



OPEN

## Relevance to the higher order structure may govern auditory statistical learning in neonates

Juanita Todd<sup>1</sup>✉, Gábor P. Háden<sup>2,3</sup> & István Winkler<sup>2</sup>

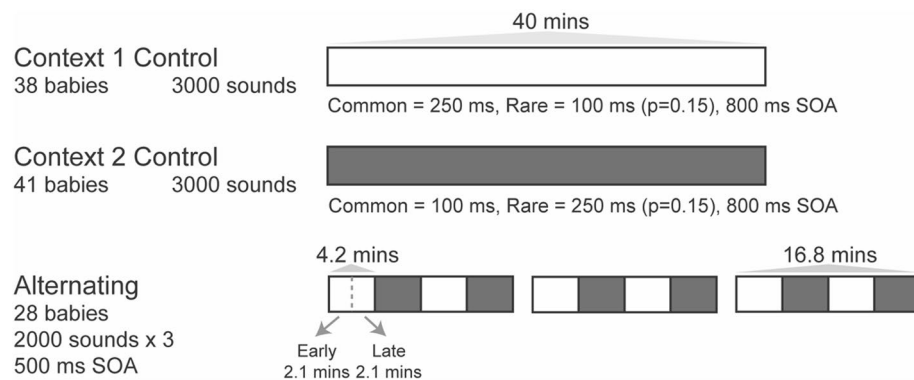
Hearing is one of the earliest senses to develop and is quite mature by birth. Contemporary theories assume that regularities in sound are exploited by the brain to create internal models of the environment. Through statistical learning, internal models extrapolate from patterns to predictions about subsequent experience. In adults, altered brain responses to sound enable us to infer the existence and properties of these models. In this study, brain potentials were used to determine whether newborns exhibit context-dependent modulations of a brain response that can be used to infer the existence and properties of internal models. Results are indicative of significant context-dependence in the responsivity to sound in newborns. When common and rare sounds continue in stable probabilities over a very long period, neonates respond to all sounds equivalently (no differentiation). However, when the same common and rare sounds at the same probabilities alternate over time, the neonate responses show clear differentiations. The context-dependence is consistent with the possibility that the neonate brain produces more precise internal models that discriminate between contexts when there is an emergent structure to be discovered but appears to adopt broader models when discrimination delivers little or no additional information about the environment.

Hearing is one of the earliest senses to mature with the developing fetus being capable of registering sound from as early as 18 weeks gestation<sup>1</sup>. By the time a new life exits the mother, it has quite a lot of experience with sound information<sup>2</sup>. We know newborns can recognize familiar voices from birth or earlier<sup>3,4</sup>, localize sound sources<sup>5</sup>, and even separate concurrent sound sources<sup>6</sup>. However, less is known about how sound is processed in relation to its context in the early stages of brain development. This study was designed to further our knowledge on how the neonate brain utilizes sound information within structured sequences.

The recording of auditory event-related brain potentials in this study utilizes a methodology that can be applied in neonates in the same way as in adults<sup>7</sup>. Electrodes are placed over the scalp surface while a structured sequence of sound is presented. This methodology has provided insight into statistical learning in the neonates that enables extraction of the kinds of abstract pattern learning required for music and language learning<sup>8–10</sup>. For sound sequences of predictable structure, with increased exposure, the brain becomes less responsive to predictable events and more responsive to events that mismatch one's predictions<sup>11</sup>. This is reflected in the event-related potential (ERP) responses. The ability to discriminate between responses to predictable and prediction-violating sounds can therefore be used to infer the existence of internal models of the environment representing a prediction or inference about the sensory input most likely to be encountered next<sup>11–13</sup>. This methodology is ideal for use in neonates because the process occurs automatically and does not require a task or explicit attention<sup>6,7,14–21</sup>.

In adults, the amplitude of the response to sounds that violate established patterns is proposed to reflect a precision-weighted signal influenced by the goodness of evidence upon which the internal model is built (e.g., the stability and reliability of the patterning<sup>22,23</sup>). A significant literature places the precision-weighting in the gain on output from specific cell populations in responsive cortical areas (namely the superficial pyramidal cells) acknowledging that this reflects a hierarchical network interaction both within and between brain regions. In other words, a seemingly simple deviance-detection indicator is subserved by a sophisticated iterative inferential process that operates over many different timescales<sup>23–27</sup>. Within scalp-recorded responses, this process has been linked to change over time in several components of the auditory ERP<sup>28</sup>. In adults, the key difference between pattern-matching and mismatching sound emerges within 100–250 ms after the point of deviance and is maximally recorded over fronto-central electrode sites when a nose or mastoid reference is used. This “mismatch”

<sup>1</sup>School of Psychological Sciences, University of Newcastle, Callaghan, Australia. <sup>2</sup>Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, Budapest, Hungary. <sup>3</sup>Department of Telecommunications and Media Informatics, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, Budapest 1117, Hungary. ✉email: Juanita.Todd@newcastle.edu.au



**Figure 1.** Depiction of the sound sequences used for the control and alternating experiments. The control context sequence was a single stream of sounds about 40 min in length. The alternating sequence contained four  $\times$  4.2 min blocks of 500 sounds (16.8 min total length) and it was presented three times with breaks between. For data analysis, the alternating sequence data was further divided to capture early and late periods of time within the two context blocks with indicative periods marked in the figure.

sensitivity (mismatch negativity or MMN) is captured in a difference waveform but in the tone responses it often overlaps early components known as the N1 and P2 and is considered part of the N2 complex<sup>29–31</sup>. The precision-weighting can be seen both as changes in the response to pattern matching sounds that become less negative with repetition and pattern stability over time and in the mismatch response which becomes more negative over this period.

