SCIENTIFIC REPORTS

OPEN

Enhanced auditory disembedding in an interleaved melody recognition test is associated with absolute pitch ability

Teresa Wenhart [1,2], Ye-Young Hwang[1] & Eckart Altenmüller[1,2]

Absolute pitch (AP) and autism have recently been associated with each other. Neurocognitive theories of autism could perhaps explain this co-occurrence. This study investigates whether AP musicians show an advantage in an interleaved melody recognition task (IMRT), an auditory version of an embedded figures test often investigated in autism with respect to the these theories. A total of N = 59 professional musicians (AP = 27) participated in the study. In each trial a probe melody was followed by an interleaved sequence. Participants had to indicate as to whether the probe melody was present in the interleaved sequence. Sensitivity index d′ and response bias c were calculated according to signal detection theory. Additionally, a pitch adjustment test measuring fine-graded differences in absolute pitch proficiency, the Autism-Spectrum-Quotient and a visual embedded figures test were conducted. AP outperformed relative pitch (RP) possessors on the overall IMRT and the fully interleaved condition. AP proficiency, visual disembedding and musicality predicted 39.2% of variance in the IMRT. No correlations were found between IMRT and autistic traits. Results are in line with a detailed-oriented cognitive style and enhanced perceptional functioning of AP musicians similar to that observed in autism.

One of the most complex functions of the auditory system lays in the ability to disentangle and extract distinct streams ("auditory streaming") of information out of the mixture of sounds that reach the ear (e.g. two interleaved and overlapped melodies played by different instruments in an orchestra or isolating a voice among noise in a crowded room). This is a key component of "auditory scene analysis"[1] with two different mechanisms: pre-attentive so-called "primitive" (bottom up) processes and, often but not always attentive, "schema-based" (top down) mechanisms[1]. Primitive processes sequentially or simultaneously separate the incoming sound mixture based on sensory information, i.e. according to regularities of harmonicity, or loudness, in relation to spatial orientation[1], or according to similarity in pitch[2–4], timbre[3], rhythm[5], or meter[6] and seem to be innate[7]. These factors correspond to the famous Gestalt principles[8] of similarity and proximity and have reached wide acceptance among visual perception researchers. Additionally, music or speech schemas learned via experience are used to attentively compare, extract and structure the auditory environment based on previous acquired knowledge[4,6]. Auditory streaming has been investigated using an interleaved melody recognition task (IMRT), first developed by W.J. Dowling (1973)[4] and later adapted and extended by Bey and McAdams (2002, 2003)[4,6]. In this paradigm a known[4,6] or unknown melody[2,3] is presented first, followed by an interleaved sequence, in which either the target or a modified melody is interleaved with distractor tones in the same pitch range.

The indirect measurement of auditory stream segregation with the use of interleaved melody recognition tests is very similar to the concept of disembedding in autism research and has indeed been described as an "auditory hidden figures test" by Dowling, Lung & Herrbold (1987)[6]. Autism-spectrum-conditions, which are characterized by social and communication difficulties, sensory abnormalities, narrow interests and problems with unexpected change[9] sometimes co-occur with special abilities[10–13] one of the more frequent of which is absolute pitch[12,14–16].

[1]Institute of Music Physiology and Musicians' Medicine, University for Music, Drama and Media, Hannover, Neues Haus 1, 30175, Hannover, Germany. [2]Center for Systems Neuroscience, Bünteweg 2, 30559, Hannover, Germany. Correspondence and requests for materials should be addressed to E.A. (email: eckart.altenmueller@hmtm-hannover.de)

Absolute pitch (AP) possessors have the unique (<1%)[17] ability to name or produce a musical tone without a reference[18,19]. AP occurs far more often in professional musicians (7–15%)[20–22] than in the general population, but it is still unclear why few people develop absolute pitch while most humans are relative pitch (RP) possessors. Relative pitch, in contrast is defined as the ability to make judgements on pitches in relation to other pitches (e.g. within musical intervals). It is very common and also often explicitly trained (e.g. naming of intervals) in professional musicians. However, AP and RP abilities might develop independently from each other and can also co-occur[23]. Among possible predictive factors for AP are an early onset of musical training especially before the age of 7[22,24] and related to this a critical period[25–29], genetic contributions[25], the type of musical education method[22] and ethnicity[20,21,30]. There is general agreement that absolute pitch is the result of an interaction between genes and environmental influences, which makes the ability a fascinating research topic for human cognitive neuroscience[29]. Interestingly, research has also shown increased autistic traits in absolute pitch possessors[31,32]. It has been suggested, that this joint occurrence of autism respectively autistic traits and absolute pitch ability might be explained by similar brain network structure and function[12,33–35], and a detail-oriented cognitive style[12,26].

A large amount of studies has investigated detail-oriented perception and cognition (e.g. disembedding and hierarchical letters) in autism[36–40] and are summarized in famous frameworks like the weak central coherence (WCC) theory[41,42], the enhanced perceptual functioning theory[43], the theory of veridical mapping[12] and the empathizing-systemizing theory[44,45].

People with autism-spectrum-conditions (ASC) show enhanced abilities to extract small visual or auditory figures out of a meaningful whole (disembedding) respectively a fully interleaved auditory sequence[30,46–49]. However, autistic people seem to fail to use pitch separation cues (i.e. separation of target and distractor melody by presentation on different average pitch) in auditory streaming experiments[30,47]. With respect to absolute pitch possessors evidence on perceptual and cognitive differences compared to non-AP possessors is restricted to visuo-spatial abilities[24] and investigation of auditory digit span[50]. It is not surprising that musicians in general exhibit an enhanced ability to extract musical streams out of complex interleaved melodies[3], as they belong to a population with far greater than average hearing abilities and knowledge of music theory[51,52]. However, it remains to be investigated, if AP-possessors differ in auditory streaming from relative pitch possessors. An investigation of auditory streaming in absolute pitch possessors could help to understand the above mentioned joint occurrence of absolute pitch ability and autism respectively higher autistic traits. Since the IMRT may serve as an auditory embedded figures test[6,30] it might be helpful to get insight if and to what extend absolute pitch possessors exhibit a similar auditory weak central coherence or detail-oriented perception as autistic people.

In summary, the present investigation addresses the questions (a) whether absolute pitch possessors exhibit an enhanced ability for disembedding in vision and audition (auditory streaming), (b) whether this is correlated to autistic traits in AP possessors and (c) whether having absolute pitch might explain different strategies and performance during auditory streaming experiments compared to relative pitch possessors. The present study aims to shed new light onto the discussion by investigating professional musicians with and without absolute pitch as they take part in an auditory streaming experiment created after Bey & McAdams[2,3] and a visual embedded figures test. We also determine their scores for autistic traits. We hypothesize that the ability of auditory stream segregation can be predicted by absolute pitch proficiency, autistic traits and musicality and that the IMRT reflects an auditory equivalent of disembedding in vision.

Methods

This study is part of a larger project consisting of various experiments investigating cognitive performance in audition and vision of absolute vs. relative pitch possessors. Therefore parts of the methods (sample description, general procedure, description of absolute pitch assessment, covariate measurements) are by their nature similar to previous publications of our lab (Wenhart, Bethlehem, Baron-Cohen & Altenmüller, 2019[53]; Wenhart & Altenmueller, 2019[54]).

Participants. Thirty-one AP musicians (16 female, mean age = 25.7 (SD = 9.7)) and 33 RP (relative pitch) musicians (15 female, mean age = 24.0 (SD = 7.13), see Table 1) - primarily students or professional musicians at the University for Music, Drama and Media, Hanover - were recruited via an online survey using UNIPARK software (https://www.unipark.com/). Recruitment was initially based on self-reports with respect to AP ability. Participants had the choice to select "I have absolute pitch", "I do not have absolute pitch" and "I do not know, whether I have absolute pitch". A first pitch identification test during the online survey (see general procedure below) was then used to evaluate the self-reports for group assignment. Four AP and two RP were amateur musicians. Non-native German speakers (4 AP) had the choice between a German and an English version of the experiments. All participants denied any history of psychiatric or neurological disorders, however, one AP reported taking Mirtazapine. The primary instruments played by the AP were piano (15), string instruments (9), woodwind instruments (3), voice (2), and brass instruments (2); for RP they were piano (13), string instruments (4), woodwind instruments (6), voice (3), brass instruments (3), bass (1), guitar (1), accordion (1), and drums (1). All AP but one were consistently right handed according to the Edinburgh Handedness Inventory[55] three RP were left-handed, two RP were ambidextrous. A total of five participants (4 AP, 1 RP; final sample: AP = 27, RP = 32) were excluded during analysis due to missing data (1 case), extreme values for reaction times (2 cases) or because they did not follow the instructions (2 cases). The study was approved according to Helsinki Declaration by the ethic committee of the Hanover Medical School (Approval no. 7372, committee's reference number: DE 9515). The methods were carried out in accordance with their guidelines and regulations. All participants gave written informed consent and received financial compensation for their participation.

