Schizophrenia Research xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Schizophrenia Research

journal homepage: www.elsevier.com/locate/schres

Time as context: The influence of hierarchical patterning on sensory inference

Juanita Todd [a,b,*], Anne Petherbridge [a], Bronte Speirs [a], Alexander Provost [a,b], Bryan Paton [a,b]

[a] *University of Newcastle, Callaghan, NSW 2308, Australia*
[b] *Hunter Medical Research Institute, New Lambton Heights, NSW 2305, Australia*

ARTICLE INFO

Article history:
Received 27 January 2017
Received in revised form 10 March 2017
Accepted 13 March 2017
Available online xxxx

Keywords:
Mismatch negativity
Temporal processing
Context
Hierarchical inference
Predictive coding

ABSTRACT

Time, or more specifically temporal structure, is a critical variable in understanding how the auditory system uses acoustic patterns to predict input, and to filter events based on their relevance. A key index of this filtering process is the auditory evoked potential component known as mismatch negativity or MMN. In this paper we review findings of smaller MMN in schizophrenia through the lens of time as an influential contextual variable. More specifically, we review studies that show how MMN to a locally rare pattern-deviation is modulated by the longer-term context in which it occurs. Empirical data is presented from a non-clinical sample confirming that the absence of a stable higher-order structure to sound sequences alters the way MMN amplitude changes over time. This result is discussed in relation to how hierarchical pattern learning might enrich our understanding of how and why MMN amplitude modulation is disrupted in schizophrenia.

1. Introduction

Time is the canvas upon which patterns in sound evolve. Time is therefore a very important contextual variable in determining how the relevance of sound will be filtered. In this paper we emphasise why careful consideration of temporal structure on multiple timescales can provide insight into how the brain modulates sensory beliefs. The paper features a test of two hypotheses related to how sensory filters[1] are modified based on a long time course of accumulated information. Both the hypotheses and results are discussed with a particular focus on how these factors might contribute to an understanding of abnormal sensory filtering in persons with schizophrenia.

In a laboratory setting, the exploration of auditory sensory filtering typically involves presenting experimental participants with a sequence of sounds over headphones. Through careful composition of the sound sequences, the experimenter can assess how the brain's response alters over time. The alterations in auditory system response are used to infer the action of sensory filtering processes (Naatanen, 1992). These responses can be assessed using scalp-recorded measures of auditory evoked potentials (AEPs) by having participants wear an electrode cap (Coles and Rugg, 1996). Research of this kind has provided substantial insight into how auditory sensory filtering processes work (Naatanen et al., 2007; Naatanen, 2008), and has revealed reliable differences in how AEPs are altered in persons with schizophrenia (see Naatanen and Kahkonen, 2009; Näätänen et al., 2015; Todd et al., 2013a, b; Light and Swerdlow, 2015 for reviews).

A great deal of interest in the study of sensory filtering centres on a component of the AEP called mismatch negativity or MMN. MMN is employed to study the integrity of sensory filtering, with both its elicitation and amplitude being useful indicators of the underlying process (Kujala et al., 2007). MMN is elicited whenever the brain has detected an unexpected change in a regular repeating pattern in sound. MMN amplitude is thought to reflect some quantification of how important this deviation is, with this quantification impacted by an accumulated estimate of "confidence" in the degree to which it is a reliable deviation from expectations (Winkler, 2007; Naatanen et al., 2001). For example, in the simplest sequence the pattern may be an identical repeating tone and the deviation, a rare and unexpected change in a physical feature. In this case the amplitude of the AEP to the deviation will be influenced by at least three factors: the degree of physical difference (the more it departs from the repeated features, the larger the response e.g., Javitt et al., 1998); the rarity of the occurrence of the deviant (the more common the standard, the larger the response to deviations e.g., Csepe et al., 1987); and the period of stability in the pattern (generally the longer a sound has been a regularity, the larger the response to deviants e.g., Todd et al., 2011 but see further discussion below).