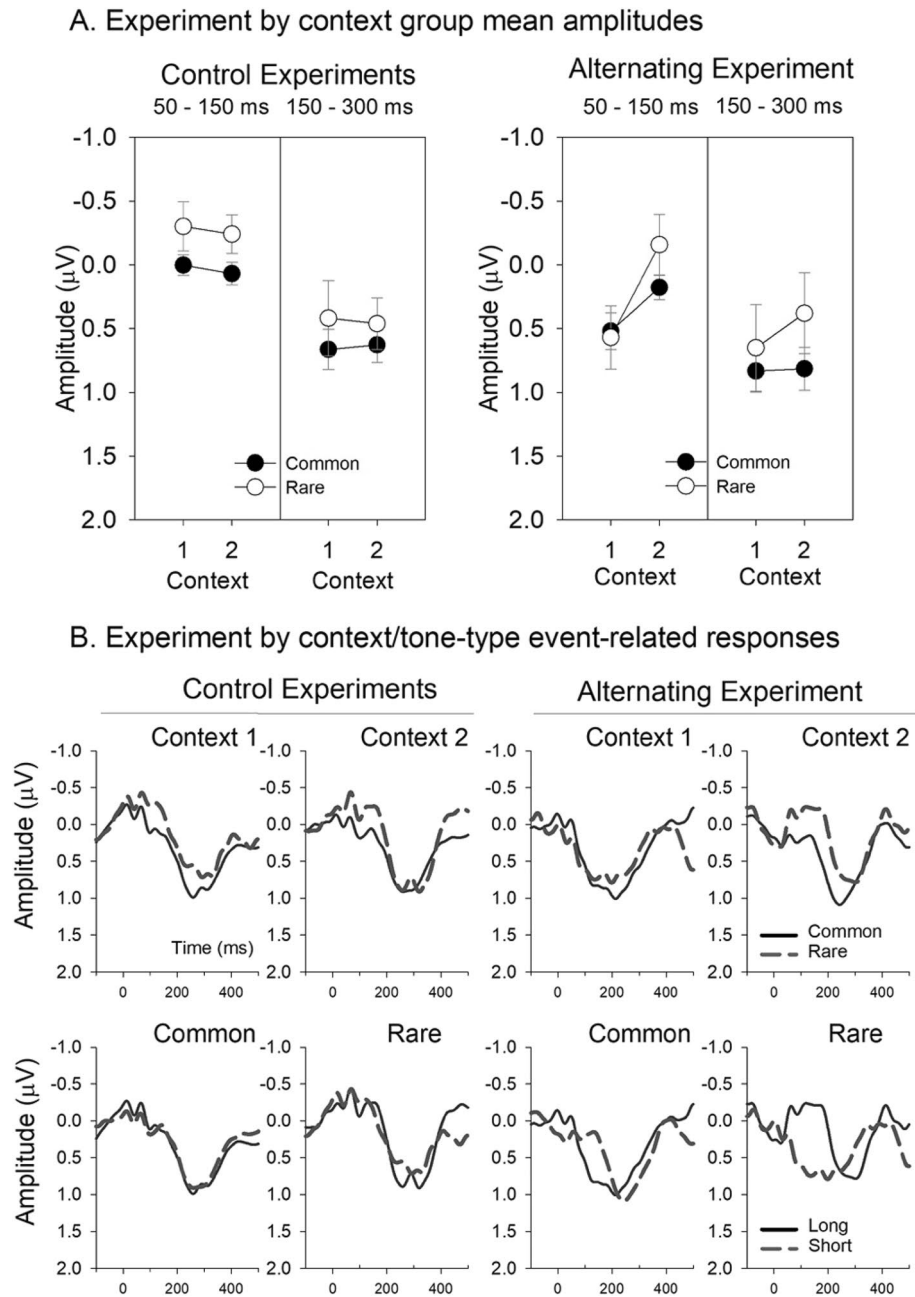
While the morphology of ERPs is rapidly changing within the first year of life and the stable componentry of adulthood only emerges during childhood, ERP signs of detecting auditory deviance are present right from birth<sup>32</sup>. The neonatal layout of the deviance-related ERP response (derived by subtracting the response to regular sounds from irregular ones) is typically dominated by a positive waveform appearing 200–400 ms from the onset of the deviation (termed the mismatch response; MMR<sup>33</sup>), which may be surrounded by an early and a late negative difference, depending on the acoustic properties of the stimuli, the amount and type of separation between the regular and irregular sounds, etc<sup>34</sup>. Although initially understood to be an indicator of probability<sup>29,30</sup>, the mismatch process shows sensitivity to violations of established sound transitions<sup>27,35,36</sup> requiring a more contextualized form of learning. This type of statistical learning involves extraction of regularities in how features and objects co-occur in the environment over space and time. Importantly, previous studies showed that properties of the infantile electric brain response to auditory deviance suggest that it reflects a prediction error<sup>37</sup>, similarly to the analogue response in adults<sup>38</sup>. The infantile MMR response is elicited both by violating regularities based on individual sounds and on short sound patterns<sup>39</sup>, bringing up the possibility that, again similarly to adults<sup>40</sup>, they may also reflect the large-scale structure of the auditory context.

The study was designed to determine whether the differential precision-weighting based on large scale structure seen in adults<sup>41–44</sup> might also be observed in neonates. Here we provide evidence that the brain of newborn infants can indeed form internal models that differentiate sounds based on probability, and weight that response based on precision, but that the propensity to do this depends on contextual factors that may be related to a determination of the inferred value of tracking information over time. Differential responses to sound may only occur when a sound provides information about longer-term structures in the environment. Thus, the same acoustic differentiation may or may not be automatically made, depending on its relevance for revealing important information about the context.

## Results

The study featured three experiments with three separate groups of neonates with the sound sequence design depicted pictorially in Fig. 1. In the first two experiments, the sequences heard by babies contained just two sounds; a long tone (250 ms) and a shorter tone (100 ms) that both comprised an overlay of 500 Hz, 1000 Hz and 1500 Hz frequencies. The sounds were organized as traditional “oddball” sequences<sup>11</sup> such that the two sounds were delivered with largely different probabilities (long  $p=0.85$  and short  $p=0.15$ , respectively). Experiment 1 is referred to hereafter as the context 1 control. A different group of newborns heard similar sequences but with the sounds in reverse probabilities (Experiment 2), which is hereafter referred to as the context 2 control. When adults hear sequences of this kind, the responses to the common and rare tones differentiate such that the response becomes more “suppressed” or more positive at fronto-central scalp sites with increased exposure, and “larger” or more negative to rare deviations<sup>11</sup>. Stable sequences of this kind are also considered ideal for formation of precise internal models and were designed based on sequences for which probability-dependent response were observed previously in neonates<sup>16</sup>. However, neither group of infants produced responses to rare tones that were significantly different to the common tones, and indeed the responses from each group were remarkably similar, both in amplitude and morphology.

In auditory ERPs, the early period of response is considered to reflect relatively more sensory-based processes with the later components progressively more cognitive in nature<sup>45</sup>. In newborn infants, the latency of responses elicited by auditory sensory deviance have been found to depend on both the specific acoustic feature



**Figure 2.** (A) Group mean amplitudes over 50–150 ms and 150–300 ms from tone onset for the common and rare tone responses in the two control experiments (left) and the alternating experiment (right), separately for the two contexts. The error bars represent the standard error of the mean. (B) Group averaged event-related brain responses at Cz for the common and rare tones in the two control experiments (left plots) and the alternating (right plots) experiment. Event-related brain responses are presented in two ways: (I) overlaid tone type by context (top) and (II) overlaid tone duration by probability (bottom).

and the amount of separation<sup>32</sup>. In previous neonate studies of tone duration sensitivity, responses to common and rare sounds tended to differentiate at about 150 ms through to 300 ms<sup>16</sup>. To capture this window, responses were quantified using the mean amplitude over 150–300 ms at the Cz scalp electrode. An additional period of analysis was included to capture earlier components using a mean amplitude over 50–150 ms. The component amplitudes are presented in Fig. 2A (left) and the accompanying responses in Fig. 2B (left) and see also Supplementary figures for full scalp montage. There was no overall condition effect, nor any main effects or interactions for either analysis window when response amplitudes were compared in separate mixed model ANOVAs with a between-group factor for *context* (context 1, context 2) and a within-subject factor for *probability* (common, rare).