General procedure. The overall project consisted of three parts: one online survey and two appointments in the lab of the Institute for Music Physiology and Musicians' Medicine, Hanover. The online survey was used for

	AP (n = 27)			RP (n = 32)			
	Mean	SD	Range	Mean	SD	Range	t-test (two-tailed)
Age	25.7	9.7	17–58	24.0	7.13	17–57	t(46.9) = −0.756; p = 0.453
SPM-IQ	114.67	12.59	74.5–132.25	114.2	13.3	86.5–134.5	t(56.2) = −0.138; p = 0.891
ZVT-IQ	123	12.47	101.5–145	120.5	13.89	97–143.5	t(56.7) = −0.728; p = 0.470
Hours main instrument	11551.27	9540.95	1642.5–39785	14107.82	17263.78	1606–77617.25	t(49.7) = 0.437; p = 0.664
AMMA total	65.59	6.04	53–78	63.28	7.14	46–76	t(57) = −1.348; p = 0.183
AMMA tonal	32.37	3.67	25–39	30.31	3.86	22–37	**t(56.1) = −2.096; p = 0.041**
AMMA rhythmic	33.19	2.69	28–39	31.94	3.26	23 37	t(57) = −1.610; p = 0.113
MSI (total)	209.04	14.79	182–234	211.12	15.24	185–246	t(55.8) = 0.533; p = 0.596
PIS	28.92	6.25	15–36	5.44	4.33	0–21*	**t(43.1) = −16.25; p < 2.2e-16*****
AQ	20.11	6.23	10–36	16.88	5.52	6–27	**t(52.5) = −2.093; p = 0.041**
MAD	42.16	39.02	9.8–200.57	297.77	87.33	91.04–467.52	**t(44.4) = 14.890; p < 2.2e-16*****
SDfoM	51.57	47.27	7.41–235.69	330.04	124.72	134.37–811.73	**t(41.0) = 11.675; p = 1.272e-14*****
Starting age	5.56	2.67	2–17	7.19	2.22	3–12	**t(55.8) = 2.527; p = 0.015**

Table 1. Participants' characteristics. Age, nonverbal IQ (SPM), information processing capacity (ZVT), musical training (total hours during life span on main instrument), musicality (AMMA; MSI) and online pitch identification screening (PIS) for each group; *Two RP reported not having absolute pitch but reached a screening score of 13 respectively 21. Because of this and their weak performance in the pitch adjustment test, the subjects were assigned to the RP group; Significant group differences (uncorrected) highlighted in bold. p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001 (Bonferroni-Holm corrected).

demographic questions, Autism Spectrum Quotient (AQ[56], German version by C.M. Freiburg, available online: https://www.autismresearchcentre.com/arc_tests), Musical Sophistication Index (GOLD-MSI[57] (question-naire)), an in-house questionnaire assessing musical education as well as total hours of musical training within the life span and a preliminary assessment of absolute pitch ability in order to allocate participants to groups (AP vs. RP) and evaluate their self-reports. The absolute pitch assessment was performed using an in-house pitch identification screening (PIS) consisting of 36 categorical (i.e. semitone distance between the tones based on equal-tempered tuning) sine waves in the range of three octaves between C4 (261.63 Hz) and B6 (1975.5 Hz). Participants exceeding 33% correct trials (>12/36 tones named correctly) were assigned to the AP group. The cutoff of 33% correct trials admittedly reflects a low threshold to select absolute pitch possessors compared to much more rigorous cutoffs (e.g. 50%, 80% or higher) in other studies[23,58,59]. However, it is still an ongoing debate, if and to what extend AP is distributed dichotomously or perhaps more gradually in professional musicians[23,58]. Furthermore, the use of relative strategies in pitch identification tests cannot be securely excluded[23]. Therefore we inspected scatterplots between PIS and Pitch adjustment test (PAT, see below) after the appointments in the lab and reassigned subjects that had differing performance on PIS and PAT based on their self-reports in the online survey. From our experience, professional musicians as investigated within this study usually know whether they themselves have absolute pitch. Due to this 2 RP's with initially high scores on the PIS (13 respectively 21 of 36 tones named correctly) were assigned to the RP group because they reported not to have AP and performed like RP possessors in the PAT. Likewise one AP possessor, who reported having technical issues with the presenta-tion of the sine waves in the online study and therefore a low PIS Score, was reassigned to the AP group. The low cutoff of 33% therefore is caused by this issue and might most likely not apply to other samples investigating AP with pitch identification tests. Tests on general intelligence (Raven's Standard Progressive Matrices (SPM))[60], information processing speed ("Zahlenverbindungstest" (ZVT))[61], a musical ability test (Advanced Measures of Music Audiation (AMMA))[62], a visual Embedded Figures Test (Group Embedded Figures Test (GEFT))[63] and all experiments were conducted in the lab (see Table 1 for general group differences regarding covariates).

Experiments and material. *Pitch adjustment test (PAT).* A pitch adjustment test (PAT) developed after Dohn *et al.*[64] was used to quantify fine-graded differences in absolute pitch proficiency. The test consisted of 108 target notes, presented as letters on a PC screen in semi-random order in 3 Blocks of 36 notes each (3*12 different notes per block) with individual breaks between the blocks. Participants' task was to adjust the frequency of a sine wave with random start frequency (220–880 Hz, 1 Hz steps) and to try to hit the target note (letter presented centrally on PC screen, e.g. "F#/Gb") within at most 15 seconds. Participants were allowed to choose their octave of preference. The tones were presented through sound isolating Shure 2-Way-In-ear Stereo Earphones (Shure

SE425-CL, Shure Distribution GmbH, Eppingen, Germany). Participants were explicitly asked to try to adjust each tone as precisely as possible without the use of any kind of reference and to confirm their answer with a button press on a Cedrus Response Pad (Response Pad RB-844, Cedrus Corporation, San Pedro, CA 90734, USA). If no button was hit, the final frequency after 15 seconds was taken and the experiment proceeded with the next trial. In both cases, the Inter Trial Interval (ITI) was set to 3000 ms. Online pitch modulation was performed by turning a USB-Controller (Griffin PowerMate NA16029, Griffin Technology, 6001 Oak Canyon, Irvine, CA, USA) and implemented in Python according to Dohn et al.[64]. Participants could choose between rough (10 cent, by scrolling the wheel) and fine tuning (1 cent, by pressing down and scrolling the wheel). The final or chosen frequencies of each participant were compared to the nearest target tone (<6 semitones/600cent). For each participant, mean absolute deviation (MAD[64] Eq. (1)) from target tone

$$MAD = \frac{\sum_{i=1}^{N_{adjustment}} |C_i|}{N_{adjustment}}$$

(1)

was calculated as the mean of the average absolute deviations c_i of the final frequencies to the target tone (referenced to a 440 Hz equal tempered tuning).

MAD reflects the pitch adjustment accuracy of the participants. The consistency of pitch adjustments (SDfoM, standard deviation from own mean), possibly reflecting the tuning of the pitch template[64], was then estimated by taking the standard deviation of the absolute deviations (Eq. (2)).

$$SDfoM = \sqrt{\frac{\sum_{i=1}^{N_{adjustment}} (C_i - MD)^2}{N_{adjustment} - 1}}$$

(2)

For regression analyses (see below), we performed a z-standardization of the MAD (Z_MAD, Eq. (3)) and SDfoM (Z_SDfoM, Eq. (4)) values relative to the mean and standard deviation of the non-AP-group, as originally proposed by Dohn et al.[64].

$$Z\_MAD_i = \frac{MAD_i - \mu(MAD)_{Non-AP}}{\sigma(MAD)_{Non-AP}}$$

(3)

$$Z\_SDfoM_i = \frac{SDfoM_i - \mu(SDfoM)_{Non-AP}}{\sigma(SDfoM)_{Non-AP}}$$

(4)

Interleaved Melody recognition test (IMRT). The experiment consisted of 16 training trials followed by 128 test trials presented in two blocks of 64 trials each with a break in between. Each trial was made up of a 6-tone probe melody (2400 ms, 400 ms per tone; ~74.5 bpm) followed by an interleaved sequence (see Fig. 1), in which the probe melody was embedded within a distractor melody (1st and then every 2nd tone belonging to the probe melody; 12 tones, 2400 ms, 200 ms per tone). In half of the trials, a modified version of the probe melody was embedded within the interleaved sequence. The general task for the participants was to press the right button on the Cedrus Response Pad (Response Pad RB-844, Cedrus Corporation, San Pedro, CA 90734, USA), if the probe melody was embedded inside the interleaved sequence, and the left button if not (required responses were counterbalanced; button colors (yellow and blue) were counterbalanced across participants). Each trial was preceded by the German or English word for "Attention!" lasting for 1000 ms, and probe melody and interleaved sequence were separated by an ISI (Inter-Stimulus-Interval) of 1000 ms. The ITI (Inter-Trial-Interval) was set to 2200 ms.

Probe melodies, modified probe melodies and distractor melodies were constructed according to Bey & McAdams (2002, 2003)[2,3] as follows:

The mean frequency of each probe melody was within a range from −3 to +2 semitones (ST) around the equal-tempered A_4 (440 Hz), i.e. from $F^\#_4$ to B_4. Modified melodies were then created by altering the 2nd and 4th or the 3rd and 5th note of the probe melody within a range of +−4 ST, which always also altered the melodic contour (see Appendix A of Bey & McAdams (2003)[3]). All melodies ranged from 5 to 11 ST and intervals within the melodies were between 1 and 8 ST. The 36 probe melodies and 36 modified melodies comprised 46 diatonic and 26 non-diatonic melodies. For each of the 36 melody pairs (probe and modified version) one distractor melody was composed (see Appendix B of Bey & McAdams (2003)[3]). All notes of distractor melodies were maximally 1–2 ST (alternating) above respectively below the two neighboring tones of the target melody to ensure, that the task could not be solved by the global contour of the sequence. In half of the melodies the first distractor note started below the target melody and the reverse was true for the other half. Finally, 8 interleaved sequences were built for each of the 36 melody pairs: distractor and either probe or modified melody were interleaved and – relative to the note closest to the mean frequency - separated by 0 (fully interleaved), 6 (separated by augmented fourth), 12 (one octave) or 24 ST (two octaves, see Fig. 2).

