NOVA

University of Newcastle Research Online

nova.newcastle.edu.au

Fitzgerald, Kaitlin; Provost, Alexander & Todd, Juanita. "First-impression bias effects on mismatch negativity to auditory spatial deviants" Published in *Psychophysiology*, Vol. 55, Issue 4, no. e13013, (2018).

Available from: <http://dx.doi.org/10.1111/psyp.13013>

This is the peer reviewed version of above article, which has been published in final form at <http://dx.doi.org/10.1111/psyp.13013>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

Accessed from: <http://hdl.handle.net/1959.13/1349411>

First-impression bias effects on MMN to auditory spatial deviants.

Kaitlin FITZGERALD¹⁻², Alexander PROVOST¹⁻², Juanita TODD¹⁻³

1. School of Psychology

University of Newcastle

University Drive, Callaghan, NSW, Australia, 2308

2. Priority Research Centre for Translational Neuroscience and Mental Health Research

University of Newcastle

University Drive, Callaghan, NSW, Australia, 2308.

3. Schizophrenia Research Institute

Darlinghurst, NSW, Australia, 2010.

Corresponding author:

Kaitlin Fitzgerald

School of Psychology

University of Newcastle

University Drive, Callaghan, NSW, Australia, 2308

Email: Kaitlin.Fitzgerald@uon.edu.au

Phone: +61 2 4921 7769

Fax: +61 2 4921 6980

Abstract

Internal models of regularities in the world facilitate perception as redundant input can be predicted and neural resources conserved for that which is new or unexpected. In the auditory system this is reflected in an evoked potential component known as mismatch negativity (MMN). MMN is elicited by the violation of an established regularity to signal the inaccuracy of the current model and direct resources to the unexpected event. Prevailing accounts suggest that MMN amplitude will increase with stability in regularity, however observations of *first-impression bias* contradict stability effects. If tones rotate probabilities as a rare deviant ($p = 0.125$) and common standard ($p = 0.875$), MMN elicited to the initial deviant tone reaches maximal amplitude faster than MMN to the first standard when later encountered as deviant – a differential pattern that persists throughout rotations. Sensory inference is therefore biased by longer-term contextual information beyond local probability statistics. Using the same multi-context sequence structure, we examined whether this bias generalises to MMN elicited by spatial sound cues using monaural sounds ($n = 19$ right first deviant and $n = 22$ left first deviant) and binaural sounds ($n = 19$ right first deviant). The characteristic differential modulation of MMN to the two tones was observed in two of three groups, providing partial support for the generalisation of first-impression bias to spatially deviant sounds. We discuss possible explanations for its absence when the initial deviant was delivered monaurally to the right ear.

First-impression bias in MMN to auditory spatial deviants

The quick and accurate processing of sound is a complex task given the multitude of auditory information which is received at any one time. In an environment which is both ever-changing yet constrained by regularity (for example, thunder predictably follows lightning), optimal processing involves the ability to detect regular patterns in the environment whilst maintaining a sensitivity to when these change. According to predictive coding theory, this is achieved by comparing incoming input with a stored model of the world generated from recent regularities, such that further processing only occurs when there is a discrepancy between actual input and the model prediction. This account assumes a hierarchical organisation of the cortex where these representations, variously referred to as “internal models” (Friston, 2005) or “prediction models” (Winkler, 2007) are generated by higher brain regions from information necessarily collected over longer time periods, and via backward connections exert a “top-down” influence over the responding of lower-level sensory cortices to current input (Friston, 2005). When redundant information is accurately predicted neural resources for the processing of this information are conserved, yet if there is an over-emphasis on predicting the future state of the world, sensitivity to detect potentially important changes is impaired. Efficient processing of sound thereby involves maintaining a stored representation not only of how the world is likely to be, but also how it is likely to change over time (Behrens, Woolrich, Walton & Rushworth, 2007; Friston, 2005). Subsequently salient “prediction errors” are communicated from these lower levels bottom-up only when an existing model is violated, and trigger an update to the existing model to account for this unexpected event (Friston, 2009; Meyniel & Dehaene, 2017; Mumford, 1992; Rao & Ballard, 1999). This hierarchical dissociation of primary sensory areas in temporal cortex and higher brain regions (including secondary and frontal cortices) in detecting novelty over local versus global timescales has been supported across various methodologies including EEG, MEG, and intracranial studies (e.g. Dürschmid et al., 2016; El

Karoui et al., 2015; Wacongne et al., 2011). The terms “bottom-up” and “top-down” as they are used throughout this paper refer to the influence of each of these components of the hierarchy on cortical responding.

To effectively balance these two sources of information, both are weighted according to estimates of “confidence” or “precision” - a quantification of certainty in their reliability or accuracy (Friston, 2005; Winkler, 2007). Like most predictions, confidence will increase as supporting evidence is accumulated, therefore the weighting given to predictions should increase the longer a pattern has remained stable. Conversely, when a prediction is incorrect, a sudden drop in predictive confidence should lead to reduced reliance on top-down predictions and increased processing of current input by sensory cortices in order to maximise accuracy (Nassar, Wilson, Heasley & Gold, 2010). This process of predictive coding has been reflected in both behavioural studies (e.g. Behrens et al., 2007; Meyniel, Schlunegger & Dehaene, 2015; Nassar et al., 2010) and via observation of brain responses thought to be related to prediction generation and updating (Garrido, Kilner, Stephan & Friston, 2009).

In electrophysiological studies this is observed via a component of the auditory event-related potential known as mismatch negativity (MMN; Garrido, Kilner, Stephan et al., 2009; Lieder, Stephan, Daunizeau, Garrido & Friston, 2013; Wacongne, Changeux, & Dehaene, 2012). MMN is the quantification of a negative deflection in scalp-recorded event-related potential (ERP) amplitude at frontal electrodes which occurs to any deviation from an established regularity in sound (Näätänen, Gaillard & Mäntysalo, 1978; Sams, Paavilainen, Alho & Näätänen, 1985). It is most easily observed in oddball sound sequences when a rare

First-impression bias in MMN to auditory spatial deviants

deviant tone irregularly interrupts a series of highly probable and repetitive *standard* tones. It is thought to signal the prediction error elicited by this pattern deviation and subsequently direct neural resources to process the unexpected sound and update the existing prediction (Winkler, 2007).

As evidence for this, the amplitude of the MMN is linked to estimates of confidence or precision in predictions, with the largest MMN observed to deviations of highly repetitive sequences where predictive confidence is presumed to be highest and deviations are most surprising (Friston, 2005; Lieder, Stephan, Daunizeau, Garrido, & Friston, 2013; Näätänen, 1992; Ritter et al., 1992). Meanwhile, MMN will decrease in amplitude rapidly with repetition of this same deviant, demonstrating the flexibility of this system to update predictions to account for this previously unexpected event and re-establish the weighting given to top-down predictions (Friston, 2005; Garrido, Kilner, Kiebel, Stephan & Friston, 2007; Näätänen, 1992; Winkler, 2007; Winkler et al., 1996). In this way MMN amplitude reflects the balance between top-down and bottom-up influences on cortical responding during sensory processing and perception.

These propositions translate to the assumption that MMN will always be larger to deviations of stable sequences which are unchanged over longer time periods as repetition leads to increase evidence for the accuracy of the prediction model and subsequently higher precision estimates (Friston, 2005; Winkler, 2007). However, studies employing sequences that contain predictable patterns on multiple timescales suggest that MMN amplitude modulations on the basis of local stability can be substantially distorted by information about the more global structure of the sequence collected over longer time periods (Frost, Winkler,

Provost & Todd, 2016; Mullens et al., 2014; Mullens et al., 2016; Todd, Heathcote, Mullens, et al., 2014; Todd, Heathcote, Whitson et al., 2014; Todd, Provost & Cooper, 2011; Todd, Provost, Whitson, Cooper & Heathcote, 2013).