In adults and neonates, the absence of a differentiation in response to a rare and common sound with different physical properties is typically considered evidence that the two tones could not be discriminated<sup>12,46–48</sup>.

Certainly, the brain responses of our two groups of infants show no significant evidence of a discrimination being made between the rare and common sounds in either control sequence. Thus, by these data alone, one would infer that these neonates were not able to discriminate the duration difference between 100 and 250 ms long tones. However, contexts 1 and 2 were quite clearly distinguished when the same tone sequences were concatenated alternately in a continuous sequence. Further, there was also indication of the rare properties being differentiated over time.

In the alternating contexts experiment (Experiment 3), another group of newborns were presented with the same tones as those used in the control contexts, but organized into an alternating sequence, switching back and forth between periods of context 1 and 2 modelled on sequences used in adults<sup>41,49,50</sup> (see Fig. 1). The responses were remarkably different from those obtained in the control contexts. In Fig. 2A (right), B (right), the responses to tones presented in context 2 appear very similar to those in the control experiments, but the responses to tones in context 1 were clearly different. The morphological difference is evident in an earlier positive deflection in response to tones in context 1 of the alternating experiment: a positive shift begins approximately 100 ms earlier (that is, at about 50 ms in context 1 compared to ca. 150 ms in context 2). This morphological difference is not only evident between the context 1 and context 2 within the alternating experiment, but also between the control and the alternating context 1.

A repeated measures ANOVA of amplitudes obtained in the alternating sequence with within-subjects factors of *context* (context 1, context 2) and *probability* (common, rare) revealed a significantly larger positive polarity response over 50–150 ms in context 1 ( $F(1,27) = 6.269, p < 0.019, \eta^2 = 0.188$ ) reflecting this earlier positive response to tones in context 1 than 2. A mixed model ANOVA comparing context 1 early responses between the corresponding control and alternating experiments also indicated significantly more positive responses over 50–150 ms in the alternating sequence group than in the control group ( $F(1,64) = 16.968, p < 0.001, \eta^2 = 0.210$ ). There were no significant main effects or interactions for the equivalent comparisons of response amplitudes for context 2. These differences are clear in Fig. 2 where the mean amplitude for this period is clearly more positive in the alternating context 1. The same repeated measures ANOVA was applied to the responses over 150–300 ms revealing no significant differences nor interactions. In summary, unlike for the two control contexts, the infants responded differently to the same two contexts when they occurred within a single sequence of sound: the dominant positivity began earlier in alternating context 1, than in alternating context 2, with only context 1 differing from either control context. There was, however, no overall significant differentiation between the common and rare tones in either context, for either window, in either experiment.

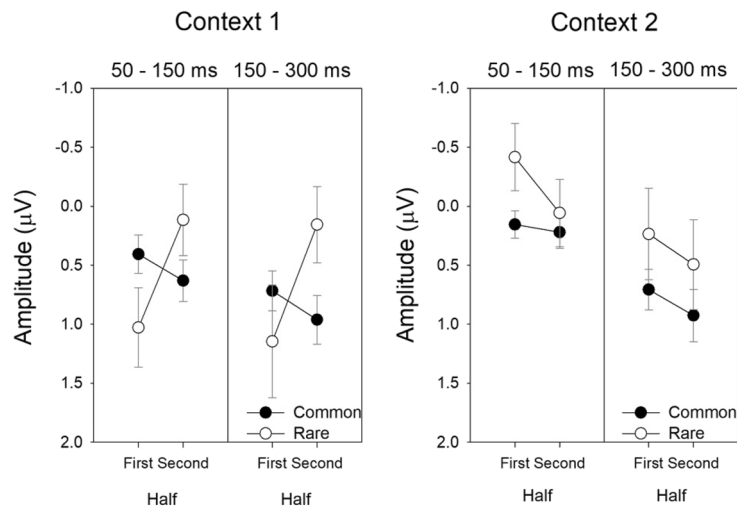
Given evidence that the infants could differentiate the contexts, and evidence in adults that the precision-weighting on auditory responses change dramatically over time within alternating sequences<sup>41,49,50</sup>, the data was further explored for evidence of probability differentiation by separately examining responses obtained from the first 2.1 min and second 2.1 min periods (1st and 2nd half, respectively). The component amplitudes for these divisions of the data are presented in Fig. 3A and the accompanying responses in Fig. 3B and see also Supplementary figures for full scalp montage. The data in Fig. 3, once again, display the much larger early positivity in response to context 1 tones in general, but there are also signs of differentiation between the rare and common tones in the second 2.1 min period over the later (150–300 ms) measurement window. A repeated measures ANOVA [*probability* (common, rare)  $\times$  *context* (context 1, context 2)] was conducted to test the responses for effects of probability difference, separately for the first and the second 2.1 min periods and for the early and late measurement intervals. Examination of the responses in the first 2.1 min period revealed a significant effect of context over the 50–150 ms time window ( $F(1,27) = 11.393, p < 0.002, \eta^2 = 0.297$ ), where responses were significantly more positive in the first than the second context. The effect was further modified by tone probability ( $F(1,27) = 5.464, p < 0.027, \eta^2 = 0.168$ ) with the differences being larger between rare than between common tones, of which only the former reached significance ( $t(27) = 4.034, p_{\text{holm}} < 0.001$ ). There were no significant main effects or interactions for amplitudes measured in the 150–300 ms window over this first 2.1 min. However, importantly, no significant main effect was found for tone probability over this first 2.1 min period for either context.

In contrast, examination of responses in the second 2.1 min period revealed no differences over 50–150 ms. However, there was a main effect of probability for the 150–300 ms period ( $F(1,27) = 7.284, p < 0.012, \eta^2 = 0.212$ ). As can be seen in Fig. 3, the responses to the rare tones were significantly less positive (more negative) than those to the common tones ( $m = 0.325 \mu\text{V}$  vs.  $m = 0.944 \mu\text{V}$ , respectively). In summary, the differentiation in the event-related brain responses within the first 2.1 min primarily reflected the context that the tones were presented in, but by the second 2.1 min the tone probabilities were differentiated with the rare ones eliciting more negative response over the 150–300 ms; the same window as previously observed to capture a negative mismatch response<sup>16</sup>.