Two experimental versions (A, B) were created, with half of the participants receiving the trials related to probe melody 1–18 and the other half receiving the trails related to probe melodies 19–36. The 16 trials of the first two (Version A), respectively 19th and 20th target melodies (Version B) served as training trails. Order of trials was pseudo-randomized with the same probe melody never occurring twice in a row. The experiment was programmed and controlled via the Python toolbox PsychoPy[65,66]. Due to a technical error in the first version, one trial occasionally appeared twice during the experiment while another one was missed. The error occurred randomly with respect to probe vs. modified melody and semitone separation. To control for possible effects on
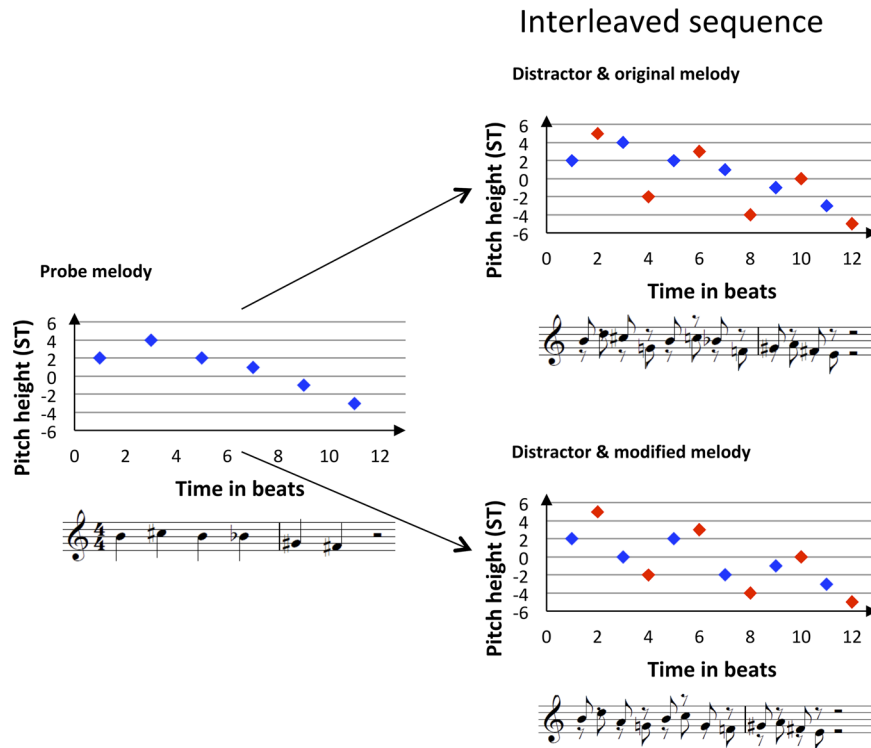
Figure 1. Graphical and musical notation of experimental material within the IMRT (interleaved melody recognition test). Each trial begins with a probe melody (A, blue) followed by an interleaved sequence with either the same probe melody (B, blue) or a modified melody (C, blue, here 2nd and 4th note modified compared to probe melody) interleaved with a distractor melody (red). Graphical notation of musical melodies in this example shows the pitch height of each note in ST relative to the A_4 (440 Hz). The distractor notes always encompassed the neighboring melody notes in terms of the pitch height.

the statistical tests, an equal number of participants in each group performed this imperfect version (N = 12; final percentage after exclusions: RP: 12/31, AP: 9/27), and all other participants afterwards performed the correct version. Calculation of Hit rates and False alarm rates was adapted accordingly, as described below. The total experiment without instructions lasted approximately 21.6 minutes.

Group embedded figures task (GEFT). The group embedded figures test (GEFT)[63], a paradigm often used to investigate cognitive style in autism[38,46,47], was chosen to additionally assess participants ability to find smaller shapes in drawings of a meaningful whole. The test is performed with pen and paper and consists of 18 trials in two blocks of 9. The participant has to find a geometric form in a large and more complex whole and mark it with a pencil. As we observed ceiling effects as to the total amount of correct trials in the GEFT during pilot experiments, we additionally measured the time participants needed to find each form so that we could better discriminate between participants' abilities at this task. Performance on the GEFT was therefore expressed as average time (over correct trials) needed to solve an item. As items increase in difficulty similarly to intelligence tests, a maximum time of 3 minutes per item was provided to avoid extreme reaction times on difficult items. Further, inspection of histograms was used to ensure that response times were not skewed within participants. Two participants have been excluded because they represented extreme values on individual or group level.

Statistical analysis. All statistical analyses were performed using the open-source statistical software package R (Version 3.5, https://www.r-project.org/). As melodies unfold over time, no reaction times were analyzed with respect to the IMRT. Instead, sensitivity index d' and response bias (decision criterion) c were calculated according to signal detection theory[67,68]. Signal detection theory is a suitable method for analyzing decision processes in recognition or discrimination tasks like in the IMRT, in which only two responses are possible (here: whether or not the probe melody is embedded). Therefore, responses fall into one of four categories: hit (correct identification of probe melody in the interleaved sequence), miss (miss of the probe melody in the interleaved sequence), false alarm (false "yes" response to modified melody in the interleaved sequence) or correct rejection (correct "no" response to modified melody in the interleaved sequence). According to Macmillan[69] the difference of the z-scores of proportions of hits (H in %) and false alarms (F in %) gives the sensitivity index d', which is independent of the decision criterion c.

$$d' = z(H) - z(F)$$

The latter is a measure of the tendency to say "yes" or "no", i.e. the response bias, and is calculated as:
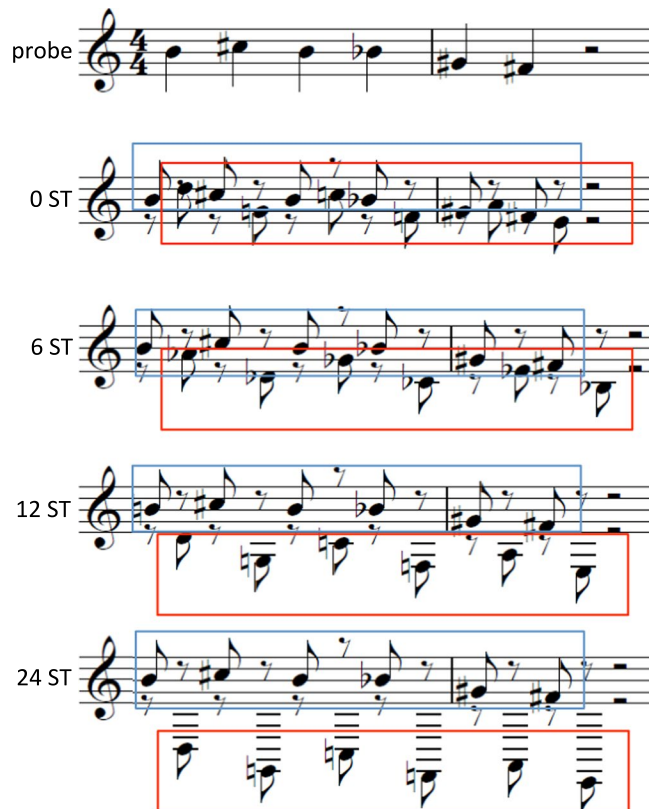
Figure 2. Interleaved sequences of IMRT. Probe melody respectively modified probe melody (here probe melody, blue) and distractor melody (red) were interleaved and four separation conditions created. The two melodies were either fully interleaved and their mean frequency separated by 0 semitones (0 ST) or the distractor melody was shifted by 6, 12 or 24 ST downwards. Therefore melodies were sequentially separated further apart, increasing the possibility of hearing two musical streams.

$$c = -\frac{1}{2}[z(H) + z(F)]$$

Therefore, positive values of c indicate a tendency towards "no" responses, while negative values indicate a tendency towards "yes" responses. As z-values for perfect proportions yielding $H = 1$ and $F = 0$ cannot be calculated (infinite values), Macmillan[69] recommends to reduce the proportion of $N_{correct}/N_{total}$ by subtracting 0.5 point from the number of correct trials within the affected condition: $H = (N_{correct} - 0.5)/N_{total}$, respectively to increase the proportion of false alarms by adding 0.5 points: $F = (N_{false} + 0.5)/N_{total}$. The IMRT comprised (in the errorless version) 16 trials per condition (8 conditions: 4 separation conditions x 2 melody conditions (probe/modified)); with this number of trials, some perfect scores were likely. As 21 participants had unequal trial numbers per condition (see Fig. A1, Supplementary Material A) hit rates and false alarm rates were always calculated according to the true number of trials of each participant per condition. However, since different numbers of trials affect calculated proportions and z-values, we performed additional calculations (e.g. using the number of trials of the participant with the most trials per condition as N_{total}) of proportions for hit rates and false alarm rates. These alternative analyses did not affect the direction or effect sizes of the reported statistical tests.

