



UvA-DARE (Digital Academic Repository)

Newborn infants detect the beat in music

Winkler, I.; Háden, G.P.; Ladinig, O.; Sziller, I.; Honing, H.

Published in:

Proceedings of the National Academy of Sciences of the United States of America

DOI:

[10.1073/pnas.0809035106](https://doi.org/10.1073/pnas.0809035106)

[Link to publication](#)

Citation for published version (APA):

Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences of the United States of America*, 106(7), 2468-2471. <https://doi.org/10.1073/pnas.0809035106>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Newborn infants detect the beat in music

István Winkler^{a,b,1}, Gábor P. Háden^{a,c}, Olivia Ladinig^d, István Sziller^e, and Henkjan Honing^d

^aInstitute for Psychology, Hungarian Academy of Sciences, H-1394 Budapest, P.O. Box 398, Hungary; ^bInstitute of Psychology, University of Szeged, H-6722 Szeged, Petőfi S. sgt. 30-34, Hungary; ^cDepartment of Cognitive Science, Budapest University of Technology and Economics, H-1111 Budapest, Stoczek u. 2, Hungary; ^dMusic Cognition Group, Institute for Logic, Language and Computation, Universiteit van Amsterdam, Postbus 19268, 1000 GG Amsterdam; and ^eFirst Department of Obstetrics and Gynaecology, Semmelweis University, H-1088 Budapest, Baross u. 27, Hungary

Edited by Dale Purves, Duke University Medical Center, Durham, NC, and approved December 23, 2008 (received for review September 10, 2008)

To shed light on how humans can learn to understand music, we need to discover what the perceptual capabilities with which infants are born. Beat induction, the detection of a regular pulse in an auditory signal, is considered a fundamental human trait that, arguably, played a decisive role in the origin of music. Theorists are divided on the issue whether this ability is innate or learned. We show that newborn infants develop expectation for the onset of rhythmic cycles (the downbeat), even when it is not marked by stress or other distinguishing spectral features. Omitting the downbeat elicits brain activity associated with violating sensory expectations. Thus, our results strongly support the view that beat perception is innate.

event-related brain potentials (ERP) | neonates | rhythm

Music is present in some form in all human cultures. Sensitivity to various elements of music appears quite early on in infancy (1, 2–4), with understanding and appreciation of music emerging later through interaction between developing perceptual capabilities and cultural influence. Whereas there is already some information regarding spectral processing abilities of newborn infants (5, 6), little is known about how they process rhythm. The ability to sense beat (a regular pulse in an auditory signal; termed “tactus” in music theory; 7, 8) helps individuals to synchronize their movements with each other, such as necessary for dancing or producing music together. Although beat induction would be very difficult to assess in newborns using behavioral techniques, it is possible to measure electrical brain responses to sounds (auditory event related brain potentials, ERP), even in sleeping babies. In adults, infrequently violating some regular feature of a sound sequence evokes a discriminative brain response termed the mismatch negativity (MMN) (9, 10). Similar responses are elicited in newborns (11) by changes in primary sound features (e.g., the pitch of a repeating tone) and by violations of higher-order properties of the sequence, such as the direction of pitch change within tone pairs (ascending or descending) that are varying in the starting pitch (12). Newborns may even form crude sound categories while listening to a sound sequence (13): an additional discriminative ERP response is elicited when a harmonic tone is occasionally presented among noise segments or vice versa, suggesting a distinction between harmonic and complex sounds.

Neonates are also sensitive to temporal stimulus parameters [e.g., sound duration (14)] and to the higher-order temporal structure of a sound sequence [such as detecting periodical repetition of a sound pattern (15)]. Because the MMN is elicited by deviations from expectations (16), it is especially appropriate for testing beat induction. One of the most salient perceptual effects of beat induction is a strong expectation of an event at the first position of a musical unit, i.e., the “downbeat” (17). Therefore, occasionally omitting the downbeat in a sound sequence composed predominantly of strictly metrical (regular or “nonsyncopated”) variants of the same rhythm should elicit discriminative ERP responses if the infants extracted the beat of the sequence.