The “multi-timescale” sequence is a variant of an oddball paradigm in which two tones alternate in the role of predictable standard and rare deviant across blocks at different rates (Todd et al., 2011). The arrangement at sequence onset will hereafter be called “context A” and the blocks where these initial roles reverse will be referred to as “context B”. The role reversals happen at regular intervals to create sequences that have either a high local stability in sound patterns (e.g., *stable sequences* where tone roles change every 2.4 minutes to form two blocks of context A and two of context B), or less local stability (e.g., *unstable sequences* where tone roles change every 0.8 minutes to form six blocks of context A and six of context B). Overall the sequences span the same total duration, contain the same number of each sound, and the relative probability of a standard within any given block is the same. The only difference is the period of time over which a reliable pattern of tone probabilities has been stable, and therefore the period over which precision estimates can accumulate.

If precision estimates accumulated solely on the basis of local information, MMN amplitude to deviations would be expected to be larger in more stable sequences, and larger at the end than at the beginning of sequence blocks due to the longer time tone roles had been stable. However, multi-timescale studies have exposed MMN modulation patterns that cannot be reconciled with local pattern stability. If participants hear a stable sequence first, MMN only increases over time within blocks for the deviant in context B (i.e. the second deviant). MMN to the deviant in context A (i.e. the first deviant) is large and shows no significant

First-impression bias in MMN to auditory spatial deviants

change across the early and later portions of the blocks (Frost et al, 2016; Todd, Heathcote, Mullens et al., 2014). If participants hear the unstable sequence first, the same modulation pattern occurs but to a lesser extent (Frost et al., 2016). These results suggest that different precision values and therefore learning rates are associated with context A and context B, with those for the former being uniformly high (or rapidly accumulating) resulting in a consistently large MMN, and those for the latter exposing slowed learning in the new context leading to small MMN initially and an observed growth in MMN over time. This effect has been termed a *primacy bias* or *first-impression bias* as it appears to represent a resistance to re-evaluate what is first learned, which subsequently biases the way in which future sounds are processed.

Cumulative evidence continues to demonstrate the primacy bias as a reflection of this preference for higher-order, global contextual information in determining resource allocation and attention. This is supported by observations of MMN amplitude modulation being influenced when the behavioural relevance of tones is pre-established. When the two tones from the multi-timescale sequence are used with equal probability in a preceding go/no-go task, the bias does not initially occur (Mullens et al., 2014). Their equal likelihood appears to interfere with the order effect. However, with repeated exposure the bias returns in a task-dependent manner suggesting that the bias interacts with tone salience. Bias was only present when the first deviant corresponded with the “go-tone” assigned for that participant (either a 30ms or 60ms tone, counterbalanced across participants). This is assumed to reflect the experientially established prior of this tone as important for behaviour (i.e. requiring a response) and the “no-go” tone as relatively uninformative. The mechanism producing this bias was therefore thought to be linked to an emphasis on the “information value” of a tone

First-impression bias in MMN to auditory spatial deviants

for behaviour, established when the tone is first experienced either as a salient change (high information value) or a redundancy (low information value).

To deepen understanding of auditory predictive processes and their relationship to attention and implicit learning, the generality of this finding to multiple sound properties must first be established. Multi-timescale sequences have previously been utilised with two tones of different duration (Todd et al., 2011; Todd, Provost, et al., 2013) and two tones of different frequency (Todd, Heathcote, Whitson et al., 2014) and have produced consistent bias patterns in stable sequences with less consistency in unstable sequences (Todd, Provost, Whitson & Mullens, 2017). However, top-down modulatory effects have been most widely studied in the area of spatial attention and spatial manipulations are most widely used in the investigation of early attention and learning (e.g. Gilbert & Sigman, 2007; Posner, 1980; see Vecera & Rizzo, 2003 for a review). The present study was therefore designed to investigate whether these same bias patterns would be present to manipulations of spatial location, in order to best direct further studies of this phenomenon.

The study consisted of a modification of the stable and unstable sequences used by Todd, Provost et al. (2013), with sounds differentiated as localised to a left or right spatial location. Working under the assumption that the bias was driven by the relative importance of sounds for behaviour, we initially chose an extreme manipulation of location via monaural presentation of sound energy to the left or right ear only to maximise the salience of these change cues and thereby the MMN amplitude differences. We expected to see the primacy bias reflected in an effect of tone order on MMN amplitude. Specifically, evidence would accumulate differently in contexts A and B as described above. We anticipated this effect

First-impression bias in MMN to auditory spatial deviants

would be more apparent in the stable than relatively more unstable sequences. It is here premised that these patterns were absent from the first set of data collected using this manipulation. Subsequently, a full counterbalancing of this monaural condition and an alternative manipulation of spatial location using a more subtle inter-aural level difference led to a complex pattern of results to be discussed.

Method

Participants

Participants were 62 healthy adults (43 female; age 18-28, $M = 21.10$, $SD = 2.36$), recruited from undergraduate psychology students at the University of Newcastle and community volunteers. Individuals were excluded from participation if they were under the age of 18 or over 35, or had a history of conditions known to affect the MMN including neurological disorder, head trauma or surgery, regular recreational drug use, heavy alcohol use, current diagnosis or treatment of mental illness, hearing impairment, or a first-degree relative with schizophrenia. Reimbursement was provided in the form of course credit for students and gift vouchers for community volunteers. Written informed consent was collected from all participants prior to participation to complete the protocol as approved by the Human Research Ethics committee of the University of Newcastle.

Stimuli

Participants were assigned to one of three conditions that determined the sound stimuli which were delivered over binaural headphones. The first 20 participants were recruited to a “monaural cue” condition in which the spatial location of each tone was

First-impression bias in MMN to auditory spatial deviants

manipulated by delivering sound energy only to one ear. These stimuli were 1000Hz pure tones of 60ms duration (created using 5ms rise/fall time and a 50ms pedestal), presented at 75dB SPL. This created a perceived location of approximately 90° from the centre of the head to the side of space to which the sound was delivered, whilst providing a maximal segregation of the tones in space. Other sound parameters were chosen to maximise consistency with previous studies of the bias, which have used pure tones of 1000-1500Hz frequency, 30-100ms duration, and 75dB loudness (Frost et al., 2016; Mullens et al., 2014; Todd et al., 2011; Todd, Provost et al., 2013; Todd, Heathcote, Whitson et al., 2014).

The two additional conditions were subsequently administered to further explore the apparent absence of the bias in the first group. A second monaural cue group was collected using these same stimuli to counterbalance and examine any effect of the relative probability of left and right-localised tones at sequence onset. The third group of participants were recruited to a “binaural cue” condition, where tones were presented to both ears with an interaural level difference in the ratio of 70:86dB SPL between the ears to explore whether the absence of bias was related to monaural stimulation. This intensity difference was modelled on that used by Paavilainen, Karlsson, Reinikainen and Näätänen (1989) to similarly give the perception of the tone being laterally localised approximately 90° from the centre of the head whilst providing auditory input to both ears to resemble more natural listening conditions. These stimuli were 60ms (5ms rise/fall time and 50ms pedestal) complex tones comprised of a 600Hz and 3000Hz component, frequencies considered optimal for study of interaural phase and level differences respectively (Paavilainen et al., 1989) and which are consistent with previous research on MMN to spatial location changes using binaural headphones (Matthews, Todd, Budd, Cooper & Michie, 2007).

Sequences

Sounds were organised in sequences identical to those described by Todd, Provost et al. (2013) and depicted in Figure 1.

Figure 1. Representation of sound sequence arrangement, where shaded blocks represent “context A” and spotted blocks “context B” for each group. Blue markings indicate rate of block change in slow and fast sequences, whilst orange markings demonstrate division of blocks into halves for analysis.