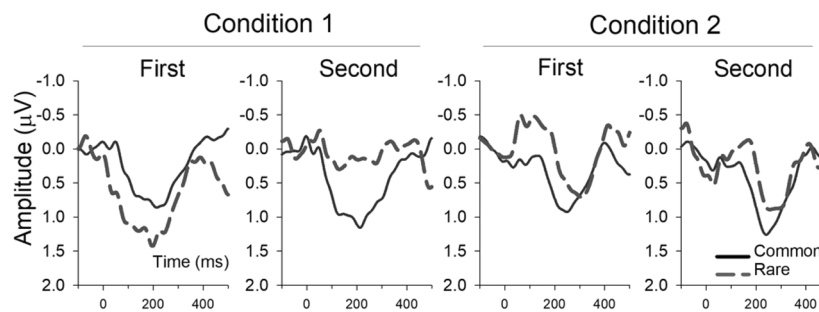
## Discussion

The context-dependence of neonate responses to sounds imply that the infant brain should be considered an intelligent consumer of information rather than as a passive processor of sound sequence material. Internal models of the environment are proposed to contain an estimate of how volatile the environment is (cf. “precision” in predictive coding theories<sup>23,25</sup>). The estimate of volatility/precision then weights the brain response leading to changes in response amplitude over time. Where a model has been formed based on highly reliable information, the response to a model-matching sound should be very suppressed and responses to model-mismatches highly salient<sup>23,25,51</sup>. The classic oddball sequences used for control contexts 1 and 2 are considered ideal learning environments for building precise internal models provided that the brain can distinguish between the common and rare elements. Earlier studies indeed suggest that the infant brain can distinguish the durations used here<sup>15,16</sup>, and the present data from the alternating sequences indicates that our infants did also discriminate the two durations. The key question is why the neonates did not do so in the classic oddball control sequences. The main findings

### A. Context by half mean amplitudes in the alternating contexts experiment



### B. Context by half event-related responses



**Figure 3.** (A) Group mean amplitudes in the alternating context experiment over 50–150 ms and 150–300 ms from tone onset for the common and rare tone responses for context 1 (left) and context 2 (right) for the first and the second 2.1 min long periods within the context blocks. The error bars represent the standard error of the mean. (B) Group averaged event-related brain responses at Cz for the common and rare tones in context 1 (left) context 2 (right).

of the present study are discussed below in relation to a thesis that the infant brain is sensitive and responsive to the information carried by sound, and that brain responses are adjusted to this determination accordingly. This sensitivity to the information carried by sound would be reliant upon the extraction of regularities in how features and objects co-occur in the environment over longer timescales which expands the view of statistical learning to the large scale structure.

The general morphology of the neonatal ERPs recorded in the current study conformed to the pattern observed in many previous studies: a negative waveform in the 50–150 ms time window followed by a dominant centrally positive response in the 150–400 ms time range. The two contexts used in these experiments were characterized by different tendencies towards two sounds that differed from each other in duration (100 ms and 250 ms). The responses elicited to the common and rare instances of these two sounds were remarkably similar in the control experiments. In contrast, the juxtaposition of the two contexts in the alternating sequence revealed an ability to discern the boundaries within the sequence where the probabilities of the two sounds changed, moving from a stronger tendency towards encountering the shorter tone after a period of hearing a stronger tendency towards encountering the longer tone and vice versa. The key indicator that the babies detected these boundaries lies in the point at which the positive shift onsets in the response morphology, which is much earlier for context 1 than 2, and much earlier in context 1 for the alternating experiment than for the control experiment. There was therefore a strong dependence of the morphology of the infant response based on the very long-term structure of the listening environment.

There are a number of properties that could influence the responses to tone changes at the boundaries through a process called adaptation (note, here we refer to adaptation in larger scale responses not at individual neuronal level, for distinctions<sup>52,53</sup>). Stimulus specific adaptation is an important mechanism through which the auditory system tunes responsivity to reflect redundancy and sharpen discrimination ability<sup>54</sup> and distinguish information carrying signals from background noise<sup>55</sup>. The processes underlying how sound statistics affect adaptation are complex and sophisticated<sup>52,53</sup>, and while they are argued to be insufficient to fully account for deviance sensitivity<sup>24</sup>, they could contribute to the general context effect in the alternating sequence. Each of these is



discussed in turn with respect to how well they fit the current data. Firstly, sounds of different duration carry different levels of sound energy<sup>56</sup>, and as such, it is possible that context 1 and 2 could have been differentiated based on a general shift in sound intensity over time. This sensory property-based distinction could be considered consistent with an effect appearing early within the event-related response<sup>45</sup>, possibly even earlier than the point of difference in tone length at 100 ms (see Fig. 2, alternating condition). The differences in sound duration used here can affect the degree of adaptation raising the possibility that context 1 responses are more positive in the early period because of a higher level of adaptation. However in adults, longer duration tones have been shown to increase the amplitude of the positive component that occurs at about 200 ms after tone onset, but not the earlier components<sup>57</sup>. Another feature of the contexts that could affect adaptation is the silent period between tones in the alternating sequence which was typically 400 ms in context 2 (100 ms common tones with 500 ms stimulus onset asynchrony (SOA)) but only 250 ms in context 1 (250 ms common tones with a 500 ms SOA). In adults, this difference in silent interval changes the degree of adaptation. However, once again in adults this is evident mostly approximately 200 ms after tone onset<sup>57</sup>. Finally, the silent intervals are slightly longer in the control conditions (550 ms for context 1 and 700 ms for context 2 which) which also could be associated with different levels of adaptation and could again affect the amplitude of the positive component that occurs at about 200 ms after tone onset. However, a difference in adaptation is not evident between the two control conditions. Further, this effect would create a more positive response in the control context 1 than in the alternating context 1<sup>57</sup>. In contrast the opposite of this has been found here. In sum, the current data does not support the notion that adaptation alone is likely to explain the effects in neonates distinguishing the different contexts or the alternating and the control sequences. Therefore, instead it is proposed that the distinction in the responses only emerges where the statistical properties of sound carry information about a longer-term structure within the sound sequence. In other words, we propose that the babies can distinguish between the two sound contexts, but they do so only when differentiation is meaningful to isolate information needed to encode larger scale structure.

It was also only when the sounds were embedded within a longer structure that evidence of an ability to discern one sound as a local rare outlier emerged. When responses from the first and second half of the contexts were separately averaged, it became evident that responses appearing initially within a “new” context tended to reflect the differentiation of the context only (i.e., the difference in early positivity). Later on, when the context had been stable for a while, a significant longer-latency negative shift was seen in response to the rare tone in context 1, consistent with prior observations of deviance detection<sup>16</sup>. The infants were detecting the sound that was out-of-place or unexpected in that context and responding in a similar way to that which we see in adults (i.e., more negative response<sup>47,58,59</sup>). Interestingly, the data also suggest this was particular to the first context heard. The first context, regardless of whether long or short tones are common, is also associated with higher precision models in adults<sup>41–44</sup>. The alternating sequence responses therefore provide evidence that newborns can not only determine the sequence structure<sup>60</sup>, but can also form internal models that build up in precision over time enabling an outlier to be detected. A progressive refinement of cortical response in neonates over time has also been seen in statistical learning of language<sup>8</sup> which is also compatible with the concept of precision-weighting but in the present case it is demonstrated with respect to relative probability. These observations add to a body of work showing how such learning is shaped by many longer timescale characteristics including hierarchy, order, entropy, and uncertainty<sup>9</sup>.