---

* Corresponding author at: University of Newcastle, University Drive, Callaghan, NSW 2308, Australia.
E-mail address: Juanita.Todd@newcastle.edu.au (J. Todd).

[1] By sensory filtering, we refer to how meaning is extracted from sensory streams to enable perceptual synthesis. From the point of view of predictive coding, this corresponds to Bayesian (sensory) filtering. This is a perspective that we will exploit; especially in relation to the hierarchical nature of generative models implicit in Bayesian filtering.

http://dx.doi.org/10.1016/j.schres.2017.03.033
0920-9964/

Please cite this article as: Todd, J., et al., Time as context: The influence of hierarchical patterning on sensory inference, Schizophr. Res. (2017), http://dx.doi.org/10.1016/j.schres.2017.03.033

There are now two meta-analyses confirming smaller MMN amplitude in persons with schizophrenia (Umbricht and Krljes, 2005; Erickson et al., 2016) with a large effect size (Cohen, 1988). Here we wish to emphasise a potentially important observation within the more recent of the two – that is, that the rarity of the deviant does not significantly impact the effect size of the observed reduction in MMN amplitude. This meta-analytic finding is at odds with studies that have observed much larger group differences under conditions that produce the largest MMN amplitudes in controls. For example, two independent groups have demonstrated much larger group differences when the pattern deviation is very rare versus more common (Sato et al., 2003; Javitt et al., 1998). A key difference between these studies, and meta-analytic findings, is that the former feature within-subject manipulations of deviant rarity, while the meta-analysis included assessment of the impact of different deviant probabilities across studies that most often include only one probability level. This observation brings into focus the important role of context in the dynamics of AEPs, and the sensory filtering process.

As a biological signal the MMN has a limited dynamic range governed by gain control. Gain control refers to a divisive normalisation process in which the change in neural response is adjusted "to efficiently use the available dynamic range, maximizing sensitivity to changes in input" (p56, Carandini and Heeger, 2012). Cognitive Neuroscience Treatment Research to Improve Cognition in Schizophrenia (CNTRICS) has promoted MMN as an established index of gain control in the auditory system and its dysfunction in persons with schizophrenia (Carter and Barch, 2007). Predictive coding accounts of MMN explain that gain control (and the normalisation process underlying it) relies on both the updating of regularities upon which an internal model is formed, and on the estimate of the precision of those regularities (Friston, 2005). The role of precision or confidence will be crucial for the arguments below and it distinguishes our and Friston's account from more generic predictive coding explanations (Allen and Friston, 2016). In brief, not only do we have to accumulate the evidence for the perceptual content or beliefs in what causes our sensations, we also accumulate evidence for the confidence we place in those predictions. In our predictive coding account, this corresponds to updating the precision (or inverse variance) associated with prediction errors at various levels in the cortical hierarchy, assuming the distribution of beliefs is approximately Gaussian. For example, if we have been exposed to a very predictable context for a sufficient amount of time, then our higher-level predictions will be held with greater confidence (i.e., high precision), making them relatively impervious to (disconfirmatory) sensory evidence. A key aspect of hierarchal predictive coding is that the relative precision of high-level representations, relative to sensory precision, determines the degree to which sensory evidence updates higher-level beliefs. Crucially, precision itself has to not only be inferred or learned but we also form expectations for the relative precision of sensory evidence. We will refer to this as contextual learning or precision updating. Precision updating can be wrong and might explain many aspects of Autism, where there is thought to be problems with matching beliefs about sensory and top down precisions (Lawson et al., 2014). Similarly, a disorder of inference regarding precisions might help explain many puzzling aspects of schizophrenia (Friston et al., 2016).

