to the interpretation of prediction as a distributionally rather than categorically represented phenomenon, where multiple graded predictions are generated simultaneously (cf. Hansen and Pearce, 2014). While the study design did not allow differentiation between prediction as a Gaussian function or one reflecting the statistical regularities of the environment, and although the two are not necessarily mutually exclusive, evidence from music perception research support the latter model. Specifically, tonal hierarchies, acquired through long-term exposure to music, designate how certain pitches in the scale are perceived as more stable than others (Krumhansl and Cuddy, 2010). This leads highly unstable tones to evoke anticipation of a more stable tone in the hierarchy (Huron, 2006). For instance, in the key of C major, the tonal hierarchy affords that the relatively unstable pitch of B is strongly expected to be followed by the highly stable pitch of C. Neurophysiological research has demonstrated different cortical activity for stable versus unstable pitches (Krohn et al., 2007), including greater MMN responses to unstable pitch intervals compared with stable ones even if there is a greater physical distance between the notes in the stable interval than in the unstable interval (Brattico et al., 2000).

Thus, when using tones from the major scale in a listening experiment like that of Hsu et al. (2015), two potential sources of knowledge are available to participants to generate predictions about the final tone in the sequence: local probabilistic knowledge gained during the experiment itself and prior knowledge manifested in tonal hierarchies reflecting statistical properties of melodic sequences acquired over the course of a lifetime. Regarding local probabilistic knowledge, prediction was informed by the fact that predicted stimuli occurred six times more frequently than mispredicted or unpredicted stimuli. The authors therefore noted that participants could have predicted both potential outcomes, with continuation of the established ascending stepwise pattern constituting the strongest prediction and violation of the established pattern as a weaker one. This would indeed cause medium prediction error for predicted stimuli.

Moreover, given the non-equidistant nature of the major scale, combining semitone steps (e.g., E-F) with whole-tone steps (e.g., C-D, D-E, F-G, G-A, A-B), each four-tone set could be uniquely identified purely from local knowledge before the occurrence of the final tone. Specifically, tone sets were derived from a restricted pool of seven pitches (C4 to B4), giving rise to four implicative tone sets with only one possible final tone in the lowest and highest sets (i.e., C-D-E-F- could only be followed by G; D-E-F-G- could be followed by A or C; E-F-G-A- could be followed by B or D; and F-G-A-B- could only be followed by E). This would, in theory, allow participants to identify the lowest and highest four-tone sets that deterministically afforded predicted and mispredicted continuations, respectively. Similarly, the fifth tone complementing unpredicted four-tone sets was sometimes deterministic, in that a four-tone

set containing extremity tones (e.g., D-G-E-A) must be followed by the implied missing tone (i.e., F), and other times 50% predictable, in that four-tone sets not containing both extremity tones (e.g., D-G-E-F) could only be followed by the absent extremity (C or A). This may potentially have contributed to N1 attenuation in this condition. While it remains unknown whether listeners are able to internalize and retrieve such complex local probabilities, sophisticated predictive processing capabilities of non-musicians suggest that this is possible to some extent (Bigand and Poulin-Charronnat, 2006).

Regarding prior knowledge of the statistical properties of music, the assumption that stepwise ascending four-tone sequences imply continuations in the same direction is supported by both electrophysiological research, where unexpected note repetitions or contour reversals produce MMN responses (Tervaniemi et al., 1994), and by behavioral research validating the Gestalt principle of "good continuation" in melodic expectation (Krumhansl, 1995). These expectations, ascribable to long-term exposure to music (Pearce and Wiggins, 2006), are effectively capitalized upon by Hsu et al. (2015) in their predicted stimuli design. First-order conditional probabilities based on ~250,000 tone pairs from a tonally representative corpus of Germanic folksongs in the major key confirm that predicted continuations of three of the four implicative four-tone sets mentioned above indeed have higher probabilities than mispredicted ones (Huron, 2006, their Table 9.2). However, in the specific case of D-E-F-G, the continuation tone C actually has a probability of 13.3% whereas A only has a probability of 10.8%. Although in this case, local probabilistic knowledge (i.e., the fact that participants heard the predicted con-

tinuation six times more often than the mispredicted continuation) would most likely bias predictions in the intended direction. Accounting for tonal influences would offer a particularly interesting avenue for future work in this area. Indeed, use of equidistant pitches from an unfamiliar scale, such as the Bohlen-Pierce Scale (Loui et al., 2010), could potentially eliminate such influences from long-term exposure.

In conclusion, Hsu et al.'s (2015) study contributes significantly to predictive coding research by demonstrating differences in neurophysiological processing in situations where prediction fails in fundamentally different ways. Also, by demonstrating that prediction is not an all-or-none phenomenon, but is represented as probability distributions, they provide key insight regarding the nature of the underlying coding mechanisms. We have substantiated how predictive uncertainty and precision-weighting of prediction error may help further guide this work. Existing research on tonal hierarchies in music perception may, moreover, elucidate the influence of environmental statistical regularities on prediction. Accounting for predictive uncertainty and tonal hierarchies resulting from long-term acquisition of probabilistic knowledge may serve to further validate these results as well as offer potentially promising avenues for advancing scientific knowledge about prediction, both when it fails and when it succeeds.

References

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

Bigand E, Poulin-Charronnat B (2006) Are we "experienced listeners"? A review of the musical capacities that do not depend on formal musical training. Cognition 100:100–130. CrossRef Medline

Brattico E, Näätänen R, Verma T, Välimäki V, Tervaniemi M (2000) Processing of musical intervals in the central auditory system: an

event-related potential (ERP) study on sensory consonance. In: Proceedings of the Sixth International Conference on Music Perception and Cognition (Woods C, Luck GB, Brochard R, O'Neill SA, Sloboda JA, eds.). Keele, UK: Keele University.

Clark A (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav Brain Sci 36:181–204. CrossRef Medline

Friston K (2009) The free-energy principle: A rough guide to the brain? Trends Cogn Sci 13:293–301. CrossRef Medline

Hansen NC, Pearce MT (2014) Predictive uncertainty in auditory sequence processing. Front Psychol 5:1052. CrossRef Medline

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

Huron D (2006) Sweet anticipation: music and the psychology of expectation. Cambridge, MA: MIT.

Krohn KI, Brattico E, Välimäki V, Tervaniemi M (2007) Neural representations of the hierarchical scale pitch structure. Music Percept 24:281–296. CrossRef

Krumhansl CL (1995) Music psychology and music theory: problems and prospects. Music Theory Spectrum 17:53–80. CrossRef

Krumhansl CL, Cuddy LL (2010) A theory of tonal hierarchies in music. In: Music perception (Jones MR, Fay RR, Popper AN, eds.), pp 51–87. New York: Springer.

Loui P, Wessel DL, Hudson Kam CL (2010) Humans rapidly learn grammatical structure in a new musical scale. Music Percept 27:377–388. CrossRef Medline

Pearce M, Wiggins G (2006) Expectation in melody: the influence of context and learning. Music Percept 23:377–405. CrossRef

Tervaniemi M, Maury S, Näätänen R (1994) Neural representations of abstract stimulus features in the human brain as reflected by the mismatch negativity. Neuroreport 5:844–846. CrossRef Medline

Vuust P, Ostergaard L, Pallesen KJ, Bailey C, Roepstorff A (2009) Predictive coding of music-brain responses to rhythmic incongruity. Cortex 45:80–92. CrossRef Medline