

<https://doi.org/10.1038/s44271-024-00157-7>

# Psychoacoustic and Archeoacoustic nature of ancient Aztec skull whistles

**Sascha Frühholz<sup>1,2</sup>✉, Pablo Rodriguez<sup>1</sup>, Mathilde Bonard<sup>1</sup>, Florence Steiner<sup>1</sup> & Marine Bobin<sup>1</sup>**

Many ancient cultures used musical tools for social and ritual procedures, with the Aztec skull whistle being a unique exemplar from postclassic Mesoamerica. Skull whistles can produce softer hiss-like but also aversive and scream-like sounds that were potentially meaningful either for sacrificial practices, mythological symbolism, or intimidating warfare of the Aztecs. However, solid psychoacoustic evidence for any theory is missing, especially how human listeners cognitively and affectively respond to skull whistle sounds. Using psychoacoustic listening and classification experiments, we show that skull whistle sounds are predominantly perceived as aversive and scary and as having a hybrid natural-artificial origin. Skull whistle sounds attract mental attention by affectively mimicking other aversive and startling sounds produced by nature and technology. They were psychoacoustically classified as a hybrid mix of being voice- and scream-like but also originating from technical mechanisms. Using human neuroimaging, we furthermore found that skull whistle sounds received a specific decoding of the affective significance in the neural auditory system of human listeners, accompanied by higher-order auditory cognition and symbolic evaluations in fronto-insular-parietal brain systems. Skull whistles thus seem unique sound tools with specific psycho-affective effects on listeners, and Aztec communities might have capitalized on the scary and scream-like nature of skull whistles.

Various ancient sound tools and musical instruments have been discovered in archeological excavation sites throughout the world. Since Paleolithic times, humans created sophisticated musical instruments to produce sounds for various purposes ranging from imitating environmental and animal sounds up to aesthetic and symbolic manifestations<sup>1</sup>. These ancient musical instruments and tools often have specific natural and/or mythological associations in their cultural context, given their sound quality (sound iconography) as well as their construction and visual appearance (visual iconography).

We here experimentally investigated the archeological case of Aztec skull whistles, which can be dated back to 1250–1521 CE<sup>2,3</sup>. Skull whistles are uniquely crafted and made from clay with an approximate size of 3–5 cm. They have a particular sound production mechanism, which makes the Aztec skull whistles rather unique and unusual sound tools compared to historical and contemporary music instruments (Fig. 1), and they have not yet been classified according to the (revised) Hornbostel-Sachs classification of musical instruments<sup>4,5</sup>. Their configuration allows for the collision of various airstreams, which were exclusively developed in prehispanic Mesoamerica<sup>6</sup>. Skull whistles produce a specific non-linear and noisy sound (wind- or hiss-like sound)

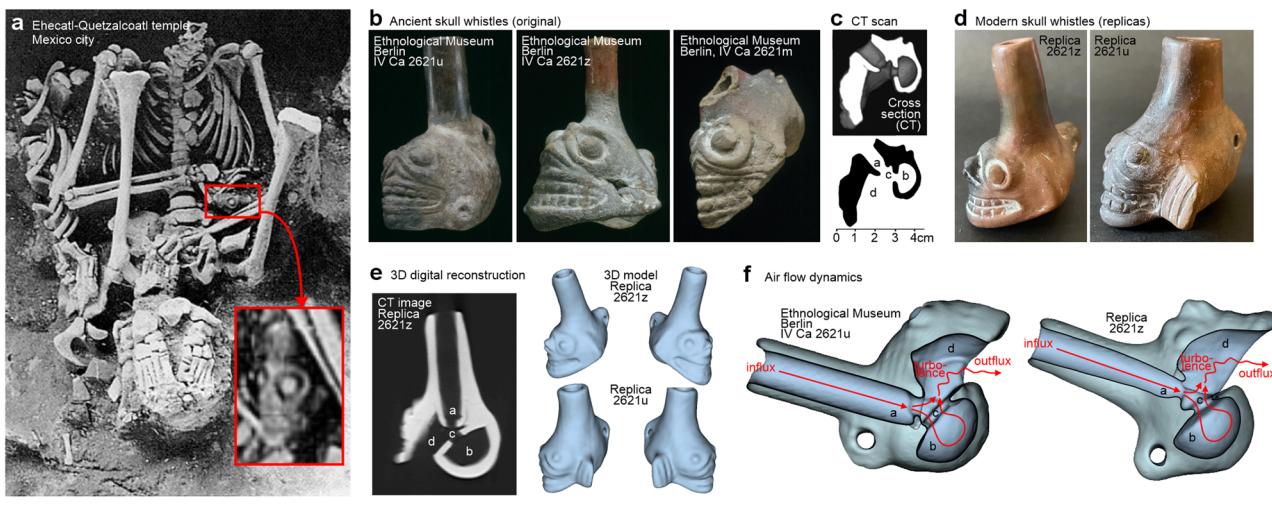
that can have a shrill, piercing, and scream-like sound quality when played with intensive air pressure<sup>2,7</sup>.

Skull whistles have not received much scientific attention so far<sup>2</sup> despite a very early reference to these instruments in the late 19th century<sup>8</sup>. However, such Aztec skull whistles are now frequently mentioned in popular contexts and have received a lot of media attention lately, given their potential spine-chilling sound similar to scary sound effects in horror movies and human screams<sup>9</sup>. In such popular contexts, they are often labeled as “death whistles”<sup>10</sup>, given a presumed but so far unsubstantiated association with collective warfare by Aztec communities to scare enemies (which we refer to as “warfare hypothesis”). We here use the more neutral label of “skull whistles” for these Aztec instruments given their general appearance as portraying a human or mystical skull that figures as the sound body of the whistle.

Only theories concerning the visual, acoustic, and contextual iconography of skull whistles exist so far, but there seems to be a multi-layered symbolism according to the cultural and mythological codex of the Aztecs<sup>7</sup>. This skull-like visual iconography might portray Mictlantecuhtli, the Aztec Lord of the Underworld, and might provide a link to Aztec sacrificial cults. The sound iconography reveals a possible association with Ehecatl, the

<sup>1</sup>Cognitive and Affective Neuroscience Unit, University of Zurich, Zurich, Switzerland. <sup>2</sup>Department of Psychology, University of Oslo, Oslo, Norway.

✉ e-mail: s.fruhholz@gmail.com



**Fig. 1 | Original exemplars and replicas of Aztec skull whistles.** **a** Human sacrifice with original skull whistle (small red box and enlarged rotated view in lower right) discovered 1987–89 at the Ehecatl-Quetzalcoatl temple in Mexico City, Mexico (burial 20; photo by Salvador Guillen Arroyo, Proyecto Tlatelolco 1987–2006, INAH Mexico). **b** Three original skull whistle exemplars from the collection of the Ethnological Museum in Berlin (Staatliche Museen zu Berlin, Germany; photo by Claudia Obrocki). **c** A computer-tomographically (CT) reconstructed cross-section of the right exemplar in (b) (IV Ca 2621m) showing the four major compartments ((**a**) tubular airduct with constricted passage, (**b**) hemispherical counterpressure

chamber, (**c**) collision chamber located between **a/b**, (**d**) bell). **d** Replicas of original skull whistles were built as a copy of original exemplars in shape and material. Exemplars “Replica 2621z” and “Replica 2621v” (manufactured by Arnd Adje Both and Osvaldo Padrón Pérez) as replicas of the original skull whistles with inventory numbers IV Ca 2621z and IV Ca 2621 v as shown in (b). **e** Digitalization and 3D reconstruction of the skull whistle replicas by using CT scans of the replicas. **f** 3D models of an original skull whistle (Ethnological Museum IV Ca 2621u) and the Replica 2621z demonstrate the air flow dynamics, construction similarity, and sound generation process.

Aztec God of the Wind, who traveled to the underworld to obtain the bones of previous world ages to create humankind (which we refer to as the “deity symbolism hypothesis”)<sup>2</sup>. Regarding the place and context of skull whistle discoveries, often involving ritual burial sites with human sacrifices, the skull whistles might have had a ritual and ceremonial iconography for the mythological descent into the Mictlan, the Aztec underworld, after sacrificial death. The fifth level of Mictlan is filled with deadly, razor-sharp, and piercing winds<sup>2</sup>, which again points to its potential sound iconography (which we refer to as the “ritual symbolism hypothesis”). For the latter, the rich Aztec sound imitation tradition also seems relevant<sup>7</sup>, with many Aztec instruments being designed to mimic environmental (wind, rain)<sup>7</sup>, animal (bird call, snake hiss)<sup>11</sup>, or human sounds (screams)<sup>9</sup>, with many of such instruments being used in ritual contexts<sup>11</sup>.

Thus, there is currently a diversity of potential archeoacoustic and iconographic theories concerning the skull whistles, which mainly relate to the acoustic and symbolic sound features of skull whistles. Experimentally exploring the perceptual effects of skull whistles could help to assess the psychoacoustic effects on human listeners as well as to discuss the potential archeoacoustic and pragmatic use of skull whistles by Aztec communities for affective (creating ritual wind-like atmospheres, scaring by scream-like quality) and/or symbolic purposes (mythical piercing winds, hostile conditions of the underworld). As for any human experimental study with a historical perspective, direct experiments cannot be performed with original Aztec humans. We however performed experiments with naïve European listeners from unbiased community samples. We specifically report data from various analytical approaches to first assess the physical and acoustic nature of skull whistles on the one hand, and second, we performed seven different psychoacoustic and neuroscientific laboratory experiments to assess the perceptual nature of skull whistles during their processing by humans.

The data reported here would provide knowledge in two major perspectives. First, from a psychoacoustic perspective, we assessed how modern humans respond to sounds produced by unique archeological artifacts that represent important historical sound tools. Second, from an archeoacoustic perspective and as outlined above, three major hypotheses exist so far concerning the cultural and practical meaning of skull whistles (warfare hypothesis, deity symbolism hypothesis, and ritual symbolism hypothesis).