So why were infants not sensitive to the outliers when presented within a listening context that should have been ideal to observe differential responses? One possible reason for the longer-term structure of the sound sequence influencing sound differentiation may relate to a trade-off between the information value of maintaining precise internal models and the associated investment of effort. In the control sequences, the babies heard 3000 tones in a constant uninterrupted probability arrangement over a 40-min period. In the alternating sequence, the exposure to the two contexts occurred in blocks of 500 stimuli: thus, the maximum period of exposure to one sound probability arrangement was shorter than 7 min. Similarly, in prior studies showing the ability of young infants to distinguish between sounds of two different durations<sup>15,16</sup> the infants always heard sequences of less than 7 min duration with breaks between successive presentations. Thus, a potential explanation for the absence of tone differentiation in the current control sequences is that rare changes to the duration of sound within a long constant sequence carried insufficient information to warrant considering a different model, as there was no alternative offered by the sound input. Once again, the babies can distinguish between the two sounds, but perhaps cease to do so when it offers no new information about structure.

When exposed to oddball sequences over a long period of time, a broader and more inclusive model that includes both sound durations would avoid attention capture by the unimportant deviations, potentially conserving energy for modelling other aspects of the environment<sup>23,25,61</sup>. Under this broader more inclusive model, all sounds heard may be captured as predicted such that the absence of differentiation in the responses does not reflect an inability to distinguish them, but rather a determination that distinguishing between them carries insufficient benefit to warrant the effort involved. In contrast, when these same two sounds occur within the alternating sequences, two different acoustic states are present in the environment and modelling these two different states may become energetically more beneficial. This suggests that the optimization of resources, a potential driver of cognitive development<sup>62</sup>, is already evident in the first few days after birth. The above account is fully compatible with a predictive coding theory of brain function suggesting that a fundamental operating principle is to seek to reduce variational free energy, that is to reduce surprise by utilizing statistical information to anticipate upcoming states<sup>23,61,63</sup>. Observations enable accumulation of a probability density over their hidden causes, and through a process of active inference, an internal model of the most likely next state can be inferred and refined over time based on the adequacy of the fit to new evidence. The optimization of a model to reduce free energy can be governed by weighting whether to aim for accuracy of predictions over the shorter or longer term by the energy required to achieve that level of accuracy. Developing or maintaining sensitivity to the two sound durations is beneficial to extracting the predictable structure in the alternating sequence, but not in

the control sequence. This kind of contextualized processing is another example of how expectations can shape responses to sound<sup>64</sup>, with the present observations suggesting that some such modulations are so fundamental to learning that evidence can be seen in the first few days of life, even in sleeping neonates.

A limitation of the present study, as noted in the introduction, is that the stable componentry of adulthood only emerges during childhood. The responses observed in neonates are incomparable to those in adults (e.g., neither the N1 nor N2 component has an equivalent in neonates; unlike the amplitude of the adult response, MMR is affected by acoustic parameters of the deviant stimulus, etc.<sup>32</sup>) which complicates the study the computational principles underlying them (see for discussion<sup>32</sup>). There are a number of studies, for example, that refer to a larger central positivity as being indicative of deviance detection for some features in infant data<sup>33</sup> but this was not observed here and may not be how duration deviance is reflected in neonate ERPs where previously, a more negative response has been observed<sup>15,16,65,66</sup>. Notwithstanding, we suggest the context-sensitivity and changes over time in responsiveness present in these neonate data are indicative of the way perception is shaped by modelling of complex patterning in the environment.

In sum, the infant data presented here provide two insights into the newborn's readiness to process sounds within context. Perhaps not so surprisingly, the newborn is a discerning listener. Far from being a passive receptacle for sensory experiences, the neonate brain appears to be governed by determinations of a trade-off between information value and energy efficiency. This cost–benefit equation seems to result in a conservation of energy in situations, where no alternative models can be considered. However, the neonates do adopt more precise models when a single model cannot adequately account for multiple states. Under these circumstances there is clear evidence of altered responsiveness that enables a gross differentiation of states (between the two contexts) and the emergence of differentiation of probable features in a precision-weighted manner. These feats are important for infants to learn from the environment, finding regular sequences of events within constantly changing stimulation of typically several sources of information/stimulation present concurrently. This ability would seem pivotal to explain the readiness with which infants learn critical interaction skills such as understanding speech and entering into communication exchange and dialogues<sup>60,67</sup>.

## Materials and methods

**Participants.** Electroencephalographic (EEG) data useful for the current analyses was recorded from 110 (56 male) healthy, full-term infants 0–6 days postpartum. Data collected from another 14 infants had to be discarded: three due to recording errors (no trigger was recorded) and eleven due to failing to meet criteria for data quality (see the “[EEG recording and analysis](#)” section below). The infants were divided into three groups (experiments): the control context 1 group (n = 38, 20 male) had a mean gestational age of 39.19 weeks (SD = 0.84), birth weight 3472 g (SD = 522.21); the control context 2 group (n = 41, 24 male) had a mean gestational age of 39.48 weeks (SD = 1.03), birth weight 3421 g (SD = 416.95) and the alternating contexts group (n = 28, 12 male) had a mean gestational age of 39.12 weeks (SD = 1.03), birth weight 3404 g (SD = 457.50). These sample size are commensurate or in excess of those in prior similar studies<sup>15,16</sup> and g\*power estimates suggest the power of observing a significant mismatch response if present exceeded 0.99. All participants had an Apgar score of 9/10. Infants were predominantly asleep during the recordings, in quiet sleep 78% of the time, and in active sleep 16% of the time. Informed consent was obtained from one (mother) or both parents. The experiment was carried out in a dedicated experimental room. Experiments 1 & 2 were carried out concurrently with babies allocated to experiments in counterbalanced order. Experiment 3 was carried out separately with all babies hearing the same sequences. Importantly, all recordings were carried out in the same manner, in the same setting. The study was conducted in full accordance with the Declaration of Helsinki and all applicable national laws and was approved by the relevant ethics committees: Medical Research Council-Committee of Scientific and Research Ethics (ETT-TUKEB), Hungary.

**Sound sequences.** Each group were presented with a sequence of harmonic tones delivered binaurally using the E-Prime stimulus presentation software (Psychology Software Tools, Inc., Pittsburgh, PA, USA) with ER-1 headphones (Etymotic Research Inc., Elk Grove Village, IL, USA) connected via sound tubes to self-adhesive ear cups (Sanibel Supply, Middelfart, Denmark) placed over the infants' ears.

The sequence presented in the control conditions were based on sounds used previously in similar infant studies yielding significantly different responses to common and rare sounds (11, 12). The two harmonic complex tones of different duration (100 ms, 250 ms with 5 ms cosine rise and fall each), comprised of three pure sine frequency components (500 Hz, 1000 Hz and 1500 Hz with – 5 dB/step intensity). All tones were delivered with a uniform 70 dB intensity. The tones were presented in a typical oddball design<sup>11</sup> with one sound being common (p = 0.85) and the other rare (p = 0.15), and the sounds occurring pseudo randomly constrained such that a minimum of three common sounds occurred between rare sounds. In context 1, the long sound was common, and the short sound was rare, and in context 2, the probabilities were reversed.