Variation in the lateral localisation of sounds occurred across two block types, within which one lateral location was highly repetitive ($p = 0.875$), and the other was rare and unpredictable ($p = 0.125$) and in the subsequent block these roles reversed between locations with the rare location becoming common, and the common location becoming rare. This role-switching occurred within stable and unstable sequences, characterised by block lengths of 2.4 and 0.8 minutes respectively. Overall both sequences contained 1920 tones: 960 left localised and 960 right localised. Stable sequences contained four blocks of 480 sounds and unstable sequences contained twelve blocks of 160 sounds. Sounds were presented at a consistent 300ms stimulus onset asynchrony.

All groups heard two versions of the stable and unstable sequences: one with a left-localised tone as the first deviant and the other with the right-localised tone as the first deviant. These arrangements are depicted in Figure 1 as orders 1 and 2. For this reason the different blocks will not be referred to as context A and B in methods and results as these are different in the two orders. Instead the term *first deviant* will be used to refer to deviants in blocks consistent with how the sequence commenced, and *second deviant* will be used for sounds that are deviant in the blocks consistent with the subsequent role-alternation.

In the monaural cue groups, 20 participants (11 female, age 18-24, $M = 21.20$, $SD = 1.64$) heard sequences in which the right-localised tone was the very first deviant in order 1 (hereafter referred to as MR) and the other 22 participants (17 female, age 18-28, $M = 21.41$, $SD = 2.91$) heard sequences in which the left-localised tone was the very first deviant in order 1 (hereafter referred to as ML). The binaural cue group ($n = 20$, 15 female, age 18-28, $M = 20.65$, $SD = 2.37$) heard sequences in which the right tone was the very first deviant in order 1 (hereafter referred to as BR). In all groups the first presentation of each sequence (stable and unstable) was followed by a second presentation of each sequence after a short silent break (see Figure 1) with the alternate sound location as first deviant. The slow sequence consistently preceded the fast sequence in both orders for all participants in all groups.

Procedure

Participants completed a screening interview for exclusion criteria prior to participation. The absence of hearing impairment was confirmed using a pure tone audiometer to assess hearing thresholds from 500Hz to 4000Hz (where minimum detection at

First-impression bias in MMN to auditory spatial deviants

$\leq 20\text{dB SPL}$ is considered adequate). Suitable participants were fitted with a 64-channel Neuroscan QuikCap[®] with Ag/AgCl electrodes and binaural Sennheiser headphones through which participants received sound sequences.

Participants were instructed to ignore the sounds and focus on a muted self-chosen movie with subtitles. There was a 40-second silent break between presentation of the stable and unstable sequences in both orders, and a two-minute break between orders 1 and 2 during which participants were permitted to move and readjust.

Continuous electroencephalogram (EEG) was taken using a SynAmps2 Neuroscience[®] system with a 1000Hz sampling rate (high-pass 0.1Hz, low-pass 70Hz, notch filter 50Hz and a fixed gain of 2010). Data was recorded from 64 scalp electrodes arranged per the International 10 $\pm$ 10 system with Modified Combinatorial Nomenclature (American Clinical Neurophysiology Society, 2006) including the nose as reference, plus two additional electrodes placed on the bilateral mastoids. Data was also taken from four electro-oculogram electrodes placed 1cm from the outer canthi of both eyes, and directly above and below the left eye in order to monitor blinks, saccades and unnecessary eye movements. Impedances were reduced to below 5k Ω prior to recording.

Data Analysis

The EEG was first visually examined using NeuroScan Edit[®] software for any sections of major artefact, which were manually removed. Eyeblink correction was applied using the same software, which follows a method of artefact averaging and regression

First-impression bias in MMN to auditory spatial deviants

(Semlitsch, Anderer, Schuster, & Presslich, 1986). To ensure adequacy the average artefact was drawn from a minimum of 30 sweeps with $< 5\%$ variance between waveforms contributing to the average.

For analysis each block of every sequence was broken into halves in accordance with prior methods (Frost et al., 2016; Mullens et al., 2014; Mullens et al., 2016; Todd, Heathcote, Mullens et al., 2014) – the first half immediately after tones changed roles, and the second half after the new tone roles had time to become established. Halves are illustrated in Figure 1 by the markings ‘1’ and ‘2’ beneath sequence blocks. Separate averages were therefore calculated for each half, based on trial numbers with a maximum of 60 deviants total for each tone type and condition. They were necessarily longer in duration in the stable (≈ 1.2 minutes, sweeps drawn from two lots of 30 deviants) than unstable (≈ 0.4 minutes, sweeps drawn from six lots of 10 deviants) sequence due to overall block lengths.

Data was epoched from 50ms prior to stimulus onset to 300ms post-stimulus and baseline corrected over the entire interval. Epochs containing amplitudes more extreme than $\pm 70\mu\text{V}$ were discarded. Separate ERPs were produced for both tones as standards and deviants, in the first and second half of blocks for all conditions. These averages were baseline corrected to the prestimulus period and a 30 Hz zerophase low-pass filter was applied with 12dB slope (as recommended for examining N100 (Kujala, Tervaniemi, & Schröger, 2007)). The first four standards at the start of a block and the first standard immediately following the presentation of a deviant were excluded from the averages, to allow time for a new memory trace for this sound as a standard to be formed (Sams, Hari, Rif, & Knuutila, 1993). Difference waveforms were then produced through the subtraction of

First-impression bias in MMN to auditory spatial deviants

the ERP to a standard sound, from that to the same sound heard as deviant, in order to reduce the influence of exogenous effects on the difference waveform used to compute MMN. A lower 20Hz zerophase low-pass filter with 12dB slope was applied to the difference waveforms as recommended for examination of MMN (Kujala et al., 2007). This resulted in 16 difference waves per participant - one for first- and second-half of blocks, for both the left and right tone as deviant, in the stable and unstable sequence, across orders 1 and 2. One participant from the MR group, and one participant from the BR group were excluded from analysis due to unsatisfactory data quality, leaving 19 data sets for analysis in both of these samples and a full sample of 22 data sets in the ML group. One grand average difference waveform was then computed for each condition in each group. The mean number of sweeps contributing to the average ranged from 49-54 of a possible 60 sweeps for deviants, and 281-324 of 410 possible sweeps for standards.

Previous studies on the primacy bias adopted frontal measures of the MMN re-referenced to the average of the left and right mastoid site in line with recommendations for maximising signal-to-noise ratio (Kujala et al., 2007). However, given the propensity for spatially deviant tones to produce more lateralised responses (Deouell, Heller, Malach, D'Esposito & Knight, 2007), we first assessed the topography of the resultant difference waveforms to determine the optimal common measures to employ to enable a comparison across conditions and groups. Justification for the choice of site and reference for MMN quantification is therefore reported below within the Results section of the manuscript, and led to the use of right-mastoid re-referenced data from Fz for analysis.

MMN was quantified as the mean amplitude 80-120ms for each participant in each condition. This latency range was selected following visual inspection of the difference waveforms for each participant as a common window that consistently captured the most negative point in the 70-170ms range.

To assess whether order-driven effects were present on MMN amplitude, omnibus analyses were conducted using mean amplitudes of right-mastoid re-referenced data at Fz in a mixed model repeated measures ANOVA with group (ML, MR, BR) as the between-subjects factor, and order (first sequence pair, second sequence pair), sequence (stable, unstable), tone (left, right) and half (first, second) as within-subjects factors. We expected that at minimum primacy effects would lead to a significant effect of order, and because order differs between groups as a function of tone we would see a group by order by tone interaction. We additionally hypothesised that half would further modify this where the differences should be maximal during the early period of blocks. A-priori planned paired *t*-tests were used to confirm significant changes in MMN amplitude over block halves. Assumptions of equal variance were met in all analyses so non-corrected statistics will be reported.