Each hypothesis would likely predict differential effects of skull whistle sounds on human listeners, and we here took a precise experimental approach to obtain confirmatory evidence that is potentially in favor of certain hypotheses.

## Methods

### Original skull whistles

Most of the known original Aztec skull whistles (SW Orig) are preserved in archeological collections of museums and research laboratories around the world. In this paper, we refer to two original skull whistles stored in INAH facilities of the Tlatelolco archeological site, Mexico City, Mexico (Burial 7, Elements 2 and 4; Inventory Nrs. 10-262662, 10-263074). Further skull whistle artifacts are stored in the Ethnological Museum Berlin, Humboldt Forum, Germany (Inventory Nrs. IV Ca 2621a; 2621m; 2621z; 2621x; 2621 v) (Fig. 1b). Some noise whistles forming part of the Aztec fire snake incense ladles are stored in the Museo Nacional de Antropología and the Museo del Templo Mayor, Mexico City, Mexico (Find Nr. V, Inventory Nrs. 10-135830; 11-2883; 11-4004).

We obtained sound recordings (16 bit, 44.1 kHz) from the two abovementioned skull whistles excavated by Salvador Guillen Arroyo at the temple precinct of Tlatelolco, Mexico City. From these recordings, we extracted a total of 20 short sound files that were included in further analyses. We also obtained sound recordings of three noise whistles forming part of Aztec fire snake incense ladles discovered by Leopoldo Batres in Find No. V, which is an offering site of the temple precinct of Tenochtitlan. The fire snake incense ladles have the same skull whistle construction at their rear end, are played in the same way, and can produce the same sounds as skull whistles. From these recordings, we extracted a total of 15 short sound files from playing these exemplars.

To assess the outer and inner construction of all original skull whistles, we acquired computer tomography (CT) scans for some of the exemplars that are part of the collection of the Ethnological Museum in Berlin and again with the permission of the Berlin State Museums. High-resolution CT scans (Siemens CT Somaris, section resolution 0.160 mm<sup>3</sup>) were obtained for the exemplars with the inventory labels IV Ca 2621m and IV Ca 2621u (Fig. 1c). The CT scans enabled us to create high-quality 3D objects based on a digital surface reconstruction of the surface of the objects, and these CT

scans allowed us to virtually explore the outer and inner architecture of the original skull whistles.

### Replica skull whistles

Experimental archeology has the purpose of re-creating archeologically relevant objects to be able to perform scientific experiments with such objects. The manufacturing process of the object is equally important as the experiments that can be performed afterward. Here we took two examples (IV Ca 2621u, IV Ca 2621z; smb.museum/en/museums-institutions/ethnologisches-museum/home/) from the Ethnological Museum in Berlin as a blueprint to rebuild the skull whistles as replicas (SW repl, referred to as “Replica 2621u” and “Replica 2621z”) using clay material (Fig. 1d). We also acquired high-resolution CT scans (Siemens CT Somatom Definition AS+ scanner, section resolution 0.128 mm<sup>3</sup>) of these two replicas and created high-quality 3D objects from the replicas. This helped to confirm the similarity between the original and the replica skull whistles.

The category of replica skull whistles that were used in this study also included four skull whistles that were artisanal creations based on the general description of skull whistles in the literature. These additional artisanal skull whistles (SW art) were also made of clay and constructed along the general principles of skull whistles including the four major compartments. We thus used a total of six replica skull whistles, and we asked  $n = 5$  humans (3 male and 2 female participants, mean age 31.60y, SD 7.92) to produce the typical sounds of the skull whistles using three different levels of air pressure (SW low, SW med/medium, SW high) when playing the skull whistles. This resulted in a total of 270 sound recordings (mono 16 bit, 44.1 kHz) from these replicas and artisanal skull whistles.

Careful handling only allowed us to record sounds from original whistles as described above with a normal air pressure intensity, but sound recordings from replica skull whistles were acquired with low (SW low), medium (SW med), and high-intensity air pressure (SW high), as air pressure intensity can produce different sound qualities. Obtaining sound recordings with different air pressure levels ensured that the acoustic analysis was not biased towards a certain sound quality of the skull whistle.

### Large dataset of sounds

To compare the original skull whistles and replica skull whistles against the acoustic and perceptual profile of a large and diverse sample of human, animal, nature, and technical sounds, we collected sounds from various databases and our own recordings<sup>9,12–14</sup>. A broad collection of sounds was taken from the ESC-50 library<sup>15</sup>, which is a labeled collection of environmental audio recordings from many sound categories in the broad sound classes of animal sounds (ANI), natural soundscape, and water sounds (labeled as “NAT” sounds), human non-speech sounds (HUM), interior/domestic sounds (INT), and exterior urban noises (EXT). The human sounds were extended by neutral and emotional voice recordings from our own databases of nonverbal<sup>12,14</sup> and whispered vocalizations<sup>13</sup>, and with short speech and speech-like utterances from established databases<sup>16</sup>. Modern instrument sounds (INS) were taken from the Philharmonia database<sup>17</sup>, and short neutral and emotional music sounds (MUS) were again taken from our own database<sup>9</sup>. We also recorded sounds from ancient Aztec instruments (e.g., ancient flute, incense pipe, trumpet; labeled as “AZT” instruments) and more contemporary and popular Mexican flutes (MEX) that produce sounds to mimic animal sounds (e.g., eagle, jaguar) and human sounds (e.g., Llorona). These sounds were recorded in a steady state and in a modulated manner. Finally, we collected synthetic sounds (SYN) from freely available sound sources including electronic sounds produced by technical devices (e.g., computer system start-up sounds, pinball machines, printers). The full list of the 86 sound categories, including the skull whistle sounds and all other sounds used in this study can be found in Tab. S1. The total number of sounds was  $n = 2567$  with a mono 16 bit encoding and a 44.1 kHz sampling rate. Each sound file was cropped to 800 ms and all sounds were normalized to have the same RMS. The final sound pressure level was set to SPL 70 dBA for all sounds with a fade-in/out of 15 ms.

### Acoustic analysis of sounds

**Acoustic feature extraction.** None of the studies reported in the following using experimental data collections and parametric data analysis has been preregistered. To perform acoustic analyses of the sounds, we extracted acoustic features with the openSMILE Toolbox<sup>18</sup>. The feature configuration file applied (emobase2010.conf) extracted a total of 1582 acoustic features in the spectral and temporal domain including the arithmetic mean as well as the first and second derivative for some of the features<sup>19,20</sup>.

To quantify the amount of spectral and temporal information contained in each sound file<sup>9,12</sup>, we also quantified the modulation power spectrum (MPS) according to a previous description<sup>21</sup> by using the MPS toolbox for MATLAB (theunissen.berkeley.edu/Software.html). To obtain an MPS for each sound, we first converted the amplitude waveform to a log amplitude of its spectrogram obtained by using Gaussian windows and a log-frequency axis. The MPS results from the amplitude squared as a function of the Fourier pairs of the time (temporal modulation in Hz) and frequency axis (spectral modulation in cycles/octave) of the spectrogram. The low-pass filter boundaries of the modulation spectrum were set to 200 Hz for the temporal modulation rate and to 12 cycles/octave for the spectral modulation rate.

**Modulation power spectrum (MPS).** A statistical difference of the MPS from the 4 skull whistle sound categories (SW orig, SW high, SW med, SW low) compared with sounds from the other 9 major sound categories (*mus* music, *nat* nature, *ani* animal, *hum* human, *int* interior, *ext* exterior, *syn* synthetic, *mex* mexican flutes, *ins* instruments) was tested using a permutation approach ( $n = 2000$ ) by shuffling category labels, resulting in a log-transformed *p*-value map for the entire spectral and temporal range of the MPS in terms of the difference between categories.

**Representational similarity analysis (RSA).** Based on the pattern of the 1582 acoustic features, we performed a representational similarity analysis (RSA) according to a procedure and code described previously<sup>22</sup>. The RSA is based on a calculation of pairwise distances between pairs of sounds across all sounds and all 1582 acoustic features. For the RSA calculation, all acoustic features were z-transformed, and distance/similarity was calculated as Pearson correlation between acoustic patterns. This resulted in an acoustic similarity matrix with values defined as [1-correlation] and thus with a range of [0-2] for highest-to-lowest similarity, or lowest-to-highest dissimilarity, respectively. This RSA analysis was performed two times, first on all 2567 sounds for getting similarity estimation across all sounds, and a second RSA was performed on the 305 skull whistle sounds only to estimate the similarity within these sounds. The skull whistle sounds were divided in 20 categories, including one category of original Aztec skull whistle sounds (orig whis), one category of Aztec fire snake incense ladles (orig snake), two replica skull whistle categories (rep1, rep2), and four artisanal skull whistle sounds (art1, art2, art3, art4).

**Hierarchical clustering analysis (HCA).** We additionally performed a hierarchical clustering analysis (HCA) to generate dendograms of how sound categories cluster together based on their acoustic patterns. For this, we calculated the Euclidean distance between sounds based on their z-transformed values of acoustic features, and the linkage in clusters was determined as the weighted average distance. Nodes were grouped with a linkage cluster threshold of  $c < 2.0$ .

### Perceptual assessment and ratings of sounds

**Dimensional ratings.** All 2567 sounds were subjected to a perceptual rating analysis by  $n = 70$  human listeners (27 male and 43 female participants, mean age 25.01 y, SD 4.55). Across all studies reported here, the adjectives female/male refer to sex (not gender), and the participants provided the information via self-report. All participants in the studies described were recruited by public announcements and volunteered to

participate. Participants were reimbursed with CHF 15 per hour or with course credits for their participation. Participants gave informed and written consent for their participation following the ethical and data security guidelines of the University of Zürich (Switzerland). All experiments were approved by the cantonal ethics committee of the Swiss canton of Zürich (#2017-01086).