A diagram depicting the structure of the sequences presented to each group is presented in Fig. 1. In experiments 1 and 2 (control) 3000 tones were presented in a single unbroken sequence at a regular 800 ms stimulus onset asynchrony (SOA) with one group receiving a sequence made up as context 1 and the other as context 2. In experiment 3 (alternating contexts), the tone sequences comprised the same two sounds of different duration but an SOA of 500 ms, which were presented in groups of 500 alternating back and forth between context 1 and 2. The shorter SOA was necessary to preserve timing similarity to alternating sequences in adults<sup>41</sup> and SOAs of 300–500 ms have been used successfully in other neonates studies<sup>68,69</sup>. Each group of 500 sounds began with a minimum of four occurrences of the new common sound. The sequence was presented three times with a 1 min silent period between them and always started with a context 1 block.

**EEG recording and analysis.** EEG was recorded with Ag/AgCl electrodes attached to the Fp1, Fp2, Fz, F3, F4, F7, F8, T3, T4, Cz, C3, C4, Pz, P3, P4 locations according to the International 10–20 system. The reference electrode was placed on the tip of the nose and the ground electrode on the forehead. Data was recorded using a direct-coupled amplifier (V-Amp, Brain Products, Munich, Germany) at 24-bit resolution and a sampling rate of 1000 Hz. Data processing was conducted using the ERPLab<sup>70</sup> plugin to EEGLAB<sup>71</sup> within MATLAB. The EEG signals were passed through a digital bandpass filter between 1 and 25 Hz and epochs of –100 to 500 ms with respect to the sound onset were extracted for each stimulus. All amplitude measurements and illustrations are referred to the 100-ms pre-stimulus interval, the baseline.

Epochs were subjected to an artifact rejection process where epochs with a voltage change exceeding 150  $\mu$ V were rejected from analysis. The data quality was considered acceptable if a minimum of 50% of the epochs were retained for every condition tested, for electrodes F3, Fz, F4, C3, Cz, C4, P3, Pz, P4. Responses were averaged separately according to sound duration and probability (context 1: long common and short rare tone; context 2: short common and long rare tone). For testing the long vs. short discrimination, the responses in the alternating condition were further divided into averages that separately captured the first and second halves (the first and second 2.1 min period, respectively; see Fig. 3) within each context block. The separation of the first and second half of the context blocks has been very useful in analyzing data from adults as the responses change significantly over time<sup>49,72</sup>.

Two time windows were used to quantify the amplitudes within the central (Cz) responses, the electrode selected as it typically best captures deviance-related effects in neonates (e.g.,<sup>32,34</sup>). The first was the mean amplitude between 50 and 150 ms after sound onset to capture early components of the response. The second was the mean amplitude between 150 and 300 ms after sound onset which captures the period over which a significant difference was present in the response to common and rare tones in prior work (12) indicating that a mismatch to the prevailing internal model was detected. Separate ANOVA analyses were conducted on the 50–150 ms and 150–300 ms amplitude measurements. These are described in the results section with some being within-subject and others between-subject as required. Post-hoc Holm-Bonferroni corrected tests were employed to explore significant interactions. Effect sizes are reported as partial eta squared and the significance criterion level was  $\alpha < 0.05$ . All statistics were computed using JASP (version 0.13.1<sup>73</sup>).

Received: 5 December 2021; Accepted: 23 March 2022

Published online: 07 April 2022

## References

- Moore, J. K. & Linthicum, F. H. Jr. The human auditory system: A timeline of development. *Int. J. Audiol.* **46**, 460–478. <https://doi.org/10.1080/14992020701383019> (2007).
- Partanen, E., Kujala, T., Tervaniemi, M. & Huotilainen, M. Prenatal music exposure induces long-term neural effects. *PLoS ONE* **8**, e78946. <https://doi.org/10.1371/journal.pone.0078946> (2013).
- DeCasper, A. J. & Fifer, W. P. Of human bonding: Newborns prefer their mothers' voices. *Science* **208**, 1174–1176. <https://doi.org/10.1126/science.7375928> (1980).
- Kisilevsky, B. S. *et al.* Fetal sensitivity to properties of maternal speech and language. *Infant. Behav. Dev.* **32**, 59–71. <https://doi.org/10.1016/j.infbeh.2008.10.002> (2009).
- Morrongiello, B. A., Fenwick, K. D., Hillier, L. & Chance, G. Sound localization in newborn human infants. *Dev. Psychobiol.* **27**, 519–538. <https://doi.org/10.1002/dev.420270805> (1994).
- Bendixen, A. *et al.* Newborn infants detect cues of concurrent sound segregation. *Dev. Neurosci.* **37**, 172–181. <https://doi.org/10.1159/000370237> (2015).
- Haden, G. P., Nemeth, R., Torok, M. & Winkler, I. Mismatch response (MMR) in neonates: Beyond refractoriness. *Biol. Psychol.* **117**, 26–31. <https://doi.org/10.1016/j.biopsycho.2016.02.004> (2016).
- Bosseler, A. N., Teinonen, T., Tervaniemi, M. & Huotilainen, M. Infant directed speech enhances statistical learning in newborn infants: An ERP study. *PLoS ONE* **11**, e0162177. <https://doi.org/10.1371/journal.pone.0162177> (2016).
- Daikoku, T. Neurophysiological markers of statistical learning in music and language: Hierarchy, Entropy, and uncertainty. *Brain. Sci.* <https://doi.org/10.3390/brainsci8060114> (2018).
- Thompson, A., Aguero, A. & Lany, J. In *Neural Circuit and Cognitive Development* (Second Edition) (eds J. Rubenstein, P. Rakic, B. Chen, & K. Y. Kwan) 319–333 (Academic Press, 2020).
- Naatanen, R., Paavilainen, P., Alho, K., Reinikainen, K. & Sams, M. Do event-related potentials reveal the mechanism of the auditory sensory memory in the human brain?. *Neurosci. Lett.* **98**, 217–221 (1989).
- Winkler, I., Karmos, G. & Naatanen, R. Adaptive modeling of the unattended acoustic environment reflected in the mismatch negativity event-related potential. *Brain Res.* **742**, 239–252 (1996).
- Alho, K., Sams, M., Paavilainen, P., Reinikainen, K. & Naatanen, R. Event-related brain potentials reflecting processing of relevant and irrelevant stimuli during selective listening. *Psychophysiology* **26**, 514–528 (1989).
- Huotilainen, M. *et al.* Auditory magnetic responses of healthy newborns. *NeuroReport* **14**, 1871–1875. <https://doi.org/10.1097/00001756-200310060-00023> (2003).
- Cheour, M., Kushnerenko, E., Ceponiene, R., Fellman, V. & Naatanen, R. Electric brain responses obtained from newborn infants to changes in duration in complex harmonic tones. *Dev. Neuropsychol.* **22**, 471–479 (2002).
- Ceponiene, R. *et al.* Event-related potential features indexing central auditory discrimination by newborns. *Brain Res. Cogn. Brain Res.* **13**, 101–113. [https://doi.org/10.1016/s0926-6410\(01\)00093-3](https://doi.org/10.1016/s0926-6410(01)00093-3) (2002).
- Ruusuvirta, T., Huotilainen, M., Fellman, V. & Naatanen, R. The newborn human brain binds sound features together. *NeuroReport* **14**, 2117–2119. <https://doi.org/10.1097/00001756-200311140-00021> (2003).
- Winkler, I. *et al.* Newborn infants can organize the auditory world. *Proc. Natl. Acad. Sci. USA* **100**, 11812–11815. <https://doi.org/10.1073/pnas.2031891100> (2003).
- Kujala, A. *et al.* Speech-sound discrimination in neonates as measured with MEG. *NeuroReport* **15**, 2089–2092. <https://doi.org/10.1097/00001756-200409150-00018> (2004).
- Stefanics, G. *et al.* Auditory temporal grouping in newborn infants. *Psychophysiology* **44**, 697–702. <https://doi.org/10.1111/j.1469-8986.2007.00540.x> (2007).