Results

An omnibus average comprising all epochs for responses to deviants on the left and right for each condition in nose-referenced data was used to examine differences in scalp topography associated with the spatial location of the deviant. These omnibus averages are presented in Figure 2A along with corresponding scalp topography maps in Figure 2B for the peak period of the difference waveform corresponding to MMN (~ 110ms). From Figure 2A it is clear that the MMN is larger at Fz/F4 for deviants to the left, and at Fz/F3 for deviants to

the right. This was confirmed in a significant tone by site interaction ($F(1,57) = 43.97, p < .001, \eta^2 = 0.61$) in a mixed model ANOVA with between subjects factor of group (MR, ML, BR) and within subjects factors of order (first sequence pair, second sequence pair), sequence (stable, unstable), tone (left, right), half (first, second) and site (F3, Fz, F4). For the purpose of choosing a common site for analysis it is relevant to note that the means were equivalent between Fz and F4 for the left deviant but were smaller at F3 (Fz = $-1.96 \mu\text{V}$, F4 = $-1.97 \mu\text{V}$ and F3 = $-1.48 \mu\text{V}$), and equivalent between Fz and F3 for the right deviant but were smaller at F4 (Fz = $-1.83 \mu\text{V}$, F3 = $-1.82 \mu\text{V}$ and F4 = $-1.38 \mu\text{V}$). For parsimony, Fz was selected as the frontal location for analysis of MMN modulation by experimental factors.

In the nose-referenced data presented in Figure 2A there is a clear tendency for the polarity inversion of MMN at the right-mastoid to be larger than that on the left for the left-location deviant. This tendency was confirmed in a main effect of site (left mastoid mean = $0.46 \mu\text{V}$, right mastoid mean = $1.14 \mu\text{V}$, $F(1,57) = 38.98, p < .001, \eta^2 = 0.41$) in a mixed model repeated measures ANOVA examining the polarity inversion at left and right mastoid over order, sequence, half and site with group as a between subjects variable for a left tone. In contrast, there is no visible difference between right and left mastoid in the size of the polarity inversion of the MMN for the right tone as deviant. This was confirmed in the absence of a main effect of site (left mastoid mean = $0.70 \mu\text{V}$, right mastoid mean = 0.65 , $F(1,57) = 0.16, p = .689, \eta^2 = .003$) in a mixed model repeated measures ANOVA examining the polarity inversion at left and right mastoid over order, sequence, half and site with group as the between subjects variable for the right tone.

Figure 2. A. Grand average MMN waveforms for nose-referenced data generated from all conditions with a left tone as deviant (red/top panel) and a right tone as deviant (blue/bottom panel) in each of the three groups; B. Topography maps of peak (~ 110 ms) grand average MMN response in nose-referenced data for all left deviant (top panel) and right deviant (bottom panel) conditions for each group.

As noted under methods, in prior studies of the primacy bias MMN amplitude was analysed using data re-referenced to the average of the left and right mastoid – a conventional technique recommended to maximise signal-to-noise ratio in estimating a “total mismatch” response. In light of the pattern of difference in polarity inversion in the present study we elected to re-reference the Fz electrode to the right-mastoid for all conditions to test primacy

hypotheses for compatibility with prior studies. We additionally report the same statistics for common-average referenced data for a comparison with a less bespoke referencing choice.

First-Impression Analysis

The omnibus mixed model ANOVA conducted on right-mastoid referenced data produced the expected significant group by order by tone by half interaction ($F(2,57) = 6.91$, $p < .005$, $\eta^2 = 0.20$; common average $F(2,57) = 5.14$, $p < .01$, $\eta^2 = 0.15$). The difference waveforms from which MMN amplitude was obtained are presented in Figure 3. In the MR group MMN amplitude displayed modulations expected of general stability where MMN increased over block half ($F(1,18) = 24.26$, $p < .001$, $\eta^2 = 0.57$; common average $F(1,18) = 17.56$, $p < .001$, $\eta^2 = 0.49$), was larger in stable than unstable sequences ($F(1,18) = 35.08$, $p < .001$, $\eta^2 = 0.66$; common average $F(1,18) = 20.67$, $p < .001$, $\eta^2 = 0.54$), and increased more over halves in the stable than unstable blocks for both tones (a sequence by half interaction, $F(1,18) = 7.55$, $p < .05$, $\eta^2 = 0.30$; common average $F(1,18) = 7.25$, $p < .05$, $\eta^2 = 0.29$).

The patterns of MMN amplitude modulation in the ML and BR groups were highly consistent with one another but differed from those in the MR group. In the ML and BR groups MMN amplitude was significantly modulated by order (larger in order 1 overall, ML: $F(1,21) = 6.21$, $p < .05$, $\eta^2 = 0.23$, BR: $F(1,18) = 8.41$, $p < .01$, $\eta^2 = 0.32$; common average ML: $F(1,21) = 5.97$, $p < .05$, $\eta^2 = 0.22$, BR: non-significant, $p = .142$, by sequence (larger in stable sequences overall; ML: $F(1,21) = 54.01$, $p < .001$, $\eta^2 = 0.72$, BR: $F(1,18) = 31.02$, $p < .001$, $\eta^2 = 0.63$; common average ML: $F(1,21) = 28.74$, $p < .001$, $\eta^2 = 0.58$, BR: $F(1,18) = 35.00$, $p < .001$, $\eta^2 = 0.66$), by half (larger second half overall, ML: $F(1,21) = 33.74$, $p < .001$, $\eta^2 = 0.62$, BR: $F(1,18) = 39.24$, $p < .001$, $\eta^2 = 0.64$; common average ML: $F(1,21) =$

First-impression bias in MMN to auditory spatial deviants

41.84, $p < .001$, $\eta^2 = 0.67$, BR: $F(1,18) = 21.15$, $p < .001$, $\eta^2 = 0.54$), and also by tone (larger for a left tone overall, ML: $F(1,21) = 6.35$, $p < .05$, $\eta^2 = 0.23$, BR: $F(1,18) = 17.90$, $p < .001$, $\eta^2 = 0.50$; common average ML: non-significant, $p = .727$, BR: $F(1,18) = 9.02$, $p < .01$, $\eta^2 = 0.33$). However, these main effects were modified by the expected order by tone by half interaction in both groups (ML: $F(1,21) = 11.46$, $p < .005$, $\eta^2 = 0.35$, BR: $F(1,18) = 4.67$, $p < .05$, $\eta^2 = 0.21$; common average ML: $F(1,21) = 5.20$, $p < .05$, $\eta^2 = 0.20$, BR: $F(1,18) = 4.60$, $p < .05$, $\eta^2 = 0.20$). The sources of this interaction are evident in the results of paired t -tests on mean amplitude present in Table 1. These were conducted to test the a-priori hypothesis that precision would accumulate more slowly for the second deviant in each sequence, resulting in a significant change over half for second-deviants only. This pattern was consistently true in stable sequences for the ML and BR groups and less consistently in unstable sequences, but was absent for the MR group.

The difference in MMN amplitude across block halves as a function of order is visible in the plot of difference waveforms in Figure 3. Figure 3A contains the MMNs produced to left- and right-location deviants in stable sequences. The tendency for MMN amplitude to be comparatively “suppressed” in the earlier half of blocks is very evident for second deviants in the ML (right for order 1 and left for order 2) and BR (left for order 1 and right for order 2) groups. In contrast the MMNs within stable sequences for the MR group are consistently smaller in the first half of blocks for all deviant types. Order-based differences are less evident in the unstable sequence MMNs depicted in Figure 3B.