Ethics approval and consent to participate. The study was approved by the ethic committee of the Hanover Medical School (Approval no. 7372, committee's reference number: DE 9515). All participants gave written consent.

Results
Participants' characteristics. Multiple t-tests (Bonferroni-Holm-corrected) revealed no differences between AP and RP groups with respect to age, intelligence (SPM-IQ), information processing speed (ZVT-IQ), practice hours on instrument and musicality (AMMA total, Gold-MSI (see Table 1)). As expected, AP outperformed RP on both absolute pitch tests (PIS: t(43.1) = −16.25, ***$p < 2.2e$-16; MAD: t(44.4) = 14.890, ***$p < 2.2e$-16; SDfoM: t(41.0) = 11.675, ***$p = 1.272e$-14). Group differences for age of onset of musical training, the tonal subscale of AMMA and autistic traits (AQ) were only significant before Bonferroni-Holm-correction (see Table 1).
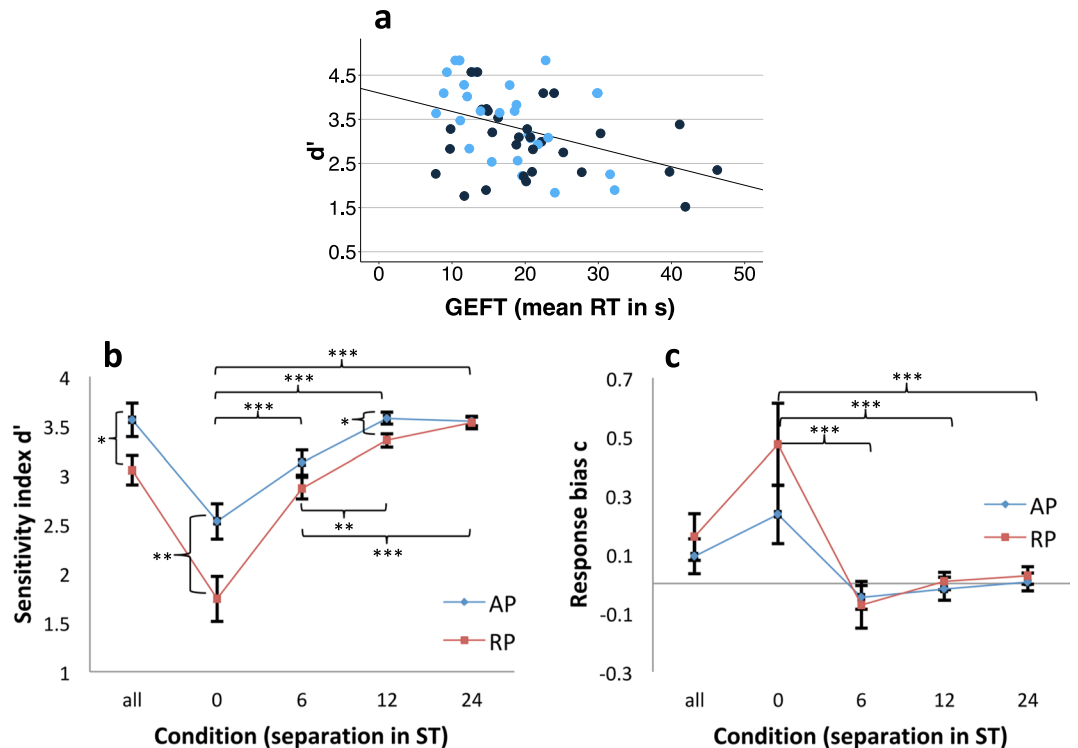
Figure 3. Relation of IMRT and visual embedded figures test (**a**) and signal detection measures by group (**b,c**). (**a**) Overall performance on IMRT (d') by mean time (in s) needed to solve an item on the visual embedded figures task (GEFT)[63]. Variables revealed a correlation of $r = -0.407$ (Pearson correlation; $p < 0.001$). Blue: IMRT version A, black: IMRT version B. (**b,c**) Results on sensitivity index d' (**b**) and response bias c (**c**)[69] by group (AP, absolute pitch; RP, relative pitch) regarding overall performance ("all") and performance within separation conditions (0, 6, 12, 24 semitone (ST) separations) on IMRT. Higher values of d' indicate better performance. Positive values of c indicate a tendency towards "no" responses, negative values a tendency towards "yes" responses. Bars represent standard errors. *p < 0.05. **p < 0.01. ***p < 0.001.

IMRT group differences.   Statistical tests were based on the signal detection measures mentioned above, with perceptual sensitivity (d')' and "response bias" (c) inspected separately. The two versions of IMRT (A, B), each containing half of the melodies, unintentionally differed in terms of d' scores. On average, version A yielded lower scores than version B (see Table B1 Supplementary Material B). Exploratory analyses revealed higher intelligence (SPM-IQ: $t(47.2) = -2.621$, $p = 0.012$), faster information processing speed (ZVT-IQ:: $t(57) = -1.870$, $p = 0.067$, marginally) and higher values on AMMA total score ($t(57) = -2.128$, $p = 0.038$) and subscales (tonal: $t(57) = -2.194$, $p = 0.032$; rhythmic: $t(57) = -2.516$, $p = 0.015$) for participants performing version B compared to version A of the test. In contrary, no group effects for version of IMRT were found for MSI, age, age of onset of musical training, hours of musical practice and performance on absolute pitch tests. Therefore, we argue that the effect of IMRT version on the performance of IMRT is most likely an artefact of accidental differences in intelligence and melody recognition (AMMA) between the participants that undertook the two versions and not due to differences in item difficulty. This is especially likely, as the musicality test AMMA is from its concept (items and recognition task) very similar to the IMRT experiment. A high correlation was therefore expected between them. Further, we have ensured that an equal number of AP and RP participants performed each version. No group differences of absolute pitch performance between the groups of the two versions confirms this procedure. Nevertheless, to correct for potential influences of version on regression models (see below), a dichotomous variable containing information of version performed per participant was included in the analyses as a covariate.

In contrast, different numbers of trials within frequency separation conditions (see Methods section) did not lead to significantly different results either in overall performance or for separate conditions (see Table A1, Supplementary Material A for further details), so number of trials was not included as a covariate.

First and importantly, the validity of IMRT (hence d' values) for measuring disembedding in audition was confirmed by Pearson correlation analysis ($r = -0.407$, $p = 0.0014$) with reaction times (RT in s) on the visual embedded figures test (GEFT)[63]. In total, roughly 15% of variance in the overall performance on IMRT was explained by the average time needed to solve the items of GEFT (linear model: intercept: 4.096***, GEFT: -0.042**; $F(1, 57) = 11.28$, $p = 0.0014$; $R^2 = 0.165$, $R^2_{adjusted} = 0.151$; see Fig. 3a).

Hence, IMRT was deemed to be an adequate test for investigating auditory disembedding. Therefore, a 2×4 repeated measurements ANOVA with separation (0, 6, 12, and 24) as within-subject factor and group (AP vs. RP) as between subject factor was run. Analyses revealed main effects of group ($F(1, 59) = 5.901$, $p = 0.0183$, $\eta^2_{partial} = 0.138$) and separation ($F(4, 228) = 76.474$, $p = 2e-16$, $\eta^2_{partial} = 0.573$) as well as a significant interaction ($F(4, 228) = 4.898$, $p = 0.000831$, $\eta^2_{partial} = 0.079$; see Fig. 3b).

	AP		RP		t-test		
	Mean	*SD*	*Mean*	*SD*	*t* (df)	*p*-value	Cohen's *d*
All	3.559	0.868	3.045	0.872	−2.262 (55.4)	<0.028*	−0.591
0 ST	2.527	0.923	1.739	1.279	−2,739 (55.8)	<0.008*	−0.697
6 ST	3.127	0.650	2.865	0.620	−1.565 (54.4)	0.123	−0.411
12 ST	3.575	0.340	3.350	0.380	−2.411 (56.8)	<0.019.	−0.624
24 ST	3.548	0.250	3.528	0.360	−0,250 (55.3)	0.804	−0.063

Table 2. IMRT group differences (N = 59). Means, standard deviation (SD) and statistical comparison (post hoc t-test, two-tailed) of IMRTperformance (perceptual sensitivity index d') by group (absolute pitch, AP, vs. relative pitch, RP) and per separation condition (in semitones, ST). p<010, *p < 0.05, **p < 0.01 (after Bonferroni-Holm correction of separation conditions).

Post hoc t-tests revealed significant differences (Bonferroni-Holm-corrected) between the 0-ST-condition and all other separation conditions (6-ST: $p = 2e\text{-}16$, 12-ST: $p = 2e\text{-}16$, 24-ST: $p = 6.3e\text{-}9$) as well as between the 6-ST-condition on the one hand and both the 12-ST- ($p = 0.00452$) and 24-ST-conditions ($p = 0.00058$), see Fig. 3b) on the other hand. Furthermore, absolute pitch possessors performed significantly better on the 0-ST and 12-ST conditions and when considering overall performance (see Table 2). We would also like to draw attention to a likely ceiling effect in condition 24-ST (see Figs 3b and 4e), the results of which should be interpreted with caution.