21. Trainor, L. J. Predictive information processing is a fundamental learning mechanism present in early development: Evidence from infants. *Int. J. Psychophysiol.* **83**, 25 (2012).
22. Friston, K. & Kiebel, S. Cortical circuits for perceptual inference. *Neural Netw.* **22**, 1093–1104. <https://doi.org/10.1016/j.neunet.2009.07.023> (2009).
23. Friston, K. A theory of cortical responses. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* **360**, 815–836 (2005).
24. Garrido, M. I., Kilner, J. M., Stephan, K. E. & Friston, K. J. The mismatch negativity: A review of underlying mechanisms. *Clin. Neurophysiol.* **120**, 453–463 (2009).
25. Lieder, F., Stephan, K. E., Daunizeau, J., Garrido, M. I. & Friston, K. J. A neurocomputational model of the mismatch negativity. *PLoS Comput. Biol.* <https://doi.org/10.1371/annotation/ca4c3cdf-9573-4a93-9542-3a62cddb8396> (2013).
26. Winkler, I. Predictive models in auditory stream segregation. *Int. J. Psychophysiol.* **77**, 215–215. <https://doi.org/10.1016/j.ijpsycho.2010.06.294> (2010).
27. Winkler, I. Interpreting the mismatch negativity. *J. Psychophysiol.* **21**, 147–163. <https://doi.org/10.1027/0269-8803.21.34.147> (2007).
28. Baldeweg, T. ERP repetition effect and mismatch negativity generation: A predictive coding perspective. *J. Psychophysiol.* **21**, 13513–13519 (2007).
29. Naatanen, R. *Attention and Brain Function*. (Lawrence Erlbaum Associates, 1992).
30. Naatanen, R. & Alho, K. Mismatch negativity—The measure for central sound representation accuracy. *Audiol. Neurootol.* **2**, 341–353 (1997).
31. Naatanen, R., Gaillard, A. W. & Mantysalo, S. Early selective-attention effect on evoked potential reinterpreted. *Acta Physiol. (Oxf)* **42**, 313–329 (1978).
32. Kushnerenko, E. V., Van den Bergh, B. R. H. & Winkler, I. Separating acoustic deviance from novelty during the first year of life: A review of event-related potential evidence. *Front. Psychol.* <https://doi.org/10.3389/fpsyg.2013.00595> (2013).
33. Choudhury, N. & Benasich, A. A. Maturation of auditory evoked potentials from 6 to 48 months: Prediction to 3 and 4 year language and cognitive abilities. *Clin. Neurophysiol.* **122**, 320–338. <https://doi.org/10.1016/j.clinph.2010.05.035> (2011).
34. Kushnerenko, E. *et al.* Processing acoustic change and novelty in newborn infants. *Eur. J. Neurosci.* **26**, 265–274. <https://doi.org/10.1111/j.1460-9568.2007.05628.x> (2007).
35. Mittag, M., Takegata, R. & Winkler, I. Transitional probabilities are prioritized over stimulus/pattern probabilities in auditory deviance detection: Memory basis for predictive sound processing. *J. Neurosci.* **36**, 9572–9579. <https://doi.org/10.1523/jneurosci.1041-16.2016> (2016).
36. Tsogli, V., Jentschke, S., Daikoku, T. & Koelsch, S. When the statistical MMN meets the physical MMN. *Sci. Rep.* **9**, 5563. <https://doi.org/10.1038/s41598-019-42066-4> (2019).
37. Haden, G. P., Nemeth, R., Torok, M. & Winkler, I. Predictive processing of pitch trends in newborn infants. *Brain Res.* **1626**, 14–20. <https://doi.org/10.1016/j.brainres.2015.02.048> (2015).
38. Winkler, I. & Czigler, I. Evidence from auditory and visual event-related potential (ERP) studies of deviance detection (MMN and vMMN) linking predictive coding theories and perceptual object representations. *Int. J. Psychophysiol.* **83**, 132–143. <https://doi.org/10.1016/j.ijpsycho.2011.10.001> (2012).
39. Basirat, A., Dehaene, S. & Dehaene-Lambertz, G. A hierarchy of cortical responses to sequence violations in three-month-old infants. *Cognition* **132**, 137–150. <https://doi.org/10.1016/j.cognition.2014.03.013> (2014).
40. Sussman, E. S. A new view on the MMN and attention debate the role of context in processing auditory events. *J. Psychophysiol.* **21**, 164–175 (2007).
41. Todd, J., Frost, J. D., Yeark, M. & Paton, B. Context is everything: How context shapes modulations of responses to unattended sound. *Hear. Res.* **399**, 107975. <https://doi.org/10.1016/j.heares.2020.107975> (2021).
42. Fitzgerald, K. *et al.* Hierarchical learning of statistical regularities over multiple timescales of sound sequence processing: A dynamic causal modelling study. *J. Cognit. Neurosci.* **33**, 1159 (2021).
43. Fitzgerald, K., Provost, A. & Todd, J. First-impression bias effects on mismatch negativity to auditory spatial deviants. *Psychophysiology*. <https://doi.org/10.1111/psyp.13013> (2018).
44. Todd, J. *et al.* Mismatch negativity (MMN) to pitch change is susceptible to order-dependent bias. *Front. Neurosci.* <https://doi.org/10.3389/fnins.2014.00180> (2014).
45. Coles, M. G. H. & Rugg, M. D. *Electrophysiology of Mind: Event-Related Brain Potentials and Cognition*, edited by M. D. Rugg and M. G. H. Coles. (Oxford University Press, 1995).
46. Novitski, N., Huottilainen, M., Tervaniemi, M., Naatanen, R. & Fellman, V. Neonatal frequency discrimination in 250–4000-Hz range: Electrophysiological evidence. *Clin. Neurophysiol.* **118**, 412–419. <https://doi.org/10.1016/j.clinph.2006.10.008> (2007).
47. Naatanen, R., Paavilainen, P., Alho, K., Reinikainen, K. & Sams, M. Do event-related potentials reveal the mechanism of the auditory sensory memory in the human brain? *Neurosci. Lett.* **98**, 217–221. [https://doi.org/10.1016/0304-3940\(89\)90513-2](https://doi.org/10.1016/0304-3940(89)90513-2) (1989).
48. Csepe, V., Karmos, G. & Molnar, M. Evoked potential correlates of stimulus deviance during wakefulness and sleep in cat—animal model of mismatch negativity. *Electroencephalogr. Clin. Neurophysiol.* **66**, 571–578 (1987).
49. Frost, J. D., Winkler, I., Provost, A. & Todd, J. Surprising sequential effects on MMN. *Biol. Psychol.* **116**, 47–56. <https://doi.org/10.1016/j.biopsycho.2015.10.005> (2016).
50. Todd, J., Provost, A., Whitson, L. R., Cooper, G. & Heathcote, A. Not so primitive: context-sensitive meta-learning about unattended sound sequences. *J. Neurophysiol.* **109**, 99–105. <https://doi.org/10.1152/jn.00581.2012> (2013).
51. Garrido, M. I. *et al.* The functional anatomy of the MMN: a DCM study of the roving paradigm. *Neuroimage* **42**, 936–944 (2008).
52. Carbajal, G. V. & Malmierca, M. S. The neuronal basis of predictive coding along the auditory pathway: From the subcortical roots to cortical deviance detection. *Trends Hear.* **22**, 2331216518784822–2331216518784822. <https://doi.org/10.1177/2331216518784822> (2018).
53. May, P. J. & Tiitinen, H. Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology* **47**, 66–122. <https://doi.org/10.1111/j.1469-8986.2009.00856.x> (2010).
54. Pienkowski, M. & Eggermont, J. J. Effects of adaptation on spectrotemporal receptive fields in primary auditory cortex. *NeuroReport* **20**, 1198–1203. <https://doi.org/10.1097/WNR.0b013e32832f812c> (2009).
55. Khalighinejad, B., Herrero, J. L., Mehta, A. D. & Mesgarani, N. Adaptation of the human auditory cortex to changing background noise. *Nat. Commun.* **10**, 2509. <https://doi.org/10.1038/s41467-019-10611-4> (2019).
56. Zwislocki, J. J. Temporal summation of loudness: An analysis. *J. Acoust. Soc. Am.* **46**, 431–441 (1969).
57. Lantings, C. P., Briley, P. M., Sumner, C. J. & Krumbholz, K. Mechanisms of adaptation in human auditory cortex. *J. Neurophysiol.* **110**, 973–983. <https://doi.org/10.1152/jn.00547.2012> (2013).
58. Naatanen, R., Tervaniemi, M., Sussman, E., Paavilainen, P. & Winkler, I. “Primitive intelligence” in the auditory cortex. *Trends Neurosci.* **24**, 283–288. [https://doi.org/10.1016/s0166-2236\(00\)01790-2](https://doi.org/10.1016/s0166-2236(00)01790-2) (2001).
59. Naatanen, R., Jacobsen, T. & Winkler, I. Memory-based or afferent processes in mismatch negativity (MMN): A review of the evidence. *Psychophysiology* **42**, 25–32. <https://doi.org/10.1111/j.1469-8986.2005.00256.x> (2005).
60. Dehaene-Lambertz, G. The human infant brain: A neural architecture able to learn language. *Psychon. Bull. Rev.* **24**, 48–55. <https://doi.org/10.3758/s13423-016-1156-9> (2017).
61. Feldman, H. & Friston, K. J. Attention, uncertainty, and free-energy. *Front. Hum. Neurosci.* **4**, 215. <https://doi.org/10.3389/fnhum.2010.00215> (2010).