In terms of the MMN, we can explain an exuberant response to an oddball stimulus as follows: as standard stimuli are repeated, we become increasingly confident in our top-down predictions and rely less on sensory prediction errors. This results in an attenuation of sensory precision and a reduction in the amplitude of stimulus bound AEP responses. However, should an oddball stimulus herald a change in context, we lose confidence in our predictions and attend to the sensory evidence at hand by suspending sensory attenuation. This produces a larger ERP and implicit MMN (see investigations of model updating to standards after deviants in Winkler et al., 1996). The key point here is that the expression of the MMN depends upon contextual learning or precision updating that may be sensitive to information derived over multiple timescales.

The estimate of precision (more specifically the posterior expectation of precision at different hierarchical levels) is something that is accumulated over a longer time course than the prediction itself, given that a new internal model for simple repetitions can be established in as little as 2–3 repetitions (Sams et al., 1983). Precision estimates are therefore contextual, based on the statistics of the immediate environment e.g. the list of stimuli delivered during an experiment. This relative nature could explain why a failure to adequately modulate MMN in accordance with such estimates can be so pronounced in within-subjects manipulations in schizophrenia groups, but not be evident in cross-study based meta-analyses.

MMN reduction in schizophrenia has been proposed as a sensory-inference level consequence of the dysconnection hypothesis, itself put forward to explain many signs and symptoms in schizophrenia (Friston et al., 2016). In accordance with the dysconnection hypothesis, the cornerstone vulnerability within schizophrenia is "not an inability to predict sensory content, but *failure to encode the relative confidence* that should be placed on sensory evidence, *relative to prior beliefs*" (p5). This idea has been applied for example to accounts of the generation of auditory verbal hallucinations in schizophrenia, where prior beliefs/expectations may be over-valued (e.g. Horga et al., 2014; Fletcher and Frith, 2009). Ideally, to test this hypothesis, one must separate prediction from confidence or precision estimates, and assess the influence of prior beliefs. An exquisite realization of this goal has recently been published for a visual behavioural task (Marshall et al., 2016). However, experimental paradigms used to elicit MMN can also achieve this in various ways.

The vast majority of studies conducted in schizophrenia employ a simple invariance oddball sequence with an identical repeating standard sound occasionally interrupted by a rare deviation (for review see Todd et al., 2012). These basic paradigms are indexing gain control in the changed response to a standard and to the deviant AEP. However, there are other ways to assess more dynamic changes. A popular alternative is the roving paradigm where each time a deviant occurs, it begins to repeat, forming the basis of a new internal model and therefore an update to the internal model predictions (Baldeweg and Hirsch, 2015; Baldeweg, 2006). The traditional invariance oddball paradigms typically reveal group difference in MMN amplitude that are due to significantly different responses to the deviant sound only, not to the repetitive standard (for discussion see Todd et al., 2012). The roving paradigm is quite different in that it exposes less suppression of the response to the repeating sound as well as a smaller amplitude response to the deviation in schizophrenia (Baldeweg et al., 2002).

At a global level the oddball and roving sequences represent different environmental scenarios. The simple oddball is ultimately a stable environment: the repetitious element remains constant across the experimental session, and the absence of a group difference in response to the repetitious element under these circumstances suggests that the expected input is suppressed to the same extent in the two groups. Over a longer time course the precision in the model should build to generate large MMN to deviations (Lieder et al., 2013, but see Sussman and Winkler, 2001 and King et al., 2014 for a examples of local and global predictable patterns), and in the oddball paradigm this is where the group difference is exposed with smaller responses in the schizophrenia groups. The roving paradigm is a comparatively volatile situation with the local environment constantly changing: volatile circumstances in which the internal model requires regular updating should engender a stronger emphasis on actual input (bottom-up information). Group differences in the rate and degree of suppression of the response to standards in roving paradigms suggest that this local timescale dynamic updating is also impaired in schizophrenia (Baldeweg et al., 2002). The period of time over which precision in a given model can accumulate in a roving paradigm is generally limited to a maximum of <20 s because the repeating strings of sound

Please cite this article as: Todd, J., et al., Time as context: The influence of hierarchical patterning on sensory inference, Schizophr. Res. (2017), http://dx.doi.org/10.1016/j.schres.2017.03.033