With respect to response bias (c) no significant group differences were found for overall performance nor for separation conditions (see Fig. 3c and Table B2, Supplementary Material B). However, a 2×4 repeated measures ANOVA with separation (0, 6, 12, 24) as within-subject factor and group (AP vs. RP) as between subject factor revealed a significant main effect of separation (F (4, 228) = 15.594, $p = 2.71e\text{-}11$, $\eta^2_{partial} = 0.215$). No main effect of group (F (1, 57) = 0.859, $p = 0.358$, $\eta^2_{partial} = 0.013$) and no interaction of group and separation (F (4, 228) = 1.452, $p = 0.218$, $\eta^2_{partial} = 0.025$) were found. Post hoc tests revealed a significant higher tendency to respond "no" on 0-ST-trials compared to all other conditions (6-ST: $p = 9.4e\text{-}6$; 12-ST: $p = 3.2e\text{-}5$; 24-ST: $p = 2.0e\text{-}7$, Bonferroni-Holm-corrected). This again confirms the impression that a greater separation between probe and distractor melody decreased the difficulty of the task and hence led to better perceptual sensitivity (d') and nearly no response bias (c, see Figs 3b,c and 4).

Linear models to predict perceptual sensitivity d'.   To investigate which musical and cognitive variables influence performance on IMRT, we performed multiple linear regressions separately for overall performance and separation conditions (0, 6, 12, 24 ST). In a first step, bivariate correlations between IMRT and variables of interest were calculated to get an overview of possibly predictive variables (see Table 3). As expected, we found a high correlation between both of our absolute pitch tests (r = −0.859, $p = 2.2e\text{-}16$). Interestingly, better absolute pitch performance in the labelling task was also associated with more autistic traits (PIS, AQ: r = 0.393, $p = 0.003$, see Table 1 and Fig. 5).

Apart from the already known correlation with GEFT (r = −0.406, $p = 0.0014$), IMRT (overall performance) was related to absolute pitch proficiency in both producing a tone (Z_MAD, standardized mean absolute deviation from target tone; r = −0.340, $p = 0.0085$) and naming a tone (PIS, Pitch identification screening; r = 0.283, $p = 0.0364$), age of onset of musical training (r = −0.289, $p = 0.0264$), information processing speed (r = 0.304, $p = 0.0199$) and marginally to musical sophistication index (MSI (total score); r = 0.238, $p = 0.077$). However, after Bonferroni-Holm-correction to adjust for multiple comparisons only the correlations with GEFT and Z_MAD were (marginally for Z_MAD) significant. Because of this and additional multicollinearity reasons age of onset and PIS were not included into the regression models. Both showed high correlations with absolute pitch performance (Z-MAD: r = 0.392, $p = 0.0021$; PIS: r = −0.361, $p = 0.00679$) respectively pitch adjustment (Z_MAD: r = −0.859, $p = 2.2e\text{-}16$). Therefore, a total of 5 variables (Z_MAD, MSI, GEFT, ZVT and version) were included into the 5 resulting regression models. Afterwards, variables with non-significant beta-weights were removed from the model leading to 5 reduced models (see Table 4). Bonferroni-Holm-corrections were performed according to the overall number of computed models. In general, IMRT performance was highly predicted by absolute pitch ability (Z-MAD) and GEFT. Performance in the 0-ST condition was additionally predicted by musical sophistication index (MSI) and the version of IMRT (A vs. B). Models of the 24-ST-condition did not reach significance and must be interpreted with caution due to ceiling effects (see Figs 3b,c and 4e). From Fig. 4 it can easily be seen that the four IMRT separation conditions (0, 6, 12, 24) decrease in variance (y-axis), indicative of decreasing task difficulty. Because at six semitones and greater separations the MSI is no longer predictive of IMRT (d'), this effect likely is due to our highly over-trained sample of professional musicians.

In summary, 39.2% ($R^2_{adjusted}$) of variance on overall performance on IMRT and 45.8% ($R^2_{adjusted}$) of variance in the fully interleaved 0-ST-condition were explained by absolute pitch ability (Z-MAD), musicality (MSI) and disembedding in vision (GEFT) (with IMRT version as a covariate). Even though $R^2_{adjusted}$ decreased for the easier conditions with a greater separation of probe and distractor melodies (6, 12), absolute pitch performance remained significant in both cases, while the other variables lost more of their influence (see Table 4). Figure 4 shows regression models for overall performance (Fig. 4a) and separation conditions (Fig. 4b–e) with regression line corresponding to a simple linear regression of absolute pitch ability (Z_MAD) on d' for simplicity reasons. Color and shape scales correspond to values on GEFT and MSI, respectively.
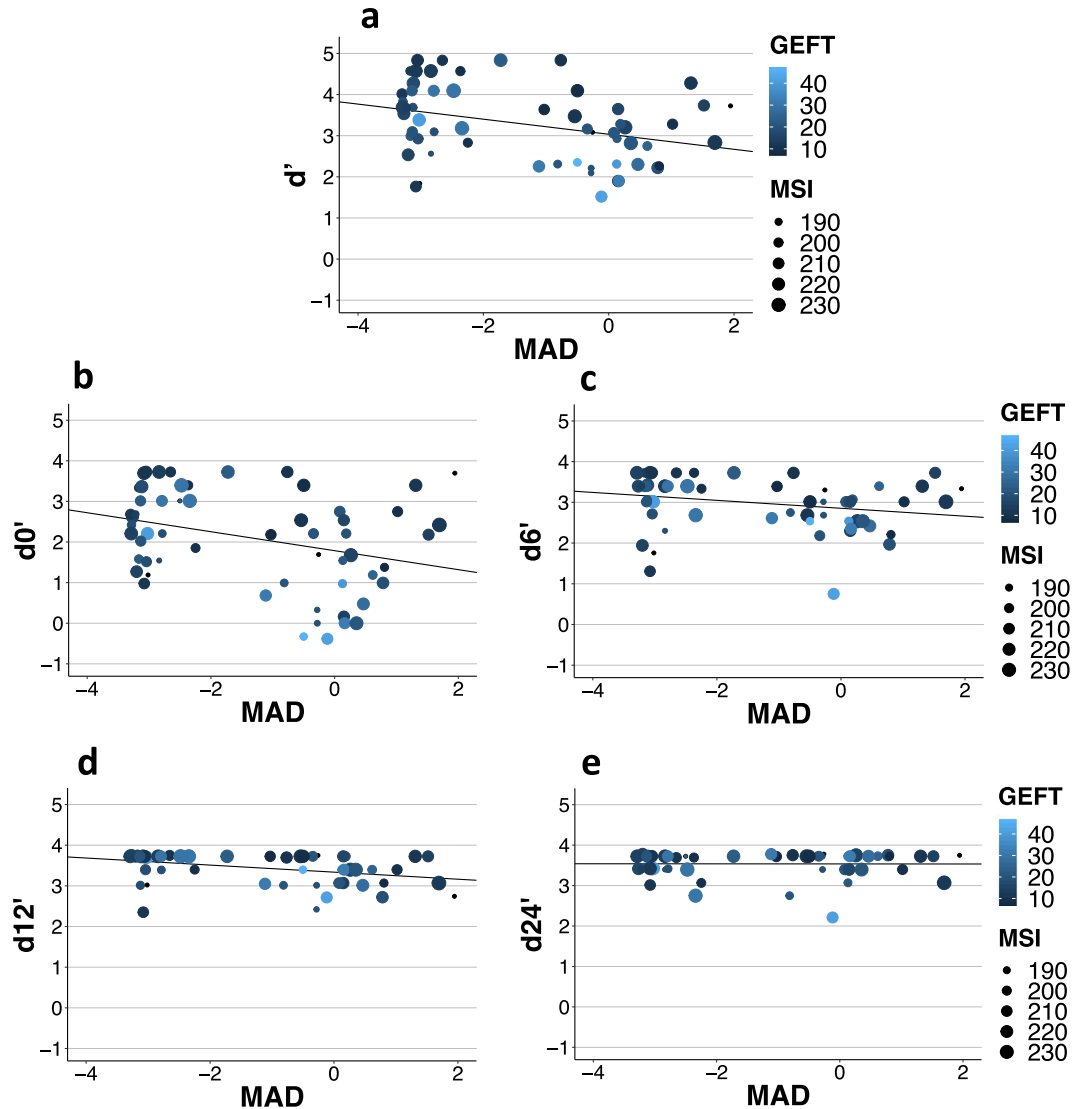
Figure 4. Influence of absolute pitch ability (Z_MAD), disembedding in vision (GEFT, time in s) and musicality (MSI) on perceptual sensitivity d'. Panel A corresponds to overall performance on IMRT, while panels B-E show the prediction of performance in different separation conditions (0, 6, 12, 24 semitone separation of probe and distractor melody). Color and shape scales correspond to disembedding in vision (GEFT, time in s) respectively MSI (score on questionnaire, higher values indicate greater musical sophistication). The regression line always takes the intercept and beta-weight of the simple linear regression of absolute pitch ability (Z_MAD, standardized to mean of RP group) on sensitivity index (d'-d24', higher values indicate better performance). Ceiling effects in the 12-ST- and especially 24-ST-condition (panel E) are clearly visible (see Table 4 for statistical values).