Results and Discussion of the Neonate Experiment

We presented 14 healthy sleeping neonates with sound sequences based on a typical 2-measure rock drum accompaniment pattern (S1) composed of snare, bass and hi-hat spanning 8 equally spaced (isochronous) positions (Fig. 1 *A* and *B*). Four further variants of the S1 pattern (S2–S4 and D) (Fig. 1 *C–F*) were created by omitting sounds in different positions. The omissions in S2, S3, and S4 do not break the rhythm when presented in random sequences of S1–S4 linked together, because the omitted sounds are at the lowest level of the metrical hierarchy of this rhythm (Fig. 1*A*) and, therefore, perceptually less salient (7). The 4 strictly metrical sound patterns (S1–S4; standard) made up the majority of the patterns in the sequences. Occasionally, the D pattern (Fig. 1*F*, deviant) was delivered in which the downbeat was omitted. Adults perceive the D pattern within the context of a sequence composed of S1–S4 as if the rhythm was broken, stumbled, or became strongly syncopated for a moment (18) (Sound File S1). A control sequence repeating the D pattern 100% of the time was also delivered (“deviant-control”).

Fig. 2 shows that the electrical brain responses elicited by the standard (only S2–S4; see *Methods*) and deviant-control patterns are very similar to each other, whereas the deviant stimulus response obtained in the main test sequence differs from them. The deviant minus deviant-control difference waveform has 2 negative waves peaking at 200 and 316 ms followed by a positive wave peaking at 428 ms. The difference between the deviant and the other 2 responses was significant in 40-ms-long latency ranges centered on the early negative and the late positive difference peaks (see Table 1 for the mean amplitudes) as shown by dependent measures ANOVAs with the factors of Stimulus (Standard vs. Deviant control vs. Deviant) × Electrode (C3 vs. Cz vs. C4). The Stimulus factor had a significant effect on both peaks (for the early negative waveform: $F[2,26] = 3.77, P < 0.05$, with the Greenhouse-Geisser correction factor $\epsilon = 0.85$ and the effect size $\eta^2 = 0.22$; for the positive waveform: $F[2,26] = 8.26, P < 0.01, \epsilon = 0.97, \eta^2 = 0.39$). No other main effects or interactions reached significance. Posthoc Tukey HSD pairwise comparisons showed significant differences between the deviant and the deviant-control responses in both latency ranges (with $df = 26, P < 0.05$ and 0.01 for the early negative and the late positive waveforms, respectively) and for the positive waveform, between the deviant and the standard response ($df = 26, P < 0.01$). No significant differences were found between the standard and the deviant control responses.

Author contributions: I.W., G.P.H., O.L., and H.H. designed research; G.P.H., O.L., and I.S. performed research; G.P.H. analyzed data; and I.W., I.S., and H.H. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

¹To whom correspondence should be addressed. E-mail: iwinkler@cogpsyphy.hu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0809035106/DCSupplemental.

© 2009 by The National Academy of Sciences of the USA

Table 1. Group averaged ($n = 14$) mean ERP amplitudes

| Group | Amplitude, μV | | | | | |
|-----------------|--------------------------|--------------|--------------|---------------------|--------------|--------------|
| | 180–220 ms interval | | | 408–448 ms interval | | |
| | C3 | Cz | C4 | C3 | Cz | C4 |
| Deviant | -0.50 (0.19) | -0.30 (0.22) | -0.41 (0.27) | 0.38 (0.17) | 0.67 (0.13) | 0.67 (0.27) |
| Deviant-control | 0.14 (0.13) | 0.06 (0.16) | 0.18 (0.25) | -0.10 (0.13) | -0.06 (0.16) | -0.18 (0.15) |
| Standard | -0.03 (0.09) | -0.06 (0.12) | -0.11 (0.09) | -0.02 (0.09) | -0.12 (0.12) | -0.16 (0.13) |

SEM values are shown in parentheses.

because this effect does not include the Stimulus factor, it is not the sign of a response distinguishing the deviant from the control response.

Thus, in adults, omission of the position-1 bass sound does not result in the elicitation of discriminative ERP responses in the absence of the rhythmic context. This result is compatible with those of previous studies showing that stimulus omissions (without a rhythmic structure) only elicit deviance-related responses at very fast presentation rates (<170-ms onset-to-onset intervals; see ref. 19). In our stimulus sequences, the omitted bass sound was separated by longer intervals from its neighbors. It should be noted that adult participants elicited the MMN discriminative ERP response, when they received the full stimulus sequence (all 3 instruments) as presented to newborn babies in the neonate experiment (18).