First-impression bias in MMN to auditory spatial deviants

Sequence	First Deviant/First Block			Second Deviant/Second Block		
	Early	Late		Early	Late	
<u>MR Group</u>						
Order 1 Stable	-1.92	-3.73	$t = 4.01 (p = .001)^{***}$	-2.02	-4.16	$t = 4.18 (p = .001)^{***}$
Order 2 Stable	-2.29	-3.86	$t = 2.75 (p = .013)^*$	-1.97	-3.18	$t = 2.07 (p = .054)$
Order 1 Unstable	-2.08	-2.20	$t = 0.24 (p = .812)$	-1.60	-2.62	$t = 1.32 (p = .203)$
Order 2 Unstable	-1.79	-2.61	$t = 2.31 (p = .033)^*$	-1.23	-2.22	$t = 1.87 (p = .078)$
<u>ML Group</u>						
Order 1 Stable	-4.48	-4.91	$t = 0.84 (p = .412)$	-2.77	-4.69	$t = 3.58 (p = .002)^{***}$
Order 2 Stable	-3.16	-3.16	$t = 0.001 (p = .999)$	-2.68	-4.59	$t = 5.07 (p < .001)^{***}$
Order 1 Unstable	-2.12	-4.17	$t = 4.22 (p < .001)^{***}$	-0.85	-2.63	$t = 3.58 (p = .002)^{***}$
Order 2 Unstable	-1.85	-2.60	$t = 1.64 (p = .116)$	-1.76	-3.15	$t = 3.10 (p = .005)^{***}$
<u>BR Group</u>						
Order 1 Stable	-3.19	-4.01	$t = 1.65 (p = .117)$	-2.56	-4.68	$t = 3.91 (p = .001)^{***}$
Order 2 Stable	-3.78	-4.63	$t = 1.52 (p = .146)$	-1.22	-2.85	$t = 2.93 (p = .009)^{**}$
Order 1 Unstable	-2.53	-2.52	$t = 0.02 (p = .985)$	-1.97	-3.45	$t = 3.40 (p = .003)^{***}$
Order 2 Unstable	-1.47	-3.15	$t = 2.92 (p = .009)^{**}$	-0.21	-2.61	$t = 4.14 (p = .001)^{***}$

Note: $*$ = $p \leq .05$, $**$ = $p \leq .01$, $***$ = $p \leq .001$

Table 1. Mean amplitude 80-120ms (μ V) and results of paired t -tests for the early (first half) and late (second half) portions of blocks for first and second deviant in stable and unstable sequences for each group. Data was re-referenced to right mastoid and taken from Fz. First deviant was a left tone in the ML group and a right tone in the MR and BR groups; second deviant was a right tone in the ML group and a left tone in the MR and BR groups.

First-impression bias in MMN to auditory spatial deviants

Figure 3. Grand average MMN waveforms for right-mastoid re-referenced data at Fz in the first- (broken line) and second-half (solid line) of blocks for the first and second deviant in the stable (A) and unstable (B) sequence for each group. Waveforms for left-localised tones are presented in red, and right-localised tones in blue.

Discussion

The present study was designed to test the hypothesis that first-impression bias is a general feature of auditory processing and would therefore be present to sequences containing two tones differentiated by spatial location. We expected to see a differential modulation of MMN to first- and second-deviant tones over time where MMN is significantly smaller in the first-half of blocks for the second deviant only. Elsewhere this has been interpreted to reflect a comparatively more rapid accumulation of confidence in the first-deviant context than the second-deviant context (Mullens et al., 2016; Todd, Heathcote, Mullens et al., 2014). These patterns were absent from the first group tested, yet were confirmed in two subsequent groups that used alternate methods or orders of spatial manipulation. In the following discussion we expand upon the potential significance of the data patterns and possible implications for our understanding of the bias.

Group Differences

In the MR condition, the only group that failed to show first-impression bias, the very first deviant type (i.e. order 1) was characterised by monaural sound delivered to the right ear and the first standard sound was presented monaurally to the left ear. In general, MMN amplitude incremented from first to second block halves consistent with a gradual build in model confidence for all sequence blocks. In essence, this group displayed MMN modulation patterns that would be expected based on a reliance on local stability patterns.

Contrary to the patterns observed for the MR group, there were clear patterns of first-impression bias in both ML and BR groups. MMN amplitude was modulated differently for the two sounds dependent on whether they were the first or second deviant. This is evident in

First-impression bias in MMN to auditory spatial deviants

the consistent finding of significantly smaller MMN to second-deviant tones in earlier than later portions of sequence blocks, whilst there was generally no significant change in MMN amplitude over block half for the first deviant. This is most evident in the stable sequences and is consistent with the expectation of comparatively slower learning rates for the second-deviant context. The inversion of these patterns within-subjects (between order 1 and 2) replicates prior studies and is consistent with these order-driven modulations of MMN amplitude being tied to learning about initial tone roles (Todd, Provost, et al., 2013; Todd, Provost et al., 2017). Evidence of first-impression bias was therefore present in two of the three groups tested – a pattern of results that complicates the conclusion that it governs sound processing in general.

The absence of the expected patterns in the MR group only is perplexing. These participants heard the exact same sequences as the ML group merely in the opposite order. If monaural sound delivery were too extreme or artificial to induce the first-impression bias, the bias should have been absent from both monaural conditions. Instead the ML and BR groups showed largely consistent patterns of modulation despite differing in both tone order (i.e. left first deviant in the first sequence pair for ML; right first deviant in the first sequence pair for BR) and stimulus delivery in terms of intensity difference between the ears (75dB for ML and 16dB for BR).

One possible explanation for group differences is participant heterogeneity. Differences in lateralisation for spatial tasks based on characteristics such as handedness and gender have long been proposed (see Vogel, Bowers, & Vogel, 2003 for a review). The MR group was indeed the most demographically distinct in terms of the sex distribution, having a

higher proportion of male participants than the ML and BR groups. However, as noted in supplementary analyses, there was no significant differences between sexes nor was sex a significant modifier of the condition effects on MMN amplitude in this study. It is a limitation of the present study that participant handedness was not recorded at the time of the experiment. However, supplementary analyses revealed that the main effects and interactions seen within the groups remained the same in a subset of data from participants later confirmed as right-handed (16 MR, 11 ML, 12 BR). If bias were reduced in left-handers, this would be unlikely to explain these differences as the proportion of confirmed right-handers was at least as high in the MR group as in the other two. We thus have no strong evidence or reason to suppose that demographic heterogeneity could explain the different pattern of results in the MR group. It is possible that it is an artefact of some unknown confound that was present only for this group, a possibility that will be ruled out only with further replication. We will briefly explore the alternative possibility that the findings relate to how differences in sound exposure might interact with the factors causing the first-impression bias, which we foreshadow with the acknowledgement that this was an unexpected effect therefore any speculation is necessarily post-hoc and further investigation is needed.

First-Impression Bias

The explanation offered for first-impression bias assumes that an emphasis on top-down influence on cortical responsiveness results in altered learning rates for the two tones (Friston, 2005; Todd, Heathcote, et al., 2014; Todd, Provost et al., 2013; Todd, Provost et al., 2017). We propose that the first impression is embodied in a binary categorisation of the information value of the two tones at sequence onset -where one sound is common, predictable and redundant, and the other is rare, unpredictable and comparatively more

important to new learning (Mullens et al., 2014; Todd et al., 2011; Todd, Provost, et al., 2014). There is also an assignment of high confidence in the prediction that this environment is stable and unlikely to change (Mullens et al., 2016; Todd, Petherbridge, Speirs, Provost & Paton, 2017). This leads to a resistance to re-evaluate this initial categorisation and accept a formerly redundant sound as rare and potentially important, resulting in a comparatively slower learning rate (observed as a more gradual increase in MMN amplitude) for the second deviant in context B.

This account is consistent with observations that the differential modulation of MMN amplitude for context A and B is most reliably seen in stable sequences – a result that is replicated in the present study where effects are less clear in unstable sequences (see Table 1). There are two related explanations for this. Firstly, the strength of the first-impression bias may be a function of the initial period of stability, which is of course longer in the stable than unstable sequence. Secondly, variable results in unstable sequences have elsewhere been discussed in terms of the impact of volatility on the weighting of top-down effects (Frost et al., 2016; Todd, et al., 2014; Todd, Petherbridge, Speirs, Provost & Paton, 2017). As outlined in the Introduction, the balance between top-down and bottom-up influences on neural response is impacted by model confidence, which is shaped by the volatility of the environment. Top-down effects will be more pronounced in stable, predictable environments whilst increased bottom-up sampling from the environment will occur in more unstable environments which are less easily predicted (Friston, 2005). In essence, the diminished evidence of first-impression bias patterns in more volatile sequences is consistent with the idea that the brain learns that change is likely and must therefore to some extent rely on bottom-up sampling, leading to a decreased emphasis on top-down modulation that is reflected in reduced or less consistent differential modulation of MMN amplitude.