Discussion

For the first time we investigated auditory streaming ability[1], an ability closely related to the embedded figures concept from many autism frameworks (e.g. WCC theory[41,42]), in absolute and relative pitch professional musicians. By using an interleaved melody recognition test[2,3], we aimed to explore, if and to what extend the auditory perception of absolute pitch possessors might have a similarly auditory weak central coherence compared to relative pitch possessors[30]. This could serve as a common framework to explain the higher proportion of absolute pitch ability in autism and autistic traits in absolute pitch possessors.

Interestingly, general performance on the IMRT correlated with a visual embedded figures test (GEFT[63], mean time needed to solve an item), which confirmed our hypothesis that the interleaved melody recognition test serves as an auditory hidden/embedded figures test[6]. As embedded figures tests were traditionally used as evidence for a detail-oriented processing style in autism[36–40] and for the weak central coherence account (WCC)[41,42], superior performance of AP possessors in our IMRT might indicates a more detail-oriented auditory processing of absolute compared to relative pitch possessors. This is especially interesting, as several studys including the present investigation have reported (slightly) increased autistic traits in AP musicians[31,32]. However, performance on IMRT was not correlated with autistic traits in our study. Further, it must be made clear that direct

Correlation coefficient (Pearson)									
IMRT(d)	−0.406**	−0.340.	0.283	0.238	0.024	−0.289	0.200	0.304	−0.140
0.001**	**GEFT**	0.030	−0.134	−0.063	0.090	0.196	−0.612***	−0.289	−0.090
0.008.	0.821	**Z_MAD**	−0.859***	0.091	0.085	0.392*	−0.026	−0.124	−0.229
0.036	0.329	2.2e-16***	**PIS**	−0.093	−0.132	−0.361*	0.153	0.088	0.393*
0.077	0.644	0.503	0.499	**MSI (total)**	0.365*	−0.227	0.086	0.192	−0.136
0.855	0.496	0.521	0.337	0.006*	**Hours**	−0.042	−0.041	0.200	<0.001
0.026	0.137	0.002*	0.007*	0.092	0.753	**Start age**	−0.272	−0.059	−0.101
0.159	1.77e-6***	0.857	0.306	0.563	0.773	0.053	**SPM-IQ**	0.215	0.138
0.019	0.026	0.349	0.523	0.156	0.128	0.654	0.129	**ZVT-IQ**	−0.040
0.919	0.507	0.090	0.003*	0.317	0.999	0.460	0.348	0.769	**AQ**

Table 3. Bivariate correlations between variables of interest. Pearson correlations (upper triangle) and corresponding p-values (lower triangle) between variables of interest; IMRT: interleaved melody recognition test; GEFT: group embedded figures test; Z_MAD: z-standardized (to the mean and standard deviation of the non-AP population) mean absolute deviation from target tone; PIS: Pitch identification screening; hours: total hours of musical training on main instrument; start age: age at the onset of musical training; SPM-IQ: Raven progressive matrices; ZVT-IQ: information processing speed; AQ: Autism-Spectrum-Quotient; significant correlation coefficients are highlighted with stars. p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001 (Bonferroni-Holm corrected0).

		Predictors (β)					comparison of models			
d'	intercept	Z_MAD	MSI (total)	GEFT	ZVT	version	F (df)	p-value	R^2	$R^2_{adjusted}$
full	−0.286	−0.214***	0.013	−0.027*	0.007	0.447*	8.004 (5.5)	1.33e-5***	0.445	0.389
reduced	0.376	−0.219***	0.014*	−0.030**	—	0.470*	9.872 (4.5)	5.42e-6***	0.436	0.392
0										
full	−2.785	−0.282***	0.019*	−0.044**	−0.009	0.610*	10.25 (5.5)	8.42e-7***	0.506	0.457
reduced	−1.891	−0.289***	0.020*	−0.048**	—	0.641*	12.63 (4.5)	3.25e-7***	0.498	0.458
6										
full	2.738.	−0.117*	−0.001	−0.020*	0.005	0.119	3.192 (5.5)	0.014*	0.242	0.166
reduced	3.358***	−0.093	—	−0.026**	—	—	6.246 (2.6)	0.004**	0.182	0.153
12										
full	2.618**	−0.088**	0.005.	−0.004	−0.003	0.151	3.419(5.5)	0.010**	0.255	0.18
reduced	2.267**	−0.087**	0.005	—	—	0.155	5.497 (3.5)	0.002**	0.241	0.197
24										
full	3.384***	−0.0004	0.002	−0.003	0.005	0.135	1.477 (5.5)	0.214	0.129	0.041
reduced	—	—	—	—	—	—	—	—	—	—

Table 4. Comparison of models predicting IMRT. Parameters, significance (F-statistics) and comparison of different models for total performance on IMRT and performance on separation conditions (0, 6, 12, 24 semitones). Models are compared using R^2 and $R^2_{adjusted}$. Higher R^2 and more parsimonious models (fewer predictors) indicate superior models. Significance after Bonferroni-Holm correction: p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.01.

measures of streaming ability using alternating high and low frequency tones[70] may not necessarily have anything to do with the embedded figures concept. In the following we therefore stick to the terms "interleaved melody recognition test" and our acronym (IMRT) which we interpret to be a paradigm for both auditory streaming and the auditory embedded figures test.

General performance (sensitivity index d', signal detection theory) as well as performance at different levels of frequency separation between probe and distractor melodies exceeded chance level (d' = 0) in both groups. This effect can be explained by above average musical knowledge in this sample which could either result in increased availability of musical schemas (schema-based auditory streaming[2]) or be based on heightened ability to recognize melodies per se. However, as we did not include a control condition in which no distractor tones were present in the second sequence of each trial, we cannot finally decide upon this issue. Ceiling effects obtained for the high separation condition (24-ST) are perfectly in line with previous reports of ceiling effects with respect to large distances between probe and distractor melodies[2]. Responses were only slightly biased towards "no" responses overall and this tendency was significantly stronger for the most difficult 0-ST condition compared to all other conditions. As RP possessors always exhibited a more pronounced response bias (however not significant) this was seen as an indicator of uncertainty of responses, especially when the task was most challenging. Failure to
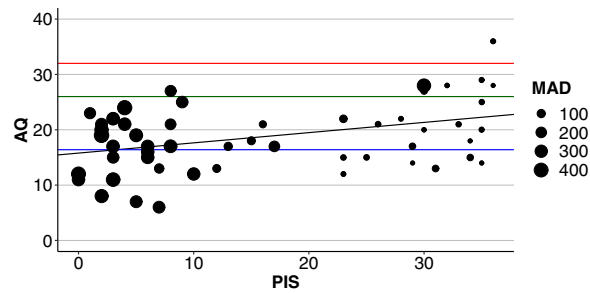
Figure 5. Relationship between absolute pitch performance and autistic traits. Higher scores on the online pitch identification screening (PIS, maximum = 36) are associated with more autistic traits on the Autism-Spectrum-Quotient (AQ, maximum = 50). Size of the dots in the scatterplot corresponds to performance in the pitch adjustment test (PAT: Mean absolute deviation (MAD) from target tone in cent (100 cent = 1 semitone)) with smaller values associated with better AP performance. It is visible by eye that except one case, all participants with good performance on PIS also succeed on the PAT. Red line: cutoff for diagnosis of autism-spectrum diagnosis[56]; green: boundary for critically/unusually high AQ[56]; blue: general healthy population mean (Baron-Cohen *et al.*, 2001)[43]. Black: regression line of PIS on AQ (r = 0.393, p = 0.003).

reach significance with respect to group comparisons on response bias might be due to inter-individual strategies to handle insecurity on melody judgements (tendency towards "no" vs. "yes" responses).

Furthermore, increasing the distance between probe and distractor melody (separation conditions) also increased performance of participants in both groups. Therefore and in accordance with other authors[2–4] pitch separation as a perceptual cue (i.e. separation of target and distractor melody by presentation on different average pitch helps to perceive two different melodies) seems to be of general importance for the processes of auditory streaming. However, in our study, we found an above chance level performance in the fully interleaved condition (0-ST). To our knowledge other authors have always found weak performance in healthy participants for this condition and have argued for the necessity of this pre-attentive pitch cue based mechanism to come into play. In their view, schema-based auditory streaming might only be possible in interaction with bottom-up sensory processes (such as using pitch cues) that structure the auditory signal[2–4,6]. In line with this, increased performance in our sample of professional musicians may have been caused by extensive musical experience which led to additional musical schema-based processes in the 0-ST condition for these participants. Musicians for example have daily exercise in extracting and focusing on melodies and fore- vs. background in orchestral or other ensemble music as well as in solo performances of harmonical instruments like piano, which can play several melodies at the same time. This explanation would be consistent with evidence that auditory experience, such as familiarity with particular voices[71] or words[72], can influence the auditory streaming of speech. This hypothesis is further strengthened by the marginal correlation of Musical Sophistication Index (MSI)[57] with overall performance and its contribution to some of the regression models. MSI is a general measure of musical sophistication aimed at the investigation of non-musicians which might explain the only small degree of correlation. It is possible that musicality tests which might be more suitable for trained musicians would have yielded higher correlations. However, as the interleaved melody recognition paradigm exhibits high similarity to famous tests of musicality (e.g. AMMA, Advanced measures of music audiation[62]) both in terms of stimuli and task (recognition of melodies) we did not include the AMMA into our regression analyses. The AMMA therefore only served to ensure equal musicality levels in both groups (AP and RP).