62. Bejjanki, V. R. & Aslin, R. N. Optimal, resource-rational or sub-optimal? Insights from cognitive development. *Behav. Brain Sci.* **43**, e4. <https://doi.org/10.1017/S0140525X19001614> (2020).
63. Mathys, C., Daunizeau, J., Friston, K. J. & Stephan, K. E. A bayesian foundation for individual learning under uncertainty. *Front. Hum. Neurosci.* **5**, 39. <https://doi.org/10.3389/fnhum.2011.00039> (2011).
64. Emberson, L. L., Boldin, A. M., Robertson, C. E., Cannon, G. & Aslin, R. N. Expectation affects neural repetition suppression in infancy. *Dev. Cogn. Neurosci.* **37**, 100597. <https://doi.org/10.1016/j.dcn.2018.11.001> (2019).
65. Kushnerenko, E. *et al.* Central auditory processing of durational changes in complex speech patterns by newborns: An event-related potential study. *Dev. Neuropsychol.* **19**, 83–97 (2001).
66. Kushnerenko, E., Ceponiene, R., Fellman, V., Huotilainen, M. & Winkler, I. Event-related potential correlates of sound duration: Similar pattern from birth to adulthood. *NeuroReport* **12**, 3777–3781. <https://doi.org/10.1097/00001756-200112040-00035> (2001).
67. Trainor, L. J. & Cirelli, L. Rhythm and interpersonal synchrony in early social development. *Ann. N. Y. Acad. Sci.* **1337**, 45–52. <https://doi.org/10.1111/nyas.12649> (2015).
68. Arenillas-Alcón, S., Costa-Faidella, J., Ribas-Prats, T., Gómez-Roig, M. D. & Escera, C. Neural encoding of voice pitch and formant structure at birth as revealed by frequency-following responses. *Sci. Rep.* **11**, 6660. <https://doi.org/10.1038/s41598-021-85799-x> (2021).
69. Kostilainen, K. *et al.* Healthy full-term infants' brain responses to emotionally and linguistically relevant sounds using a multi-feature mismatch negativity (MMN) paradigm. *Neurosci. Lett.* **670**, 110–115. <https://doi.org/10.1016/j.neulet.2018.01.039> (2018).
70. Lopez-Calderon, J. & Luck, S. J. ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Front. Hum. Neurosci.* <https://doi.org/10.3389/fnhum.2014.00213> (2014).
71. Delorme, A. & Makeig, S. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134**, 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009> (2004).
72. Todd, J. *et al.* What controls gain in gain control? Mismatch negativity (MMN), priors and system biases. *Brain Topogr.* **27**, 578–589. <https://doi.org/10.1007/s10548-013-0344-4> (2014).
73. JASP (0.13.1) v. 0.13.1 (Computer Software, 2021).

## Acknowledgements

We thank Gavin Cooper for his assistance in programming an earlier version of these experiments and Judit Roschéne Farkas for collecting the data. This research was supported by funds provided by the Australian Government through the Australian Research Council (DP200102346) and by the Hungarian National Research, Development and Innovation Office (NKFI K 115385 awarded to IW), by the New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation (ÚNKP-21-5-BME-364 for GPH) and the Bolyai Research Grant from the Hungarian Academy of Sciences (BO/00523/21/2 for GPH).

## Author contributions

J.T. project design, data analysis and interpretation, manuscript preparation. G.H. data collection, data processing, manuscript preparation. I.W. project design, interpretation, manuscript preparation.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-09994-0>.

**Correspondence** and requests for materials should be addressed to J.T.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022