Since a perpetual rating of all sounds by one listener would have been extremely time-consuming and exhausting, each listener only rated a random selection of  $n = 221\text{--}433$  sounds. Listeners were asked to rate each sound along four different dimensions on a visual analog scale setup using 10-point Likert scales. Ratings were done on the arousal level elicited by the sounds (0 not arousing at all, 10 highly arousing), the urgency to respond to the sounds (0 no urgency at all, 10 high urgency), the naturalness of sounds (0 not natural at all, 10 highly natural), and the valence of sounds ( $-5$  highly negative, 0 neutral, 5 highly positive).

Ratings were obtained and summarized for the 4 skull whistle sound categories (SW orig, SW low, SW med, SW high) and for the 9 categories of other sounds (human, animal, nature, exterior, interior, music, instrument, Mexican, synthetic). Each sound was rated ten times by a random selection of ten out of the 70 participants. The mean of these ten ratings was calculated and plotted in the 4-dimensional space of ratings. Additionally, we calculated the distribution of ratings across all participants for the 13 sound categories. These distributions were used to calculate inter-rater consistencies (intraclass correlation coefficient, ICC) and we calculated the statistical differences in ratings across the sound categories using by fitting a linear mixed model (LME) to the data:

$$\text{rating} \sim 1 + \text{category} + (1|\text{participant}) \quad (1)$$

with the “rating” being each of the 4 rating scales separately, “category” being the 13 sound categories as a fixed effect, and “participant” as a random effect factor.

**Representational similarity analysis.** Based on the pattern of the 4 perceptual rating scales, we performed a presentational similarity analysis (RSA) similar to the pattern of acoustic features (see above). This RSA analysis was again performed two times, first on all 2567 sounds for getting similarity estimation across all sounds, and a second RSA was performed on the 305 skull whistle sounds only to estimate the similarity within these sounds. We again also performed a hierarchical clustering analysis (HCA) to generate dendograms of how sound categories cluster together based on their perceptual ratings and the nodes were grouped with a linkage cluster threshold of  $c < 2.0$ .

### Free choice labeling and categorical classification of sounds

**Free choice labeling.** To obtain an unbiased estimate about how human listeners would label skull whistle sounds and sounds of the other sound categories, we performed a free choice labeling experiment with  $n = 40$  human listeners (15 male and 25 female participants, mean age 24.60 y, SD 3.43; all native English speakers). We presented a balanced selection of 200 out of the total 2567 sounds to each participant; this selection was necessary to ensure that we could obtain labels from each participant within a reasonable duration of the experiment. We selected sounds from the major sound categories including 8 SW orig, 8 SW high, 8 SW med, 8 SW low sounds as well as 32 nature (8 thunder, 8 fire/wood cracking, 8 water, 8 wind sounds), 32 animal (8 bird, 8 reptile, 8 monkey, 8 other mammal sounds), 32 human (8 baby cries, 8 whispers, 8 affective voice, 8 vowel/syllable), 8 interior, 16 exterior (8 engine, 4 horn, 4 chainsaw), 8 synthetic/music, 8 Mexican flute, 24 instrument (8 brass, 8 string, 8 wood instrument sounds), and 8 white/pink noise sounds. We asked human listeners to provide two types of written labels in English for each of these sounds: first, participants were asked to provide a single substantive label (referred to as sounds “Labels”) for each sound that would best describe the object, mechanisms, or process that most likely produced the sound; second, participants were asked to provide a single

adjective (referred to as “Adjectives”) for each sound that would best describe the affective nature and human reaction to the sound.

Raw labels for each participant and each sound were pre-processed before any statistical analysis. This pre-processing followed several rules to make labels more consistent and unified for proper statistical handling<sup>23</sup>: white space, special characters, and numbers were removed; misspelling was corrected where unambiguous (otherwise replaced with “unidentified” label); plural was converted to singular; generic names were removed (e.g., “sounds”, “emotion”) and replaced with the “unidentified” label; synonyms were unified using Microsoft’s synonym checker; and words were unified to the most common word stem. In the first step of the analysis, we quantified the relative occurrence of labels for each sound calculated as the total number of occurrences multiplied by the percentage of participants that used this label; the resulting numbers were divided by 100 for better plotting and representation of the data. We used a binomial test ( $p < 0.05$ , FDR corrected) to determine if the relative occurrence numbers were significantly greater than would be expected by chance given the probability of labels/adjectives.

In the second step of the analysis, we aimed to reduce the diverse number of labels given during the free choice labeling task to a low and consistent number of labels that would represent more general sound and affective categories. We therefore presented all labels and adjectives to a group of 7 independent human raters (3 male and 4 female participants, mean age 33.57 y, SD 7.25; all native English speakers), and asked them to assign the original substantive labels to one of 9 general sound categories (human, animal, nature, exterior, interior, music, instrument, synthetic, unknown), and to assign the adjective to one of 10 emotional labels (anger, disgust, fear, anticipation, joy, sadness, surprise, trust, neutral, unknown) according to the emotion classification scheme proposed by Robert Plutchik<sup>24</sup>. Based on the categorical re-labeling of the original substantive labels and adjectives, we again determined the relative occurrence of these new labels the same way as was done for the original labels. We again used a binomial test ( $p < 0.05$ , FDR corrected) to determine if the relative occurrence numbers of labels/adjectives were significantly greater than would be expected by chance given the probability of labels/adjectives.

Based on these re-labeled data, we also performed a Correspondence Analysis (CA) to estimate the principal dimension that could represent data on a more general level. First, we used Pearson’s Chi-squared test to evaluate the level of dependency of row (sounds) and column elements (labels, adjectives) of the data matrix. Second, after performing the CA analysis, we retained 3 dimensions for the substantive labels (explained variance 61.31%) and 3 dimensions for the adjectives (explained variance 69.39%) to obtain a 3-dimensional representation of all sounds in a sound category and a sound affective space, respectively.

### Forced choice sound classification (3-alternative forced choice, 3AFC)

To assess how humans classify skull whistle and other sounds into three major sound categories as a basic process of sound processing and classification, we presented a selection of 72 sounds to human listeners in two separate experiments with separate samples of human participants. In the first experiment with 76 participants (24 male and 52 female participants, mean age 24.17 y, SD 4.46; four out of the original 79 datasets were not included because of non-recorded responses due to technical response device errors), we included sounds from 6 sound categories consisting of 12 SW orig, 12 animal, 12 nature, 12 exterior/interior, 12 human, and 12 music sounds. These sounds were presented twice in random order, resulting in a total of 144 trials. Using three buttons on a keyboard, participants listened to each sound and classified it afterward as belonging to one of three categories (animated, technical, environment). The category “animated” sound had to be chosen if participants believed that the sound was produced by a living organism, the category “technical” sound had to be chosen if the sound originated from a technical device or process, and the category “environment” had to be chosen if the sound originated from a non-living environmental object or process. The next trial started 1 s after a category had been chosen for the current trials. The second experiment was the same as the first experiment, but instead of

SW orig sounds we presented SW repl sounds here. A different sample of 58 participants took part in the second experiment (17 male and 41 female participants, mean age 25.60 y, SD 6.29; 3 of 61 datasets could not be included again due to technical response device errors). Data for the two experiments were analyzed separately, and classification data were entered into a repeated measures analysis of variance with the within-subject factors sounds category (6 levels) and classification category (3 levels).

**Dichotic listening experiment.** Sounds can have different meanings and affective importance for human listeners. The more meaningful and effectively engaging sounds are the more they attract attention and distract from other ongoing tasks. We tested the meaningfulness and importance of skull whistle and other sounds using a dichotic listening experiment. We again performed two different experiments here including the same 72 sounds as described in the section above, but also including two additional categories of noise sounds (pink noise) and silent trials (no sounds presented on the unattended ear, see below) as baseline conditions. In the first experiment, 47 human listeners (13 male and 34 female participants, mean age 23.91 y, SD 4.21) listened to sounds from original skull whistles and other sounds that were presented on one ear, while one of two simple sine wave tones was presented on the other ear. The two sine wave tones were a 350 Hz (low tone) and a 370 Hz sine wave tone (high tone). Participants were asked to attend to the ear where these tones were presented and decide if the tone was “high” or “low” using two buttons and their right-hand index and middle finger. The skull whistle and other sounds were presented on the unattended ear, and each of these sounds was presented two times on the left (1 trial with low tone, 1 trial with high tone) and two times on the right ear (1 trial with low tone, 1 trial with high tone). There was a total of 384 trials split across two runs. The first run was preceded by 16 training trials to familiarize participants with the task and the high/low tone. Within each run, there were 8 blocks of 24 trials with the attended ear and tone discrimination task switching with every block. Button assignment and laterality of the first attended ear were counter-balanced across participants. The second experiment included an independent sample of 47 human listeners (14 male and 33 female participants, mean age 25.02 y, SD 3.91), with the only difference being that we presented SW repl sounds instead of the SW orig sounds. Data for the two experiments were analyzed separately, and reaction time and classification data were entered into a repeated measures analysis of variance with the within-subject factors sounds category (8 levels) and attended ear (2 levels).

### Neural decoding of skull whistle sounds

**Experimental setup.** The experiment included 32 human participants (14 male and 18 female participants, mean age 26.00 y, SD 5.47). All participants had normal hearing and normal or corrected-to-normal vision. Exclusion criteria were hearing and visual impairments as well as psychiatric or neurological disorders in life history. Participants were invited to take part in a functional magnetic resonance imaging experiment (fMRI) to quantify brain activity during the processing of skull whistle and other sounds.