The even higher performance of absolute pitch possessors in the fully interleaved and all other separation conditions as well as in overall performance allows us to assume that additional processes related to the absolute perception and/or naming ability of pitches lead to an advantage of AP's in this test. We cannot rule out the possibility, that AP possessors might also use relative pitch strategies in this experiment. Despite we did not control for RP strategies, it can be assumed that AP possessors as well have a wide variety of relative pitch abilities[23]. The present group differences therefore indicate specific additional perceptual mechanisms of AP that enhance the performance of AP possessors. Perhaps stable pitch-label associations comprise an additional schema-based mechanism, by which probe melody and tones in the interleaved sequence can be compared and extracted. However, as AP's performance on IMRT also increases with the separation of probe and distractor melody, pitch-label associations might also increase auditory streaming ability due to an additional pre-attentive perceptual quality (pitch chroma) similar to e.g. tone-color perceptions of colored hearing synesthetes. Indeed, some authors have argued to explain absolute pitch ability as a form of synesthesia[12,73], which also leads to enhanced low level perception[13,74]. Interestingly, enhanced low-level perceptual functioning respectively a tendency to focus on details are two main elements of famous theories to explain autistic symptoms[41,43,44]. Furthermore, nearly all special abilities up to savant skills in autism share certain features like enhanced low-level perception, focus on details and mapping of two cognitive or perceptual structures according to their inherent elemental structure (Theory of Veridical Mapping[12]). The one-to-one mapping of pitches to pitch labels might therefore be similar to the one-to-one mappings of pitches to colors or letters to colors. This could lead to enhanced performances on experiments in the affected sensory modalities. This enhancement might in turn be caused by pre-attentive processes that make use of the additional sensory quality that other participants lack.

On top of that, autistic people can perform above chance with fully interleaved sequences while faring relatively poorly with respect to separation conditions[30]. However, in our sample, none of the performances on IMRT was explained (even in part) by autistic traits. Despite that, we cannot rule out the possibility that the lack of correlation with autistic symptoms is due to very heterogeneous factors that play a role in the development of absolute pitch (e.g. heritability, onset of musical training, ethnicity). A tendency towards higher autistic traits in AP, as was also present in our sample might form only one out of various different influences on the acquisition of AP. Therefore the lack of a correlation might be in part because the relation of AP and autistic traits is not true for all AP's.

In general, performance on IMRT (sensitivity index d'), i.e. the ability to recognize a melody in an interleaved sequence was highly ($R^2 = 15$–39.2%, smaller R^2 for easier conditions) dependent on absolute pitch proficiency as measured by a pitch adjustment test (PAT, developed after Dohn *et al*.[64]), visual disembedding (GEFT)[63] and in the more difficult conditions musical sophistication (MSI)[57]. We therefore conclude that the interleaved melody recognition test is an auditory streaming paradigm[5,6] which also serves as an auditory hidden figures test (as for the correlation with GEFT). Furthermore, performance on the IMRT is enhanced by musical training[2,3] and absolute pitch ability. The latter might be due to an additional sensory quality leading to enhanced low-level perceptual functioning in general. This study therefore to our knowledge is the first to show enhanced auditory disembedding in absolute pitch possessors. Nevertheless some constraints have to be mentioned: First, we did not include a control condition for melody recognition (see above) or a non-musical control group, second, due to a technical error the number of trials per condition and subject was not counterbalanced (see Methods section). However, additional statistical analyses indicated that results were not affected by this issue (see Methods section and Supplementary Material A). Third, with respect to the discussed similarities to autism, a subsample including autistic people (and a matched control group) would have been desirable.

To conclude we would like to draw attention to goals that future studies should address. Because of the partial, but nevertheless re-occurring, similarities between absolute pitch possessors and autism in terms of cognition, perception, personality traits and neurophysiology and anatomy it remains to be investigated whether and to what extend one is the cause or a side effect of the other, or which external factors lead to the coincidence. Such investigation would not only increase the knowledge about both phenomena but also help to understand fine-graded differences in human perception and its relation to other cognitive functions and personality traits.

Data Availability

The datasets generated and/or analysed during the current study are not publicly available due to specifications on data availablity within ethics approval. Data are however available from the corresponding author upon reasonable request and with permission of the ethics committee of the Hanover Medical School.

References

1. Bregman, A. S. Auditory scene analysis: The perceptual organization of sound. (MIT press, 1994).
2. Bey, C. & McAdams, S. Schema-based processing in auditory scene analysis. *Percept. Psychophys.* **64**, 844–854 (2002).
3. Bey, C. & McAdams, S. Postrecognition of interleaved melodies as an indirect measure of auditory stream formation. *J. Exp. Psychol. Hum. Percept. Perform.* **29**, 267–279 (2003).
4. Dowling, W. J. The perception of interleaved melodies. *Cognit. Psychol.* **5**, 322–337 (1973).
5. Devergie, A., Grimault, N., Tillmann, B. & Berthommier, F. Effect of rhythmic attention on the segregation of interleaved melodies. *J. Acoust. Soc. Am.* **128**, EL1–EL7 (2010).
6. Dowling, W. J., Lung, K. M.-T. & Herrbold, S. Aiming attention in pitch and time in the perception of interleaved melodies. *Percept. Psychophys.* **41**, 642–656 (1987).
7. McAdams, S. & Bertoncini, J. Organization and discrimination of repeating sound sequences by newborn infants. *J. Acoust. Soc. Am.* **102**, 2945–2953 (1997).
8. Wertheimer, M. Drei Abhandlungen Zur Gestalttheorie. *Ann. Philos. Philos. Krit.* **5**, 87–87 (1925).
9. Lai, M.-C., Lombardo, M. V., Chakrabarti, B. & Baron-Cohen, S. Subgrouping the Autism "Spectrum": Reflections on DSM-5. *PLoS Biol.* **11**, e1001544 (2013).
10. Bouvet, L. *et al*. Veridical mapping in savant abilities, absolute pitch, and synesthesia: an autism case study. *Front. Psychol.* **5** (2014).
11. Howlin, P., Goode, S., Hutton, J. & Rutter, M. Savant skills in autism: psychometric approaches and parental reports. *Philos. Trans. R. Soc. B Biol. Sci.* **364**, 1359–1367 (2009).
12. Mottron, L. *et al*. Veridical mapping in the development of exceptional autistic abilities. *Neurosci Biobehav Rev* **37** (2012).
13. Mottron, L., Dawson, M. & Soulieres, I. Enhanced perception in savant syndrome: patterns, structure and creativity. *Philos. Trans. R. Soc. B Biol. Sci.* **364**, 1385–1391 (2009).
14. Heaton, P. Pitch memory, labelling and disembedding in autism. *J. Child Psychol. Psychiatry* **44**, 543–551 (2003).
15. Heaton, P., Davis, R. E. & Happé, F. G. E. Research note: Exceptional absolute pitch perception for spoken words in an able adult with autism. *Neuropsychologia* **46**, 2095–2098 (2008).
16. Heaton, P., Hermelin, B. & Pring, L. Autism and Pitch Processing: A Precursor for Savant Musical Ability? Music Percept. *Interdiscip. J.* **15**, 291–305 (1998).
17. Profita, J., Bidder, T. G., Optiz, J. M. & Reynolds, J. F. Perfect pitch. *Am. J. Med. Genet.* **29**, 763–771 (1988).
18. Takeuchi, A. H. & Hulse, S. H. Absolute pitch. *Psychol. Bull.* **113**, 345–361 (1993).
19. Ward, W. D. Absolute pitch. In *The Psychology of Music* (ed. Deutsch, D.) 265–298 (Academic Press, 1999).
20. Deutsch, D., Dooley, K., Henthorn, T. & Head, B. Absolute pitch among students in an American music conservatory: Association with tone language fluency. *J. Acoust. Soc. Am.* **125**, 2398–2403 (2009).
21. Gregersen, P. K., Kowalsky, E., Kohn, N. & Marvin, E. W. Absolute Pitch: Prevalence, Ethnic Variation, and Estimation of the Genetic Component. *Am. J. Hum. Genet.* **65**, 911–913 (1999).
22. Gregersen, P. K., Kowalsky, E., Kohn, N. & Marvin, E. W. Early childhood music education and predisposition to absolute pitch: Teasing apart genes and environment. *Am. J. Med. Genet.* **98**, 280–282 (2001).
23. Wengenroth, M. *et al*. Increased Volume and Function of Right Auditory Cortex as a Marker for Absolute Pitch. *Cereb. Cortex* **24**, 1127–1137 (2014).
24. Costa-Giomi, E., Gilmour, R., Siddell, J. & Lefebvre, E. Absolute Pitch, Early Musical Instruction, and Spatial Abilities. *Ann. N. Y. Acad. Sci.* **930**, 394–396 (2006).
25. Baharloo, S., Johnston, P. A., Service, S. K., Gitschier, J. & Freimer, N. B. Absolute Pitch: An Approach for Identification of Genetic and Nongenetic Components. *Am. J. Hum. Genet.* **62**, 224–231 (1998).