This explanation of the bias pattern is in contrast to a primarily bottom-up account of the differential growth rate in MMN amplitude. A similar order effect observed by an independent group was attributed to a form of long-term stimulus-specific adaptation (Costa-Faidella, Grimm, Slabu, Díaz-Santaella, & Escera, 2011). In this study, participants demonstrated the same pattern of MMN modulation differences in context A and B using much shorter blocks (even shorter than the unstable blocks in the present study), and only one block of A and of B. The slower increase in MMN amplitude over time for context B was attributed to a high level of repetition of the first standard leading to strong adaptation in the responding neurons when the tone was the standard in context A, that took time to overcome in context B.

The design employed in the present study does not enable us to determine whether a top-down or bottom-up explanation is more likely. However, other designs in previous studies have provided strong evidence for a top-down explanation. For example, bottom-up explanations alone cannot account for the task-dependence of the first-impression bias effects in Mullens et al. (2014) reviewed in the introduction. Participants in this study heard the exact same sound sequences and yet displayed different first-impression bias patterns (i.e., MMN modulation patterns) based on which of the two sounds they had been asked to respond to in the preceding go, no-go task. Bottom-up explanations also fail to adequately account for why the differential modulation of MMN amplitude in context A and B of stable blocks is reversed if as many as twelve unstable blocks precede the four stable blocks (Todd, Petherbridge et al., 2017; see also Frost et al., 2016 and Mullens et al., 2016) The authors instead propose that learning over a much longer time course could enable the auditory

system to extract an additional model of regularity in alternation rate, leading to a drop in confidence and smaller MMN when tone roles fail to change after 0.8 minutes in context A of the stable sequence (also see Mullens et al., 2016 for further discussion). It is also unlikely that the results of the present study could be explained via a bottom-up account, as we would expect an equivalent degree of neural adaptation in ML and MR groups leading to similar patterns of MMN modulation. The question becomes whether a top-down explanation can be offered for the failure of MR participants to show the first-impression bias.

The only methodological property shared between the two groups showing the bias that is not present for the MR group was that both involved some input to the left ear for the deviant, and to the right ear for the standard at sequence onset. Put differently, it seems the only stimulation difference available to explain the inconsistent results in the MR group may be the relative activation of afferent pathways from each ear by each tone at the start of the sequence. As detailed in the method section, where previous studies of first-impression bias using frequency and duration deviants have revealed maximal effects over right frontal scalp sites, data in this spatial paradigm generally produced contralateral-dominant MMNs reflecting the crossing of afferent pathways providing direct sensory input to the contralateral hemisphere (Kolb & Whishaw, 2009; as shown in Figure 2). If the bias is right-hemisphere dominant when tones are not spatially differentiated, then perhaps the MR group fail to show bias because it somehow depends on direct engagement of the right-hemisphere by the first deviant via stimulation of the left ear (as occurs for the ML and BR groups). This is consistent with another striking difference between groups when data is taken from Fz - as shown in Figure 3, there is little variability in MMN amplitude to left- and right-location deviants in the MR group, whilst MMN is clearly larger to left- than right-localised tones for the ML and BR groups (shown in Figure 3 and as per main effect of tone in results). Thus,

First-impression bias in MMN to auditory spatial deviants

more variability in MMN amplitude is observed in the present study when the initial deviant involves direct input to the right hemisphere.

There is growing evidence to suggest that asymmetry favouring the right hemisphere may underlie the higher order processes responsible for MMN amplitude modulation. The MMN elicited to frequency, duration and intensity deviations has repeatedly been shown to be maximal in amplitude over right-lateralised sites, which is interpreted as evidence for stronger MMN generation processes in the right hemisphere than the left (Paavilainen, Alho, Reinikainen, Sams & Näätänen, 1991). Dynamic causal modelling (DCM) of MMN to frequency deviants has subsequently favoured a network consisting of bidirectional projections from bilateral sources in A1 and the superior temporal gyrus, with asymmetrical engagement of the right inferior frontal gyrus (Garrido, Kilner, Kiebel & Friston, 2009). Meanwhile, when the deviation is lateralised, left-sided deviants selectively activate projections to right-frontal areas, whilst right-sided deviants are associated with activation in both hemispheres (Deouell, Bentin & Giard, 1998; Dietz, Friston, Mattingley, Roepstorff, & Garrido, 2014; Giard, Perrin, Pernier & Bouchet, 1990). This is consistent with the suggestion that the right hemisphere plays a critical role in processing spatial information from both sides of the sensorium whilst the left is contralaterally limited (Mesulam, 1999; Zatorre & Penhune, 2001) and might further be interpreted as evidence that right-hemisphere engagement is necessary in higher-order processing of deviant events. This would be consistent with recent claims that the process of generating confidence estimates based on longer-term statistics is also right-hemisphere localised (Meyniel & Dehaene, 2017).

However unlikely, it remains possible that the comparatively reduced engagement of the right hemisphere by the first salient change encountered by the MR group could have implications for the bias because of how attentional resources are allocated. Concurrent work in our laboratory has indicated that the same pattern of results seen in the MR group (i.e., basic stability, no bias) is produced if the multi-timescale sequences are presented to participants while they are engaged in an attention-demanding task (Frost, McDonnell, Provost & Todd, 2015). Therefore, a potential top-down explanation for the apparent absence of first-impression bias in the MR group is the lack of an attention-based augmentation of precision assigned to the first deviant. This could be either due to reduced engagement of the right-hemisphere by deviations that form part of the first impression, or to a sampling coincidence that resulted in the MR group finding the DVD movie task more demanding. The former would be consistent with the idea that both confidence (and therefore precision) estimates and attention allocation mechanisms are also right-hemisphere dominant.

Finally, we acknowledge the necessary limitation of low epoch numbers in the averaging process that contributes to our first and second half data. There is an obvious tension between capturing learning over time in ERP studies and the need for sufficient signal to noise. Nonetheless, the confidence to interpret these effects as meaningful stems from the high consistency in the pattern of results across the ML and BR groups in this study, together with this pattern replicating many previous studies.

Conclusions

In summary, the present study provides evidence that first-impression bias influences over MMN amplitude can be observed in a multi-timescale sequence comprising tones

First-impression bias in MMN to auditory spatial deviants

differentiated by spatial location. Given that spatial manipulations are considered the optimal paradigm for investigating automatic learning and attention allocation in the brain (Driver, 2001; Posner, 1980), this finding encourages further exploration of how bias mechanisms depend on or interact with attention. The results also point to possible hemispheric asymmetries in producing the bias that warrant further investigation and may be related to similar asymmetries observed for attention and spatial processing more generally.

Although the acoustic stimulation used in the multi-timescale sequence can be considered a very artificial environment, it can provide insight into the principles of learning that govern responsiveness to sound in very early components of the auditory-evoked potential. The series of studies have provided cumulative evidence consistent with the notion that sensory inference in the auditory system is influenced by regularities present on multiple timescales, so responsiveness to a given sound cannot be predicted based on local transition statistics alone. While compatible observations date back to much earlier studies (see Sussman, Chen, Sussman-Fort & Dinces, 2014 and Winkler 2007 for reviews), the more recent framework of predictive coding provides a more coherent explanation in terms of how local predictions and model updating will be constrained by any concurrent higher-order predictions and the precision or confidence with which they are held. Given that MMN amplitude is thought to play a role in determining where our limited neural resources are directed, consideration of the first-impression bias and the conditions under which it is and is not observed can offer a pathway to a deeper understanding of influences over sensory inference and the filtering of sound relevance. Ultimately this deeper appreciation of MMN as a signal within a system balancing concurrent demands on its resources has applications in investigations in clinical groups where this component is known to be reduced (Schall, 2016).