This experiment included the same 200 sounds that are described in the section for the free choice labeling experiment (see above). The experiment consisted of 4 different runs, and each run presented all 200 sounds in random order as single trials. Each run also contained a random selection of 20 trials, in which sounds of the previous trial were repeated. The participants’ task was to detect these sound repetitions and indicate the detection of a repeated sound by a button press with their right index finger. Participants were asked to listen attentively to the sounds, and the sound repetition task ensured that participants kept an attentive listening state. All repetition trials were excluded from all further analyses.

**Brain data acquisition and pre-processing.** Structural and functional brain data were recorded on a 3 T Philips Ingenia MR scanner by using a standard 32-channel head coil. A high-resolution structural image was

acquired by using a T1-weighted scan (301 contiguous 1.2 mm slices, repetition time [TR]/echo time [TE] = 1.96 s/3.71 ms, field of view [FOV] = 256 mm, in-plane resolution  $1 \times 1 \text{ mm}^2$ ). Functional whole-brain images were recorded by using a T2\*-weighted echo-planar pulse imaging (EPI) sequence (TR 1.6 s, TE 30 ms, flip angle [FA] 82°; in-plane resolution  $220 \times 114.2 \text{ mm}$ , voxel size  $2.75 \times 2.75 \times 3.5 \text{ mm}^3$ ; slice gap 0.6 mm) covering the whole brain.

Pre-processing and statistical analyses of functional images were performed using the Statistical Parametric Mapping software (SPM12, fil.ion.ucl.ac.uk/spm). Functional data were first manually realigned to the AC-PC axis, and functional images were then motion-corrected using a 6-parameter rigid-body transformation with realignment to the mean functional image. This was followed by a slice time correction of the slices acquired within a brain volume. Each participant’s anatomical T1 image was then co-registered to the mean functional brain image, followed by a segmentation of the T1 image for estimating normalization parameters using a geodesic shooting and Gauss-Newton optimization approach<sup>25</sup> for transformations into the standard space. Based on the estimated parameters, the anatomical and functional images were then normalized to the Montreal Neurological Institute (MNI) stereotactic space. Functional images were resampled into an isotropic  $2 \text{ mm}^3$  voxel size during the normalization procedure. All functional images were spatially smoothed with an 8 mm full-width half-maximum (FWHM) isotropic Gaussian kernel.

**GLM functional brain data analysis.** Functional brain data were then entered into a fixed-effects single-subject analysis, with a general linear model (GLM) design matrix containing 26 separate regressors for each of the 25 sound categories plus an additional regressor for all repetition trials. The 25 sound categories consisted of 4 skull whistle categories (SW orig, SW high, SW med, SW low), 4 human sound categories (nonverbal vocalizations, baby cries, vowel/syllable, whisper), 4 animal sound categories (monkeys, other mammals, snakes, birds), 4 categories for technical/synthetic sounds (alarm sounds, tools, vehicles/engines, synthetic), 4 categories for nature and environment sounds (wind, thunder, fire/cracking wood, water), 4 categories of instrument sounds (Mexican flutes, brass, wood, string), and one category of white/pink noise sounds. Each sound category contained 8 individual sounds. All sound trials were modeled with a stick function aligned to the onset of each stimulus, which was then convolved with a standard hemodynamic response function (HRF). The design matrix also included six motion correction parameters as regressors of no interest to account for signal artifacts due to head motion. Contrast images for each of the 25 main sound categories and from each participant were then taken to several separate random-effects factorial group-level analyses. Different directional contrasts on the group-level were performed between conditions and were thresholded at a combined voxel threshold of  $p < 0.05$  (FWE corrected) and a cluster extent threshold of  $k = 10$ .

**Functional brain connectivity analysis.** We used the Connectivity Toolbox (CONN toolbox)<sup>26</sup> to perform functional connectivity analyses following a standard procedure for functional brain data analysis. We computed seed-to-voxel analysis based on 10 regions of interest (ROI). The ROIs included three right-hemispheric frontal regions originating from the [SW > other sounds] contrast (MFC, MFG, INS) and an additional parietal region in IPS. The ROIs also included six bilateral regions in the auditory cortex (PTe, PPo, ST) resulting from the [Other sounds > SW] contrast. From all ten ROIs we extracted the time series of data in a 3 mm square around the peak voxel coordinate. From the time series data, spurious sources of noise were estimated and removed by using an automated denoising procedure, and the residual BOLD time series was band-pass filtered in the range 0.008–0.09 Hz to minimize the influence of physiological, head-motion, and other noise sources.

We performed a generalized psycho-physiological interaction (gPPI) analysis, computing the interaction between the seed BOLD time series and a condition-specific factor when predicting each voxel BOLD time series. In

contrast to standard PPI, gPPI allows the inclusion of the interaction factor of all task conditions simultaneously in the estimation model to better account for between-condition effects and influences. We included all 24 original sound conditions in a single gPPI model based on a bivariate correlation approach between seed and target regions. For the group-level analysis, we specified seed-to-voxel analysis for the right fronto-parietal seeds regions during the task conditions including the four skull whistle sounds. Two further seed-to-voxel analyses were set up for the left and right AC seed regions separately including the task conditions of all other 20 sound categories. The significance threshold was set to a voxel-level threshold of  $p = 0.005$  combined with an FWE-corrected cluster-level threshold of  $p = 0.05$ .

**MVPA and RSA analysis.** To identify brain regions that encode the acoustic and perceptual similarity/difference of skull whistle sounds to the sounds of the other categories, we correlated the results of an RSA analysis of the sounds used in this experiment with a voxel-wise indicator of functional brain activity differences between the sound categories. This analysis procedure included several steps.

In the first step, we repeated the RSA analysis as described above for the acoustical features of sounds (1582 acoustic features) and the perceptual rating pattern (4 rating scales), but we performed the analysis here only on the 200 sounds that have been used in this experiment. This resulted in RSM patterns that encode the acoustic similarity/difference of 25 sound categories as well as the perceptual similarity/difference of the 25 sound categories. We thus obtained a  $25 \times 25$  dissimilarity matrix based on the acoustic features and a  $25 \times 25$  dissimilarity matrix based on the perceptual rating patterns.

In the second step, we performed a multivoxel pattern analysis (MVPA) implemented as searchlight decoding analysis using The Decoding Toolbox (TDT)<sup>27</sup> for the purpose of an information-based brain mapping according to multivoxel activity patterns, which are assumed to differ between the experimental conditions. The searchlight analysis was performed on normalized but unsmoothed functional data. This analysis was performed on single-trial beta images resulting from an iterative GLM analysis as recommended for better sensitivity in event-related designs<sup>28</sup>. We obtained beta images for each trial by using a GLM with one regressor modeling a single trial and a second regressor modeling all remaining trials. This GLM modeling was repeated for each trial, including movement parameters as a regressor of no interest to account for false positive activity due to head movements. For each voxel, we defined a local sphere of 6 mm radius to investigate the local multivoxel pattern information in the single-trial beta images. This procedure was repeated for every pairwise comparison between the 25 sound category conditions. For each of these iterations, we trained a multivoxel support vector machine classifier using the LIBSVM package and implemented a leave-one-run-out cross-validation design. This procedure resulted in a brain map of local pairwise decoding accuracy across the 25 experimental conditions for each participant. This was represented as a voxel-wise  $25 \times 25$  MVPA dissimilarity matrix with the assumption that the decoding accuracy indicates how well conditions can be discriminated on a brain level (i.e., higher decoding accuracy means higher neural discrimination).

In the third step, we cross-correlated (Spearman correlation) the brain MVPA dissimilarity matrices separately with the acoustic RSM and the perceptual RSM on a voxel-by-voxel level. We restricted this correlation analysis to the matrix part that represented the similarity/dissimilarity measure for the skull whistle sounds as the target sounds of this experiment and similarity analysis. This restricted matrix was the  $4 \times 25$  matrix including the 4 categories of skull whistle sounds and the 25 overall sound categories. This resulted in voxel-wise and Fisher z-transformed correlation maps for each participant. The correlation maps were spatially smoothed with an 8 mm full-width half-maximum (FWHM) isotropic Gaussian kernel and were entered into a group-level GLM analysis (voxel threshold of  $p < 0.05$ , FWE corrected; cluster extent threshold of  $k = 10$ ) separately for the acoustic RSM and the perceptual rating RSM analysis.

To test the distinctiveness of the resulting correlation maps with the specific RSM matrices for the acoustic and perceptual properties, we performed two additional permutation-based analyses. We created  $n = 30$  random permutations of the acoustic and perceptual RSM matrices and entered them into the same cross-correlation approach as described above. The resulting brain correlation maps were averaged on the group level using a conjunction analysis (conjunction null hypothesis<sup>29</sup>) across the permutations to determine if random permutation would or would not produce similar significant correlation maps as for the main analysis (voxel threshold of  $p < 0.05$ , FWE corrected; cluster extent threshold of  $k = 10$ ). No significant results were found in this permutation-based analysis.

## Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Results

### Artisanal and digital reconstruction of Aztec skull whistles

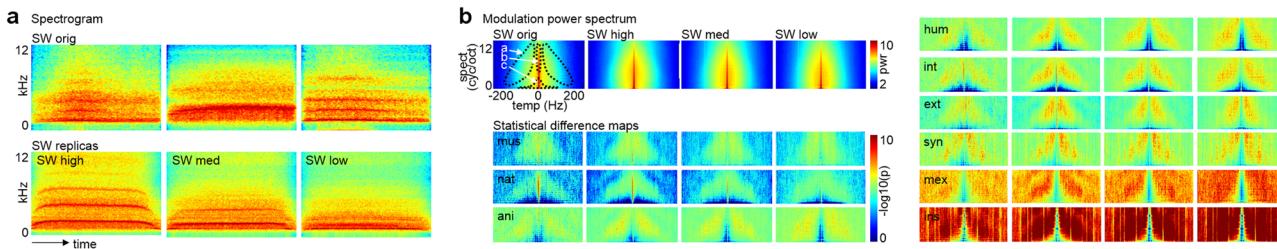
We first aimed to understand the physical structure and sound production mechanisms of original Aztec skull whistles. Skull whistles typically feature a tubular airduct with a constricted passage, a hemispherical counterpressure chamber, a collision chamber, and a bell cavity<sup>7</sup>. We acquired a high-resolution computer-tomographical (CT) image of one of the exemplars from the EMB to confirm this principle outer and inner architecture of SWs (Fig. 1c). Given this architecture, the driving acoustic mechanism is thought to rely on the Venturi effect generating a constant air aspiration, collision, and turbulence process<sup>2</sup>. At high playing intensities and air speeds, this leads to acoustic distortions and to a rough and piercing sound character that seems uniquely produced by the skull whistles.