26. Chin, C. S. The Development of Absolute Pitch: A Theory Concerning the Roles of Music Training at an Early Developmental Age and Individual Cognitive Style. *Psychol. Music* **31**, 155–171 (2003).
27. Russo, F. A., Windell, D. L. & Cuddy, L. L. Learning the 'Special Note': Evidence for a Critical Period for Absolute Pitch Acquisition. *Music Percept.* **21**, 119–127 (2003).
28. Schellenberg, E. G. & Trehub, S. E. Good Pitch Memory Is Widespread. *Psychol. Sci.* **14**, 262–266 (2003).
29. Zatorre, R. J. Absolute pitch: a model for understanding the influence of genes and development on neural and cognitive function. *Nat. Neurosci.* **6**, 692–695 (2003).
30. Bouvet, L., Mottron, L., Valdois, S. & Donnadieu, S. Auditory Stream Segregation in Autism Spectrum Disorder: Benefits and Downsides of Superior Perceptual Processes. *J. Autism Dev. Disord.* **46**, 1553–1561 (2016).
31. Brown, W. A. *et al.* Autism-Related Language, Personality, and Cognition in People with Absolute Pitch: Results of a Preliminary Study. *J. Autism Dev. Disord.* **33**, 163–167 (2003).
32. Dohn, A., Garza-Villarreal, E. A., Heaton, P. & Vuust, P. Do musicians with perfect pitch have more autism traits than musicians without perfect pitch? An empirical study. *PLoS One* **7** (2012).
33. Jäncke, L., Langer, N. & Hänggi, J. Diminished Whole-brain but Enhanced Peri-sylvian Connectivity in Absolute Pitch Musicians. *J. Cogn. Neurosci.* **24**, 1447–1461 (2012).
34. Loui, P., Li, H. C., Hohmann, A. & Schlaug, G. Enhanced Cortical Connectivity in Absolute Pitch Musicians: A Model for Local Hyperconnectivity. *J. Cogn. Neurosci.* **23**, 1015–1026 (2011).
35. Loui, P, Zamm, A. & Schlaug, G. Enhanced functional networks in absolute pitch. *NeuroImage* **63**, 632–640 (2012).
36. Grice, S. J. *et al.* Disordered visual processing and oscillatory brain activity in autism and Williams Syndrome. *Neuro Report* **12**, 2697 (2001).
37. Mottron, L., Burack, J. A., Iarocci, G., Belleville, S. & Enns, J. T. Locally oriented perception with intact global processing among adolescents with high-functioning autism: evidence from multiple paradigms. *J Child Psychol Psychiatry* **44** (2003).
38. Bölte, S., Holtmann, M., Poustka, F., Scheurich, A. & Schmidt, L. Gestalt Perception and Local-Global Processing in High-Functioning Autism. *J. Autism Dev. Disord.* **37**, 1493–1504 (2007).
39. Pring, L., Ryder, N., Crane, L. & Hermelin, B. Local and global processing in savant artists with autism. *Perception* **39**, 1094–1103 (2010).
40. Russell-Smith, S. N., Maybery, M. T., Bayliss, D. M. & Sng, A. A. Support for a link between the local processing bias and social deficits in autism: an investigation of embedded figures test performance in non-clinical individuals. *J Autism Dev Disord* **42** (2012).
41. Happé, F. & Frith, U. The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *J Autism Dev Disord* **36** (2006).
42. Happé, F. Autism: cognitive deficit or cognitive style? *Trends Cogn Sci* **3** (1999).
43. Mottron, L., Dawson, M., Soulières, I., Hubert, B. & Burack, J. Enhanced Perceptual Functioning in Autism: An Update, and Eight Principles of Autistic Perception. *J. Autism Dev. Disord.* **36**, 27–43 (2006).
44. Baron-Cohen, S. Two new theories of autism: hyper-systemising and assortative mating. *Arch. Dis. Child.* **91**, 2–5 (2005).
45. Baron-Cohen, S. Autism: The Empathizing-Systemizing (E-S) Theory. *Ann. N. Y. Acad. Sci.* **1156**, 68–80 (2009).
46. Jolliffe, T. & Baron-Cohen, S. Are people with autism and Asperger syndrome faster than normal on the Embedded Figures Test? *J Child Psychol Psychiatry* **38** (1997).
47. Lepistö, T. *et al.* Auditory stream segregation in children with Asperger syndrome. *Biol. Psychol.* **82**, 301–307 (2009).
48. Russell-Smith, S. N., Maybery, M. T. & Bayliss, D. M. Are the autism and positive schizotypy spectra diametrically opposed in local versus global processing? *J Autism Dev Disord* **40** (2010).
49. Spencer, M. D. *et al.* Atypical activation during the Embedded Figures Task as a functional magnetic resonance imaging endophenotype of autism. *Brain* **135**, 3469–3480 (2012).
50. Ziv, N. & Radin, S. Absolute and relative pitch: Global versus local processing of chords. *Adv. Cogn. Psychol.* **10**, 15–25 (2014).
51. McLachlan, N., Marco, D. & Wilson, S. Pitch and Plasticity: Insights from the Pitch Matching of Chords by Musicians with Absolute and Relative Pitch. *Brain Sci.* **3**, 1615–1634 (2013).
52. Mikutta, C. A., Maissen, G., Altorfer, A., Strik, W. & Koenig, T. Professional musicians listen differently to music. *Neuroscience* **268**, 102–111 (2014).
53. Wenhart, T., Bethlehem, R. A. I., Baron-Cohen, S. & Altenmüller, E. Autistic traits, resting-state connectivity, and absolute pitch in professional musicians: shared and distinct neural features. *Molecular Autism* **10**(1) (2019).
54. Wenhart, T. & Altenmüller, E. A Tendency Towards Details? Inconsistent Results on Auditory and Visual Local-To-Global Processing in Absolute Pitch Musicians. *Front. Psychol.* **10** (2019).
55. Oldfield, R. C. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia* **9**, 97–113 (1971).
56. Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J. & Clubley, E. The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Malesand Females, Scientists and Mathematicians. *J. Autism Dev. Disord.* **31**, 5–17 (2001).
57. Müllensiefen, D., Gingras, B., Musil, J. & Stewart, L. The Musicality of Non-Musicians: An Index for Assessing Musical Sophistication in the General Population. *PLOS ONE* **9**, e89642 (2014).
58. Athos, E. A. *et al.* Dichotomy and perceptual distortions in absolute pitch ability. *Proc. Natl. Acad. Sci.* **104**, 14795–14800 (2007).
59. Elmer, S., Rogenmoser, L., Kühnis, J. & Jäncke, L. Bridging the Gap between Perceptual and Cognitive Perspectives on Absolute Pitch. *J. Neurosci.* **35**, 366–371 (2015).
60. Raven, J., Raven, J. C. & Court, J. H. *Manual for Raven's Progressive Matrices and Vocabulary Tests.* Section 3: Standard Progressive Matrices: 2000 Edition, updated 2004. (Pearson Assessment, 2004).
61. Oswald, W. D. Zahlen-Verbindungs-Test (ZVT) - 3. überarbeitete und neu normerte Auflage. (Hogrefe, 2016).
62. Gordon, E. E. Manual for the advanced measures of music audiation. (GIA Publications, 1989).
63. Oltman, P. K., Raskin, E. & Karp, S. A. Group embedded figures test. (Consulting Psychologists Press, Inc., 1971).
64. Dohn, A., Garza-Villarreal, E. A., Ribe, L. R., Wallentin, M. & Vuust, P. Musical Activity Tunes Up Absolute Pitch Ability. *Music Percept. Interdiscip. J.* **31**, 359–371 (2014).
65. Peirce, J. W. PsychoPy—Psychophysics software in Python. *J. Neurosci. Methods* **162**, 8–13 (2007).
66. Peirce, J. W. Generating stimuli for neuroscience using PsychoPy. *Front. Neuroinformatics* **2** (2008).
67. Macmillan, N. A. & Creelman, C. D. Detection theory: A user's guide. (Cambridge University Press, 1991).
68. Rotello, C. M., Macmillan, N. A. & Reeder, J. A. Sum-Difference Theory of Remembering and Knowing: A Two-Dimensional Signal-Detection Model. *Psychol. Rev.* **111**, 588–616 (2004).
69. Macmillan, N. A. Signal detection theory. Stevens' Handb. *Exp. Psychol.* **4**, 43–90 (2002).
70. van Noorden, L. P. A. S. Temporal coherence in the perception of tone sequences. 3, (Institute for Perceptual Research Eindhoven, The Netherlands, 1975).
71. Johnsrude, I. S. *et al.* Swinging at a cocktail party: voice familiarity aids speech perception in the presence of a competing voice. *Psychol. Sci.* **24**, 1995–2004 (2013).
72. Billig, A. J., Davis, M. H., Deeks, J. M., Monstrey, J. & Carlyon, R. P. Lexical influences on auditory streaming. *Curr. Biol. CB* **23**, 1585–1589 (2013).
73. Loui, P., Zamm, A. & Schlaug, G. Absolute Pitch and Synesthesia: Two Sides of the Same Coin? Shared and Distinct Neural Substrates of Music Listening. ICMPC Proc. Ed. Catherine Stevens Al Int. *Conf. Music Percept. Cogn.* 618–623 (2012).
74. Banissy, M. J., Walsh, V. & Ward, J. Enhanced sensory perception in synaesthesia. *Exp. Brain Res.* **196**, 565–571 (2009).

Author Contributions

T.W. designed the study, collected, analysed and interpreted the data. E.A. contributed to the design of the study and interpretation of the data. Y.H. took part in data collection and analysis. All authors read, improved and approved the final manuscript.

Additional Information

Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-019-44297-x.

Competing Interests: The authors declare no competing interests.

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