References

- American Clinical Neurophysiology Society (2006). Guideline 5: Guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, 23(2), 107-110.
- Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, 10(9), 1214-1221. <https://doi.org/10.1038/nn1954>.
- Costa-Faidella, J., Grimm, S., Slabu, L., Díaz-Santaella, F., & Escera, C. (2011). Multiple time scales of adaptation in the auditory system as revealed by human evoked potentials. *Psychophysiology*, 48(6), 774-783. <https://doi.org/10.1111/j.1469-8986.2010.01144.x>.
- Deouell, L. Y., Bentin, S., & Giard, M. H. (1998). Mismatch negativity in dichotic listening: Evidence for interhemispheric differences and multiple generators. *Psychophysiology*, 35(4), 355-365. <http://doi.org/10.1111/1469-8986.3540355>.
- Deouell, L. Y., Heller, A. S., Malach, R., D'Esposito, M., & Knight, R. T. (2007). Cerebral responses to change in spatial location of unattended sounds. *Neuron*, 55(6), 985-996. <https://doi.org/10.1016/j.neuron.2007.08.019>.
- Dietz, M. J., Friston, K. J., Mattingley, J. B., Roepstorff, A., & Garrido, M. I. (2014). Effective connectivity reveals right-hemisphere dominance in audiospatial perception: Implications for models of spatial neglect. *The Journal of Neuroscience*, 34(14), 5003-5011. <https://doi.org/10.1523/jneurosci.3765-13.2014>.
- Driver, J. (2001). A selective review of selective attention research from the past century. *British Journal of Psychology*, 92(1), 53-78. <http://dx.doi.org/10.1348/000712601162103>.
- Dürschmid, S., Edwards, E., Reichert, C., Dewar, C., Hinrichs, H., Heinze, H. J., ... Knight, R. T. (2016). Hierarchy of prediction errors for auditory events in human temporal and

First-impression bias in MMN to auditory spatial deviants

frontal cortex. *Proceedings of the National Academy of Sciences - Neuroscience*, 113(24), 6755-6760. <http://dx.doi.org/10.1073/pnas.1525030113>.

El Karoui, I., King, J., Sitt, J., Meyniel, F., Van Gaal, S., Hasboun, D., ... Naacache, L. (2014). Event-related potential, time-frequency, and functional connectivity facets of local and global auditory novelty processing: An intracranial study in humans. *Cerebral Cortex*, 25(11), 4203-4212. <http://dx.doi.org/10.1093/cercor/bhu143>.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1456), 815-836. <http://dx.doi.org/10.1098/rstb.2005.1622>.

Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences*, 13(7), 293-301. <http://dx.doi.org/10.1016/j.tics.2009.04.005>.

Frost, J. D., McDonnell, K., Provost, A. & Todd, J. (2015). Does sequence foreknowledge or concurrent task affect first-impression bias in mismatch negativity? *Frontiers in Human Neuroscience Conference Abstract: ASP2015 - 25th Annual Conference of the Australasian Society for Psychophysiology*. <https://doi.org/10.3389/conf.fnhum.2015.219.00044>.

Frost, J. D., Winkler, I., Provost, A., & Todd, J. (2016). Surprising sequential effects on MMN. *Biological Psychology*, 116, 47-56. <http://dx.doi.org/10.1016/j.biopsycho.2015.10.005>.

Garrido, M. I., Kilner, J. M., Kiebel, S. J., & Friston, K. J. (2009). Dynamic causal modeling of the response to frequency deviants. *Journal of Neurophysiology*, 101(5), 2620-2631. <http://dx.doi.org/10.1152/jn.90291.2008>.

First-impression bias in MMN to auditory spatial deviants

Garrido, M. I., Kilner, J. M., Kiebel, S. J., Stephan, K. E., & Friston, K. J. (2007). Dynamic causal modelling of evoked potentials: a reproducibility study. *NeuroImage*, 36(3), 571-580. <http://dx.doi.org/10.1016/j.neuroimage.2007.03.014>.

Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: A review of underlying mechanisms. *Clinical Neurophysiology*, 120(3), 453-463. <http://dx.doi.org/10.1016/j.clinph.2008.11.029>.

Giard, M. H., Perrin, F., Pernier, J., & Bouchet, P. (1990). Brain generators implicated in the processing of auditory stimulus deviance: A topographic event-related potential study. *Psychophysiology*, 27(6), 627-640. <http://dx.doi.org/10.1111/j.1469-8986.1990.tb03184.x>.

Gilbert, C. D., & Sigman, M. (2007). Brain states: Top-down influences in sensory processing. *Neuron*, 54(5), 677-696. <http://dx.doi.org/10.1016/j.neuron.2007.05.019>.

Kolb, B. & Whishaw, I. Q. (2009). *An Introduction to Brain and Behaviour (3rd ed)*. New York, USA: Worth Publishers.

Kujala, T., Tervaniemi, M., & Schröger, E. (2007). The mismatch negativity in cognitive and clinical neuroscience: theoretical and methodological considerations. *Biological Psychology*, 74(1), 1-19. <http://dx.doi.org/10.1016/j.biopsycho.2006.06.001>.

Lieder, F., Daunizeau, J., Garrido, M. I., Friston, K. J., Stephan, K. E. (2013). Modelling trial-by-trial changes in the mismatch negativity. *Public Library of Science Computational Biology*, 9(2), <http://dx.doi.org/10.1371/journal.pcbi.1002911>.

Lieder, F., Stephan, K. E., Daunizeau, J., Garrido, M. I., & Friston, K. J. (2013). A neurocomputational model of the mismatch negativity. *Public Library of Science Computational Biology*, 9(11). <http://dx.doi.org/10.1371/journal.pcbi.1003288>.

First-impression bias in MMN to auditory spatial deviants

Matthews, N., Todd, J., Budd, T. W., Cooper, G. & Michie, P. T. (2007). Auditory lateralization in schizophrenia – Mismatch negativity and behavioral evidence of a selective impairment in encoding interaural time cues. *Clinical Neurophysiology*, 118(4), 833-844. <http://dx.doi.org/10.1016/j.clinph.2006.11.017>.

Mesulam, M. M. (1999). Spatial attention and neglect: Parietal, frontal and cingulate contributions to the mental representation and attentional targeting of salient extrapersonal events. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 354(1387), 1325-1346. <http://dx.doi.org/10.1098/rstb.1999.0482>.

Meyniel, F. & Dehaene, S. (2017). Brain networks for confidence weighting and hierarchical inference during probabilistic learning. *Proceedings of the National Academy of Sciences*, 114(9), 3859-3868. <http://dx.doi.org/10.1073/pnas.1615773114>.

Meyniel, F., Schlunegger, D. & Dehaene, S. (2015). The sense of confidence during probabilistic learning: A normative account. *Public Library of Science Computational Biology*, 11(6). <http://dx.doi.org/10.1371/journal.pcbi.1004305>.

Mullens, D., Winkler, I., Damaso, K., Heathcote, A., Whitson, L., Provost, A., & Todd, J. (2016). Biased relevance filtering in the auditory system: A test of confidence-weighted first-impressions. *Biological Psychology*, 115, 101-111. <http://dx.doi.org/10.1016/j.biopsycho.2016.01.018>.

Mullens, D., Woodley, J., Whitson, L., Provost, A., Heathcote, A., Winkler, I., & Todd, J. (2014). Altering the primacy bias--how does a prior task affect mismatch negativity? *Psychophysiology*, 51(5), 437-445. <http://dx.doi.org/10.1111/psyp.12190>.