Given this established structure of original skull whistles (SW orig), a specialized music archeologist (Arnd Adje Both<sup>2</sup>, together with the ceramist Osvaldo Padrón Pérez) planned and manufactured two replica skull whistles (SW repl) for the purpose of this study. These replica skull whistles are referred to as exemplars “Replica 2621 u” and “Replica 2621z” and they were copies of two specific exemplars from the EMB, using the same structure and material (Fig. 1d). This was done to acquire knowledge about how original skull whistles might have been manufactured by the Aztecs, and to perform in-detail psychoacoustic experiments with these replica skull whistles. We also acquired high-resolution CT scans of the replicas (Fig. 1e). Subsequently, we produced 3D digital reconstructions from the CT scans to compare the 3D models of the replica skull whistles with the 3D models of the original skull whistles (Fig. 1f). Both the original and the replica skull whistles are very similar, if not nearly identical, in their architecture and the corresponding airflow dynamics.

### Skull whistles produce a rough and piercing sound

To quantify the acoustic features as well as the similarity between the original and the replica skull whistles, we recorded and analyzed sounds produced from playing two original skull whistles stored in the EMB (IV Ca 2621 u, IV Ca 2621z; kindly provided by Arnd Adje Both). We also acquired sound recordings from playing several replica skull whistles including the two copies of the original skull whistles of the EMB (IV Ca 2621 u, IV Ca 2621z) and four additional artisanal skull whistles that were manufactured according to the general SW structure but were not exact copies of the original SWs. These recordings overall confirmed the noisy, partly rough and piercing sounds produced by skull whistles. This sound quality can also be inferred from the spectral (frequency) and temporal sound profile of skull whistles, which contains broadband pink-like noise features as well as elements of high-pitched frequencies (Fig. 2a).

A major hypothesis for the existence of skull whistles is that Aztec communities created and used skull whistles to imitate and symbolize natural and mythical sound entities. To explore this hypothesis, we compared the acoustic profile of the original and replica skull whistles with sounds from a broad soundscape. We assembled a large database of 2262 sounds from natural and artificial sources covered by 9 major sound



**Fig. 2 | Acoustic profiles of skull whistles and the MPS profile.** **a** Example spectrograms from four skull whistle sound categories, an original skull whistle (SW orig; upper panel) as well as from a replica skull whistle with sounds produced at three different levels of air pressure (SW high, SW med/medium, SW low). **b** Mean modulation power spectrum (MPS; *pwr* power level) for original ( $n = 35$ ) and replica skull whistle sounds ( $n = 90$  for each level, total  $n = 270$ ) (left upper panel). The other plots show statistical difference maps (log-transformed *p*-values,  $n = 2000$  permutations) between MPSs for the four skull whistle sound categories compared

to sounds from 9 major sound categories (*mus* musical sound effects, *nat* nature, *ani* animal, *hum* human, *int* interior, *ext* exterior, *syn* synthetic, *mex* Mexican flutes, *ins* solo instruments; total  $n = 2262$ ). The difference maps highlight three MPS patches ( $-\log_{10}(p) > \text{abs}(3)$ ) that show relative power differences in skull whistle sounds (left upper panel; patch marked as a —MPS noise patch, patch marked as b —MPS pitch patch, patch marked as c —MPS slow pattern patch). For abbreviations of sound categories, see Tab. S1.

categories and 82 subcategories. To obtain a first general indicator of the (dis)similarity of SWs to other sounds, we calculated the modulation power spectrum (MPS) that quantifies the spectral and temporal profile of sounds (Fig. 2b), which the human auditory system uses to differentiate sounds<sup>30</sup>. This revealed three major observations: first, there were only minor differences between the original and replica skull whistles in how they differed from other sounds. Second, three distinct levels of differences appeared when comparing skull whistles to other sounds, with electronic music effects and natural sounds showing the lowest difference to skull whistles, with animal, human, interior/exterior, and synthetic sounds showing a medium level of difference, and with Mexican flutes and solo instrument sounds showing the most significant difference to skull whistle sounds.

A third important observation was that these differences concerned three different patches of the MPS that have a psychoacoustic significance (Fig. 2b, left upper panel). One patch covered a broad range of spectral and temporal modulation rates and signifies the noisy and rough acoustic nature of the skull whistles. Such effects are often found with sounds that carry scary affective meanings like primate screams<sup>12</sup> or terrifying music<sup>9</sup>. A smaller patch located centrally showed a broad distribution of spectral modulation rates. Such patches typically represent the pitch patterns of sounds explaining the (piercing) pitch impression of skull whistles when compared to some other sound categories (nature, interior), while not having less pitch strength when compared with other sound categories (animal/human, Mexican flutes, instruments). The third patch was centered on low temporal and spectral modulation rates, which are typical for structured and slow-oscillating sound profiles, such as human speech<sup>30</sup> and animal vocal patterns<sup>31,32</sup>. Skull whistles showed significantly lower power in these slow-oscillating patterns, also given that the structure of skull whistles seems not to allow introduction large modulations during playing. Their overall psychoacoustic meaning of skull whistles seems to be the production of single noisy, rough, and piercing sounds.

To obtain a more detailed picture of the (dis)similarity of skull whistle sounds to sounds of the broader soundscape, we quantified 1582 acoustic features of each sound and performed two sets of acoustic similarity analyses. First, a representational similarity analysis (RSA) quantified the correlation between the acoustic feature pattern of sounds, and the RSA indicated that original and replica skull whistles were very similar in terms of acoustic patterns (Fig. 3a, b, red box). Replica skull whistles thus seem to closely match the original skull whistles in terms of their acoustic sound quality, and they both seem to form a unique category of sounds (Fig. S1a), which however also share some similarities with other sound objects. Contemporary and popular Mexican flutes were most similar to skull whistles, and many such flutes visually and acoustically resemble animal species (jaguar, owl, turtle whistles) and iconize skulls (monkey/scary skull whistle, Llorona whistle). These contemporary whistles are also aerophones with a similar playing style to skull whistles. Further similar sounds

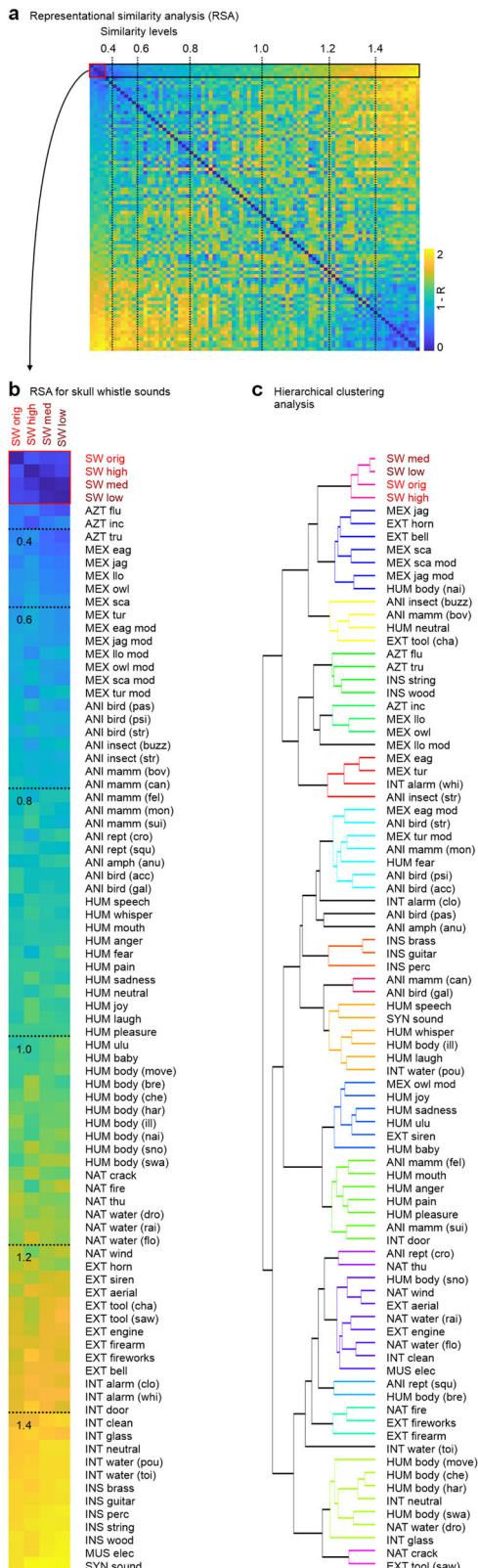
were nonspeech human voices and tool/technical sounds with a rough and shrill sound quality (alarm clocks, chainsaw, emergency horn). Medium similar sounds included animals (bovidae), human (painful bursts, nails scratching), and nature sounds (water, wind), and the most dissimilar sounds were solo instruments (guitar, brass, percussion), digital sounds, tonal human (laughing) and animal sounds (Canidae), and water sounds.