Discussion

These results demonstrate that violating the beat of a rhythmic sound sequence is detected by the brain of newborn infants. In support of this conclusion we showed that the sound pattern with omission at the downbeat position elicited discriminative electrical brain responses when it was delivered infrequently within the context of a strictly metrical rhythmic sequence. These responses were not elicited by the D pattern per se: When the D pattern was delivered in a repetitive sequence of its own, the brain response to it did not differ from that elicited by the standards. Neither were discriminative responses simply the result of detecting omissions in the rhythmic pattern. Omissions occurring in non-salient positions elicited no discriminative responses (see the response to the standards in Fig. 2). Furthermore, the discriminative ERP response elicited by the D pattern was not caused by separate representations formed for the 3 instruments: only omissions of the downbeat within the rhythmic context elicit this response.

So it appears that the capability of detecting beat in rhythmic sound sequences is already functional at birth. Several authors consider beat perception to be acquired during the first year of life (2–4), suggesting that being rocked to music by their parents is the most important factor. At the age of 7 months, infants have been shown to discriminate different rhythms (2, 3). These results were attributed to sensitivity to rhythmic variability, rather than to perceptual judgments making use of induced beat. Our results show that although learning by movement is probably important, the newborn auditory system is apparently sensitive to periodicities and develops expectations about when a new cycle should start (i.e., when the downbeat should occur). Therefore, although auditory perceptual learning starts already in the womb (20, 21), our results are fully compatible with the notion that the perception of beat is innate. In the current experiment, the beat was extracted from a sequence comprised of 4 different variants of the same rhythmic structure. This shows that newborns detect regular features in the acoustic environment despite variance (12) and they possess both spectral and temporal processing prerequisites of music perception.

Many questions arise as a result of this work. Does neonate sensitivity to important musical features mean that music carries some evolutionary advantage? If so, are the processing algorithms necessary for music perception part of our genetic heritage? One should note that the auditory processing capabilities found in newborn babies are also useful in auditory communication. The ability to extract melodic contours at different levels of absolute pitch is necessary to process prosody. Sensing higher-order periodicities of sound sequences is similarly needed for adapting to different speech rhythms e.g., finding the right time to reply or interject in a conversation (22). Temporal coordination is essential for effective communication. When it breaks down, understanding and cooperation between partners is seriously hampered. Therefore, even if beat induction

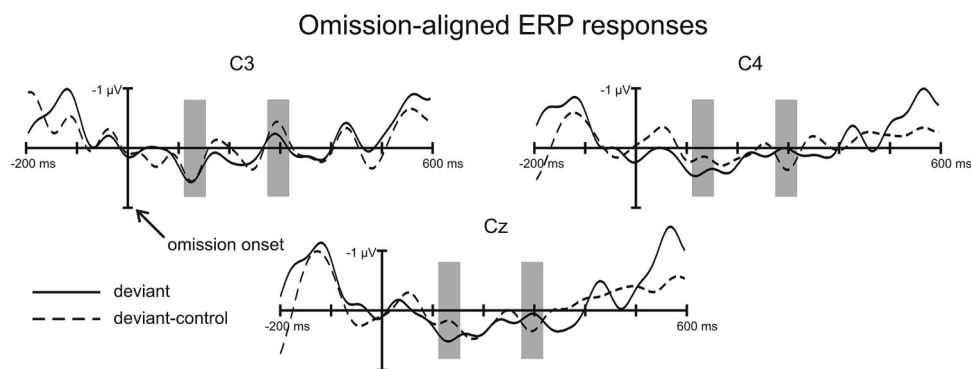


Fig. 3. Group averaged ($n = 13$) electrical brain responses elicited by the bass sound patterns in adults. Responses to deviant (D; solid line), and deviant-control patterns (D patterns appearing in the repetitive control stimulus block; dashed line) are aligned at the onset of the omitted bass sound (compared with the standard patterns: S1–S4) and shown from 200 ms before to 600 ms after the omission. Gray-shaded areas mark the time ranges in which amplitudes were measured.