Mumford, D. (1992). On the computational architecture of the neocortex. *Biological Cybernetics*, 66(3), 241-251. <http://dx.doi.org/10.1007/BF00198477>.

First-impression bias in MMN to auditory spatial deviants

Näätänen, R. (1992). *Attention and Brain Function*. Hillsdale, USA: Erlbaum.

Näätänen, R., Gaillard, A. W., & Mantysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica (Amsterdam)*, 42(4), 313-329.

[http://dx.doi.org/10.1016/0001-6918\(78\)90006-9](http://dx.doi.org/10.1016/0001-6918(78)90006-9).

Nassar, M. R., Wilson, R. C., Heasley, B., & Gold, J. I. (2010). An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *The Journal of Neuroscience*, 30(37), 12366-12378. <http://dx.doi.org/10.1523/jneurosci.0822-10.2010>.

Paavilainen, P., Alho, K., Reinikainen, K., Sams, M., & Näätänen, R. (1991). Right hemisphere dominance of different mismatch negativities. *Electroencephalography and Clinical Neurophysiology*, 78(6), 466-479. [http://dx.doi.org/10.1016/0013-4694\(91\)90064-B](http://dx.doi.org/10.1016/0013-4694(91)90064-B).

Paavilainen, P., Karlsson, M. L., Reinikainen, K., & Näätänen, R. (1989). Mismatch negativity to change in spatial location of an auditory stimulus. *Electroencephalography and Clinical Neurophysiology*, 73(2), 129-141. [http://doi.org/10.1016/0013-4694\(91\)90064-B](http://doi.org/10.1016/0013-4694(91)90064-B).

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology*, 32(1), 3-25. <http://doi.org/10.1080/00335558008248231>.

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1), 79-87. <http://doi.org/10.1038/4580>.

Ritter, W., Paavilainen, P., Lavikainen, J., Reinikainen, K., Alho, K., Sams, M., & Näätänen, R. (1992). Event-related potentials to repetition and change of auditory stimuli. *Electroencephalography and Clinical Neurophysiology*, 83(5), 306-321.

[http://doi.org/10.1016/0013-4694\(92\)90090-5](http://doi.org/10.1016/0013-4694(92)90090-5).

First-impression bias in MMN to auditory spatial deviants

Sams, M., Hari, R., Rif, J., & Knuutila, J. (1993). The human auditory sensory memory trace persists about 10 sec: Neuromagnetic evidence. *Journal of Cognitive Neuroscience*, 5(3), 363-370. <http://doi.org/10.1162/jocn.1993.5.3.363>.

Sams, M., Paavilainen, P., Alho, K., & Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalography and Clinical Neurophysiology*, 62(6), 437-448. [http://doi.org/10.1016/0168-5597\(85\)90054-1](http://doi.org/10.1016/0168-5597(85)90054-1).

Schall, U. (2016). Is it time to move mismatch negativity into the clinic? *Biological Psychology*, 116, 41-46. <http://doi.org/10.1016/j.biopsycho.2015.09.001>.

Semlitsch, H. V., Anderer, P., Schuster, P., & Presslich, O. (1986). A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. *Psychophysiology*, 23(6), 695-703. <http://doi.org/10.1111/j.1469-8986.1986.tb00696.x>.

Sussman, E. S., Chen, S., Sussman-Fort, J., & Dinces, E. (2014). The five myths of MMN: Redefining how to use MMN in basic and clinical research. *Brain Topography*, 27(4), 553–564. <http://doi.org/10.1007/s10548-013-0326-6>.

Todd, J., Heathcote, A., Mullens, D., Whitson, L. R., Provost, A., & Winkler, I. (2013). What controls gain in gain control? Mismatch negativity (MMN), priors and system biases. *Brain Topography*, 27(4), 578-589. <http://doi.org/10.1007/s10548-013-0344-4>.

Todd, J., Heathcote, A., Whitson, L., Mullens, D., Provost, A., & Winkler, I. (2014). Mismatch negativity (MMN) to pitch change is susceptible to order-dependent bias. *Frontiers in Neuroscience*, 8. <http://doi.org/10.3389/fnins.2014.00180>.

Todd, J., Petherbridge, A., Speirs, B., Provost, A., & Paton, B. (2017). Time as context: The influence of hierarchical patterning on sensory influence. *Schizophrenia Research*. <http://doi.org/10.1016/j.schres.2017.03.033>.

First-impression bias in MMN to auditory spatial deviants

Todd, J., Provost, A., & Cooper, G. (2011). Lasting first impressions: a conservative bias in automatic filters of the acoustic environment. *Neuropsychologia*, 49(12), 3399-3405.

<http://doi.org/10.1016/j.neuropsychologia.2011.08.016>.

Todd, J., Provost, A., Whitson, L. R., Cooper, G., & Heathcote, A. (2013). Not so primitive: context-sensitive meta-learning about unattended sound sequences. *Journal of Neurophysiology*, 109(1), 99-105. <http://doi.org/10.1152/jn.00581.2012>.

Todd, J., Provost, A., Whitson, L., & Mullens, D. (2017). Initial uncertainty impacts statistical learning in sound sequence processing. *Journal of Physiology Paris*.

<http://doi.org/10.1016/j.jphysparis.2017.01.001>.

Vecera, S. P. & Rizzo, M. (2003). Spatial attention: Normal processes and their breakdown.

Neurologic Clinics of North America, 21, 575-607. [http://doi.org/10.1016/S0733-](http://doi.org/10.1016/S0733-8619(02)00103-2)

8619(02)00103-2.

Vogel, J. J., Bowers, C. A. & Vogel, D. S. (2003). Cerebral lateralization of spatial abilities:

A meta-analysis. *Brain and Cognition*, 52(2), 197-204. [http://doi.org/10.1016/S0278-](http://doi.org/10.1016/S0278-2626(03)00056-3)

2626(03)00056-3.

Wacongne, C., Changeux, J. P., Dehaene, S. (2012). A neuronal model of predictive coding accounting for the mismatch negativity. *Journal of Neuroscience*, 32(11), 3665-3678.

<http://doi.org/10.1523/jneurosci.5003-11.2012>.

Winkler, I. (2007). Interpreting the mismatch negativity. *Journal of Psychophysiology*, 21(3-

4), 147-163. <http://doi.org/10.1027/0269-8803.21.34.147>.

Winkler, I., Karmos, G., & Näätänen, R. (1996). Adaptive modeling of the unattended acoustic environment reflected in the mismatch negativity event-related potential. *Brain Research*, 742(1-2), 239-252. [http://doi.org/10.1016/S0006-8993\(96\)01008-6](http://doi.org/10.1016/S0006-8993(96)01008-6).

First-impression bias in MMN to auditory spatial deviants

Zatorre, R. J. & Penhune, V. B. (2001). Spatial localization after excision of human auditory cortex. *Journal of Neuroscience*, 21(16), 6321-6328.

Author Notes

Grants and Acknowledgements

This research was supported by a National Health and Medical Research Council of Australia grant (APP1002995). K.F. acknowledges receipt of an Australian Postgraduate Award Scholarship. There was no involvement of funding bodies in the design, execution or write-up of this study. We thank Gavin Cooper for the programming of these experiments.

Ethics

Full approval of all experimental protocol was received from the Human Research Ethics committee of the University of Newcastle prior to commencement of the research, and written informed consent was received from all individual participants prior to undertaking the study.

Author Contributions

J.T. and K.F. conception and design of research; K.F. and J.T. analyzed data; K.F. and J.T. interpreted results of experiments; K.F., J.T. and A.P. prepared figures; K.F. drafted manuscript; K.F., J.T., and A.P. edited and revised manuscript; K.F., J.T. and A.P. approved final version of manuscript; K.F. and A.P. performed experiments.

Disclosures

No conflicts of interest, financial or otherwise, are declared by the authors.