A subsequent hierarchical clustering analysis (HCA) confirmed these findings (Fig. 3c), showing that the original and replica skull whistles formed their own cluster of sound objects. The acoustically closest cluster consisted of contemporary Mexican flutes, tool/technical sounds (emergency horn, bell), and human body sounds (scratching nails). The second closest cluster consisted of animal sounds (insects, bovidae), human voices, and chainsaw sounds. It thus seems that skull whistles were acoustically close to other sounds with an intended acoustic and symbolic iconography, and to sounds that potentially provoke alerting, startling, and affective responses in human listeners. We have to note that these analyses of acoustic similarity and acoustic clustering are solely based on quantitative acoustic data without any bias by human perceptual impressions. But SW sounds to be similar to other sound with well-known basic psychoacoustic effects on listeners potentially across historical ages, SW thus might have been manufactured and used by Aztec communities for this hybrid purpose of serving both a symbolic and affective meaning for potential listeners.

### Skull whistle sounds are perceived as aversive and startling

To investigate the psychoacoustic effects and affective responses in listeners, we performed a perceptual assessment study with human participants. Based on dimensional arousal (how strong does the sound trigger emotional responses in listeners), valence (negative-to-positive emotional quality), urgency (how urgent would a listener respond to the sound), and naturalness ratings (how natural/unnatural is the sound), we located the SWs in this four-dimensional psychoacoustic affective space (Fig. 4a). Skull whistles were rated as being largely of negative emotional quality (valence) with a low-to-medium level of arousal intensity. Skull whistle sounds trigger a medium level of urgency responses in listeners and are rated as sounding largely unnatural.

All four rating scales showed a main effect for differences between sound categories (valence  $F_{1,121.06} = 221.46, p < 0.001$ , eta<sup>2</sup> = 0.10, CI<sub>95%</sub> [0.10 0.11]; arousal  $F_{1,94.96} = 212.63, p < 0.001$ , eta<sup>2</sup> = 0.08, CI<sub>95%</sub> [0.07 0.09]; urgency  $F_{1,115.03} = 494.64, p < 0.001$ , eta<sup>2</sup> = 0.08, CI<sub>95%</sub> [0.07 0.08]; naturalness  $F_{1,113.99} = 106.16, p < 0.001$ , eta<sup>2</sup> = 0.16, CI<sub>95%</sub> [0.15 0.17]), but showed some differential effects when skull whistles were compared to other sound categories. All four skull whistle sound categories were rated similarly in terms of their high negative valence, and they revealed significantly the most negative valence compared with all other sound categories (Fig. 4b; for full statistics on posthoc comparisons, see Tab. S2). Similarly, skull whistles trigger significantly higher urgent tendencies than all other sound

**Fig. 3 | RSA analysis on acoustic profiles and hierarchical clustering analysis.**

**a** Representational similarity analysis (RSA) across all 80 sound categories (left panel) based on a pattern of 1582 acoustic features; dashed vertical lines indicate linear steps (0.2) in similarity. **b** Enlarged view of the four top lines of the RSA values for the four skull whistle sound categories and their acoustic similarity to sounds of the other 76 categories; sounds are sorted for similarity in descending order (lower values represent higher similarity). **c** Hierarchical clustering analysis on all sounds and their acoustic features (colored nodes with linkage cluster threshold of  $c < 2.0$ ). For abbreviations of sound categories, see Tab. S1.

other sound category. Finally, skull whistles sounded more unnatural than original biological sounds (human, animal, nature) and exterior sounds, and they largely also sounded less natural than some musical sounds (music, instrument) (Tab. S5). No significant differences were found in naturalness rating when compared with contemporary Mexican flutes as well as interior and synthetic sounds. The sound of skull whistles thus seems to carry a negative emotional meaning of relevant arousal intensity. This seems to trigger urgent response tendencies in listeners, which is a typical psychoacoustic and affective profile of aversive, scary, and startling sounds<sup>33</sup>.

As for the acoustic patterns of the sounds, we also performed two sets of similarity analyses based on the perceptual patterns from the four rating scales. Using a representational similarity analysis, the original and replica skull whistles again showed the closest perceptual similarity to each other (Figs. 5a, b, S1b). The psychoacoustically and effectively closest sounds to skull whistles were many exteriors (firearm, siren, horn), instruments (Mexican scary flute), and human sounds (fear, pain, anger, sad voice) that have an alarming, startling, and aversive effect on listeners. The psychoacoustically and effectively most distant sounds were instrument (guitar, string, brass, percussion), nature (water, fire), and some Mexican flute sounds (animal imitating flutes), and most of these sounds usually have a pleasant and relaxing listener effect. This pattern was also confirmed by a hierarchical clustering analysis (Fig. 5c), where skull whistles formed their own cluster, immediately neighbored by a cluster of exterior alarming and startling sounds (horn, siren, firearm). Further close sound cluster consisted of human (fear, pain, anger, sad voice) and some Mexican flute sounds (Mexican scary flute).

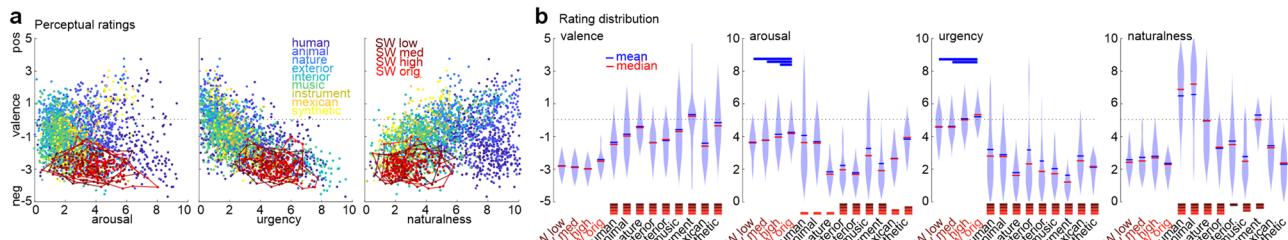
Overall, skull whistles seem to produce a perceptually unique cluster of sounds that share and mimic affective qualities with many other scary and aversive sounds. While some of these effectively similar sounds might not have been present in prehispanic Aztec environments and thus impossible to be directly mimicked by skull whistles (firearm, siren), some other biological sounds were part of Aztec environments (human voices). We might thus speculate that skull whistles have been created by Aztec communities to mimic the acoustic nature of sounds or at least mimic the psychoacoustic and affective impact that sounds can have on listeners. This might have been done for cultural and ritual purposes<sup>11</sup>, but further cross-disciplinary and archeological evidence is needed here.

### Skull whistle sounds have a hybrid natural-artificial nature

If skull whistles acoustically and effectively mimics other sounds and sound qualities, naive listeners might form associations and speculations about the origin of the sound<sup>34</sup>. We asked humans who were completely unfamiliar with SW sounds to listen to a balanced selection of 200 out of the total 2567 sounds to provide a substantive and adjective label for each sound in a free-choice labeling approach. The substantive label should specify the underlying acoustic object or process that potentially produced the sound, while the adjective should provide a verbal affective description of the sound (Fig. 6a).

The most frequent substantive label given by human listeners was “scream” (SW orig: binomial test  $p < 0.001$ , all FDR corrected, relative risk RR = 16.35, CI<sub>95%</sub> [0.24 0.38]; SW high:  $p < 0.001$ , RR = 25.95, CI<sub>95%</sub> [0.65 0.82]; SW med:  $p < 0.001$ , RR = 26.63, CI<sub>95%</sub> [0.46 0.67]; SW low:  $p < 0.001$ , RR = 25.84, CI<sub>95%</sub> [0.45 0.67]), while the two most frequent adjective labels were “aversive” (SW orig:  $p < 0.001$ , RR = 21.07, CI<sub>95%</sub> [0.24 0.39]; SW high:

categories, with original skull whistles triggering somewhat lower urgencies than some of the replica skull whistles (Tab. S3). Skull whistles were mostly significantly more arousing than other sounds, but no significant difference was found when compared to biological sounds (human, animal, nature), some contemporary Mexican flutes, and synthetic sounds (Tab. S4). Only original skull whistle sounds were significantly more arousing than any



**Fig. 4 | Perceptual ratings on sounds.** **a** All 2567 sounds were perceptually rated by  $n = 70$  human listeners along 4 dimensions (arousal, valence, urgency, naturalness) on a 10-point Likert scale. Shown are 2D plots with valence (negative -5, positive 5) on the y-axis and the other scales on the x-axis. Skull whistle sounds are color-coded in red; 9 major sound categories are color-coded in blue-to-yellow. **b** Mean, median, and distribution of ratings for each of the 13 sound categories. Rating on all four scales revealed a high inter-rater consistency (valence ICC = 0.78,  $F_{2566,23094} = 4.611$ ,  $p < 0.001$ ,  $CI_{95\%} [0.77 \ 0.80]$ ; arousal ICC = 0.71,  $F_{2566,23094} = 3.450$ ,  $p < 0.001$ ,  $CI_{95\%}$

[0.69 0.73]; urgency ICC = 0.80,  $F_{2566,23094} = 5.037$ ,  $p < 0.001$ ,  $CI_{95\%} [0.79 \ 0.81]$ ; naturalness ICC = 0.78,  $F_{2566,23094} = 4.597$ ,  $p < 0.001$ ,  $CI_{95\%} [0.77 \ 0.79]$ ). Red-colored bars at the bottom of plots indicate significant differences from the skull whistle categories ( $p < 1e^{-5}$ , FDR corrected, posthoc coefficient test; see Tab. S2–S5 for full statistics on posthoc comparisons); blue-colored bars at the top of plots indicate differences within the skull whistle categories ( $p < 1e^{-5}$ , FDR corrected, posthoc coefficient test; see Tab. S2–S5). For abbreviations of sound categories, see Tab. S1.

$p < 0.001$ , RR = 9.42,  $CI_{95\%} [0.13 \ 0.32]$ ; SW med:  $p < 0.001$ , RR = 17.37,  $CI_{95\%} [0.27 \ 0.48]$ ; SW low:  $p < 0.001$ , RR = 14.69,  $CI_{95\%} [0.22 \ 0.42]$ ) and “scary” (SW orig:  $p < 0.001$ , RR = 11.97,  $CI_{95\%} [0.12 \ 0.25]$ ; SW high:  $p < 0.001$ , RR = 17.28,  $CI_{95\%} [0.23 \ 0.51]$ ; SW med:  $p < 0.001$ , RR = 8.17,  $CI_{95\%} [0.10 \ 0.27]$ ; SW low:  $p < 0.001$ , RR = 11.75,  $CI_{95\%} [0.17 \ 0.35]$ ). Other frequent substantive labels were a mix of living (human) and technical sound sources (concerning vehicles, instruments, or kettles/whistles), with the most diverse labels found for original skull whistles (see Tab. S6–7 for a full report of statistics). Although natural sound sources were mentioned, they were not frequent enough for statistical significance (see Tab. S6). All original substantive and adjective labels were then re-labeled to the labels of the 9 major sound categories according to our major sound dataset, and according to 10 emotional labels from a common taxonomy of human affect<sup>24</sup>. The most frequent substantive label was “human” (SW orig:  $p < 0.001$ , RR = 3.21,  $CI_{95\%} [0.30 \ 0.40]$ ; SW high:  $p < 0.001$ , RR = 6.97,  $CI_{95\%} [0.71 \ 0.83]$ ; SW med:  $p < 0.001$ , RR = 4.19,  $CI_{95\%} [0.40 \ 0.53]$ ; SW low:  $p < 0.001$ , RR = 5.17,  $CI_{95\%} [0.50 \ 0.64]$ ), followed by exterior/interior sound sources (Fig. 6b, see Tab. S8 for full report of statistics). Again, skull whistle sounds were most diverse here, also including “instrument” and “nature” labels. “Fear” (SW orig:  $p < 0.001$ , RR = 6.27,  $CI_{95\%} [0.57 \ 0.68]$ ; SW high:  $p < 0.001$ , RR = 8.68,  $CI_{95\%} [0.82 \ 0.91]$ ; SW med:  $p < 0.001$ , RR = 7.89,  $CI_{95\%} [0.74 \ 0.85]$ ; SW low:  $p < 0.001$ , RR = 8.15,  $CI_{95\%} [0.76 \ 0.86]$ ) was by far the most frequent emotional label, with some original SW sounds also being labeled with “joy” (SW orig:  $p < 0.001$ , RR = 2.51,  $CI_{95\%} [0.21 \ 0.30]$ ) (Tab. S9).

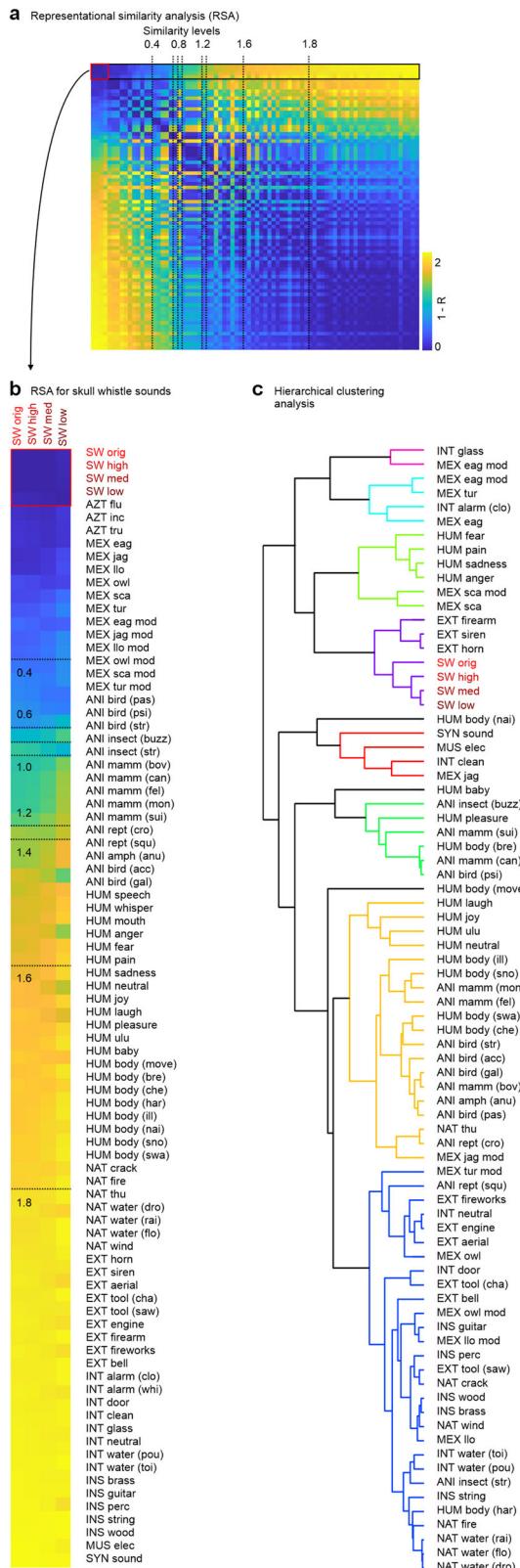
Based on these re-labeled data, we performed a correspondence analysis to assess how labels provided for the SWs would correspond to labels given to the other sounds (Fig. 6c, d). Skull whistle sounds were located in an acoustic object space between human sounds on one end and a cluster of exterior/interior/synthetic sounds on the other end, with also a space of “unknown” labeling close by. “Unknown” labels were provided if listeners could not associate any object or process with the sound. For the affective space, skull whistle sounds were centered around the fear and anger label, as some other human, exterior/interior, and electronic music effects also did. These labeling data together point to a hybrid natural-artificial status of skull whistle sounds, such that they are primarily associated with a human origin, but also form associations with sounds produced from technical objects and processes.

This hybrid status was also confirmed in an independent experiment, where we asked participants to classify sounds into three possible categories regarding the potential origin of the sound as being either from an animated (living organism), technical (technical object/process), or environmental sound source (Fig. 6e). The experiment was set up as a 3-alternative forced choice (3AFC) task. Whereas most of the other sounds were classified into a specific category (except for nature sounds being both technical and environmental) (sound-by-class interaction,  $F_{10,750} = 467.322$ ,  $p < 0.001$ ,

$\eta^2 = 0.09$ ,  $CI_{95\%} [0.08 \ 0.10]$ ; Greenhouse-Geisser (GG) corrected  $p$ -value based on the Mauchly test for checking sphericity violations), skull whistles received a somehow mixed classification both for the original and replica SWs (Fig. 6e left and right panel). For original skull whistles, classifications were mixed between the animated ( $p < 0.001$ ,  $\eta^2 = 0.59$ ,  $CI_{95\%} [0.40 \ 0.56]$ ; comparison against “environment”) and technical categories ( $p < 0.001$ ,  $\eta^2 = 0.42$ ,  $CI_{95\%} [0.24 \ 0.37]$ ; comparison against “environment”), with classifications as “animated” being higher than classifications as “technical” ( $p = 0.003$ ,  $\eta^2 = 0.10$ ,  $CI_{95\%} [0.05 \ 0.29]$ ). An almost identical pattern was found in a second experiment where we used replica SWs instead of the original skull whistle sounds (sound-by-class interaction,  $F_{10,570} = 374.785$ ,  $p < 0.001$ ,  $\eta^2 = 0.68$ ,  $CI_{95\%} [0.65 \ 0.71]$ ; GG corrected). Replica skull whistle sounds again received a mixed classification between the animated ( $p < 0.001$ ,  $\eta^2 = 0.53$ ,  $CI_{95\%} [0.32 \ 0.50]$ ; comparison against “environment”) and technical categories ( $p < 0.001$ ,  $\eta^2 = 0.67$ ,  $CI_{95\%} [0.47 \ 0.65]$ ; comparison against “environment”), but the comparisons between animated and technical classifications did not reveal a significant difference ( $p = 0.092$ ,  $\eta^2 = 0.07$ ,  $CI_{95\%} [-0.02 \ 0.33]$ ).

The pattern of emotional labeling again pointed to a rather negative quality of skull whistle sounds, and the origin might stem from socially relevant human voices and sounds produced by technical devices that have biological significance. Sounds with social and biological relevance often receive increased attention from the cognitive system to prioritize their processing, and this often interferes with other ongoing mental processes. We tested this in a dichotic listening experiment, where humans performed low-high discriminations on simple tones (high- or low-pitch tone) presented on their attended ear, while we presented skull whistle and other sounds on the unattended ear (or no sound, silent condition) (Fig. 6f). For an experiment including original skull whistle sounds, all sounds presented on the unattended ear interfered with the tone discrimination task performed on the attended ear ( $F_{7,322} = 9.075$ ,  $p < 0.001$ ,  $\eta^2 = 0.01$ ,  $CI_{95\%} [0.003 \ 0.011]$ ; GG corrected), especially in terms of slowing down the reaction time (posthoc tests against silent condition; all  $p < 0.020$ ; for full statistics see Tab. S10). Comparing original skull whistles with all other sound categories regarding the interference level did not reveal significant differences ( $p > 0.473$ , Tab. S10), but only exterior sounds showed a significantly higher interference level than skull whistle sounds ( $p = 0.005$ ). The same pattern of results was obtained in a second experiment with replica skull whistles instead of the original skull whistle sounds ( $F_{7,308} = 13.503$ ,  $p < 0.001$ ; GG corrected; posthoc tests against the silent condition, all  $p < 0.003$ , for full statistics see Tab. S11), including significantly stronger interference effects of exterior sounds compared to skull whistle sounds ( $p = 0.046$ , Tab. S11).

Being able to interfere with other ongoing mental processes highlights the notion that skull whistle sounds carry some relevant social and/or biological meaning and associations to human listeners. We have to highlight the notion here again that we only investigated these psychoacoustic



**Fig. 5 | RSA analysis on perceptual ratings on sounds and hierarchical clustering analysis.** **a** Representative similarity analysis (RSA) on the pattern of perceptual ratings (left panel); dashed vertical lines indicate linear steps (0.2) in similarity. **b** Enlarged view of the four top lines are the RSA values for the four skull whistle sound categories and their perceptual similarity to sounds of the other 76 categories; sounds are sorted for similarity in descending order (lower values represent higher similarity). **c** Hierarchical clustering analysis on all sounds and their perceptual ratings (colored nodes with linkage cluster threshold of  $c < 2.0$ ). For abbreviations of sound categories, see Tab. S1.

evolutionary closer modern and Aztec humans, and psychoacoustic effects of SWs might have been capitalized on by Aztec communities.

### Neural decoding of skull whistle sounds requires elaborated and associative processing

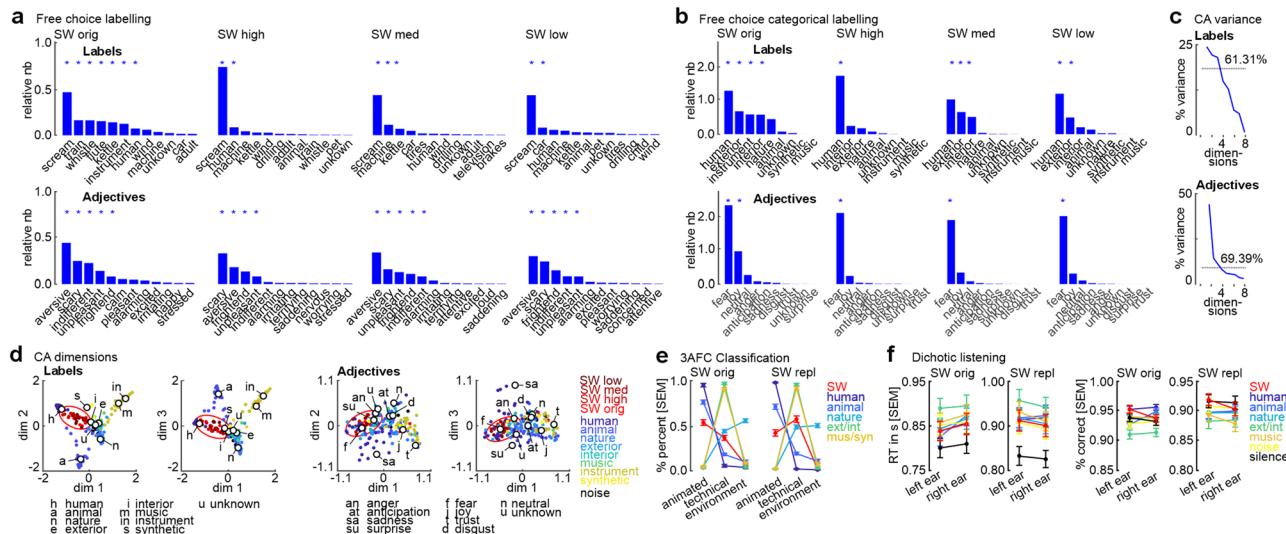
We finally performed a neuroimaging experiment with human listeners using functional magnetic resonance imaging. We investigated the neural processes that support decoding of such associated and affective meanings from SW sounds, examining whether this decoding happens at the level of basic auditory processing<sup>37</sup> or rather at the level of higher-order cognition<sup>38</sup>. We again presented the selection of 200 sounds from the broader sound dataset, including original and replica skull whistle sounds, sounds from five major sound categories (human, animal, technical, nature, instrument) as well as additional eight white and pink noise stimuli. Common and known sounds typically elicit activity in auditory cortical regions (AC)<sup>39</sup> for basic acoustic analysis and auditory cognition<sup>40</sup>. Sounds from the five major sound categories elicited significantly higher activity in many low- and higher-order AC regions compared to skull whistle sounds (Fig. 7a, Tab. S12).

Contrarily, neural activity for skull whistle sounds was rather located in regions for higher-order cognition. Skull whistle sounds compared to all other sounds elicited brain activity in the lateral (IFC inferior frontal cortex) and medial frontal cortex (MFC) as well as the insula. The IFC typically provides elaborated sound evaluation and classification processes<sup>38</sup>, with potential support of sensory-affective integration mechanisms in the insula, and associative processing in the MFC<sup>37</sup>. Brain activity in the insula and MFC as well as additional activity in the intraparietal sulcus (IPS) for skull whistle sounds were especially found when compared to brain activity for animated sounds (human, animal) but not when compared to artificial sounds (technical, instrument) (Fig. 7b; see also Fig. S2). This might further highlight the notion that skull whistle sounds are of a hybrid nature, such that they appear as partly coming from an animated source but also having a wider symbolic and associative meaning. This wider meaning might warrant increased processing in the fronto-insular brain systems as well as an elaborated sound mapping and acoustic evidence integration in the IPS<sup>41</sup>.

This right fronto-parietal-insular network for skull whistle sound processing was largely interconnected to itself, as revealed by a functional brain connectivity analysis, with additional connections to the left insula and the anterior cingulate cortex, which supports the MFC in the cognitive appraisal of sounds<sup>42</sup> (Fig. 7c, Tab. S13). Interestingly, the functional network for processing the other sounds, which was centered especially in left AC, showed links with the skull whistle processing network, especially to the right fronto-insular system and the MFC. Whereas for other sounds, this connection might point to a natural forward mapping from sound analysis (auditory cortex) to sound evaluation (fronto-insular), the neural network for skull whistle sound processing suggests a challenging level of elaborated processing and higher-order cognition.

While the previous analysis mainly quantified the neural differences in the processing of skull whistle sounds compared with other sounds, we finally also performed an analysis that focused more on the neural similarity in processing such sounds. Identical to the RSA of the acoustic and perceptual patterns for all 2567 sounds (Figs. 3a, b, 5a, b; Tab. S14), we performed an RSA for the 192 selected sounds (excluding noise sounds) of the

effects in samples of naïve European listeners from modern cultures. Basic psychoacoustic distraction effects of meaningful sounds are however even present in primate species close to humans, which share basic auditory and affective sound processing dynamics<sup>35,36</sup>. Despite cultural differences, there might also be large psychoacoustic processing similarities between the



**Fig. 6 | Free categorical and emotional labeling of sounds.** **a** Human listeners ( $n = 40$ ) labeled 200 sounds (4 skull whistle categories, 8 other sound categories) with a substantive label (“Labels”) describing the origin or source of the sound (upper panel) as well as with an adjective label (“Adjectives”) describing their emotional response to the sound (lower panel). Plots show the sorted frequency of labels used as the relative number (relative nb = total nb of label multiplied by the percentage of listeners using the label; y-axis is relative nb/100); plots only show results for the four skull whistle categories. Asterisks indicate significance  $p < 0.05$  (significant difference of observed frequency above expected probability), binomial test, FDR corrected; see Tab. S6-7. **b** To perform a correspondence analysis (CA) on the labels, we re-labeled the original labels along the 9 major sound categories and 8 major emotional dimensions according to the model by Plutchik<sup>24</sup>. Re-labeling was done based on the classification probabilities of  $n = 7$  independent raters. \* $p < 0.05$ ,

binomial test, FDR corrected; see Tab. S8-9. **c** Explained variance of the dimensions resulting from the CA; dimensions 1–3 explained >61% of the variance in the data.

**d** Dimensions 1–3 of the CA for categorical labels and emotional adjectives.

**e** Probability for classifying sounds skull whistle sounds and sounds from 5 other categories as originating from an “animated”, “technical”, or environmental source. Left plot is for the experiment including original skull whistles ( $n = 76$ ), right plot is for the experiment including replica skull whistles ( $n = 58$ ). **f** Reaction time and accuracy data from the dichotic listening experiment, where humans discriminated low/high tone on the attended ear while presenting other sounds or silence on the unattended ear. Left plot is for the experiment including original skull whistles ( $n = 47$ ), the right plot is for the experiment including replica skull whistles ( $n = 47$ ).

neuroimaging experiment. Again, the skull whistles were most similar to each other, and all other sounds showed a variable degree of similarity to the skull whistle based on the acoustic and perceptual patterns (Fig. 7d). We combined these parametric acoustic and perceptual data with a parametric analysis of the neural similarity of activation patterns between sound categories<sup>43</sup>. For the latter, we performed a multivariate decoding analysis at every brain location to quantify the likelihood that each location can separate between any of two sound categories based on their activation pattern. The reverse of this separation likelihood was taken as a measure of neural similarity<sup>43</sup>. The resulting neural cross-correlation data as described below were quite distinctive to the observed acoustic and perceptual similarity patterns in relation to skull whistles, as no neural significance was found when randomly permuting the similarity matrixes (see “Methods”).

Cross-correlating the acoustic similarity for the sound categories with neural similarity revealed a broadly extended neural significance in the AC, IFC, and MFC, and (pre)motor cortex (Fig. 7e). Cross-correlating perceptual similarity with neural similarity revealed neural significance in bilateral low-order AC and right higher-order AC. This focus of neural significance for the perceptual similarity analysis on the AC might have been driven by similarity to very specific sounds, such as baby cries, screams, or alarm sounds, which have distinctive rough and piercing sound features as well as aversive affective features, which receive low-level decoding in the neural auditory system<sup>44,45</sup>. The neural similarity analysis based on acoustic patterns seems to involve a more extensive brain network and various levels of processing, potentially ranging from acoustic analysis (AC), and acoustic object classifications (IFC), to elaborate evaluations (MFC). The specific role of the (pre)motor cortex might be to represent embodied acoustic recognitions<sup>46</sup> and to discriminate acoustic patterns that are critical for behavioral choices<sup>47</sup>.

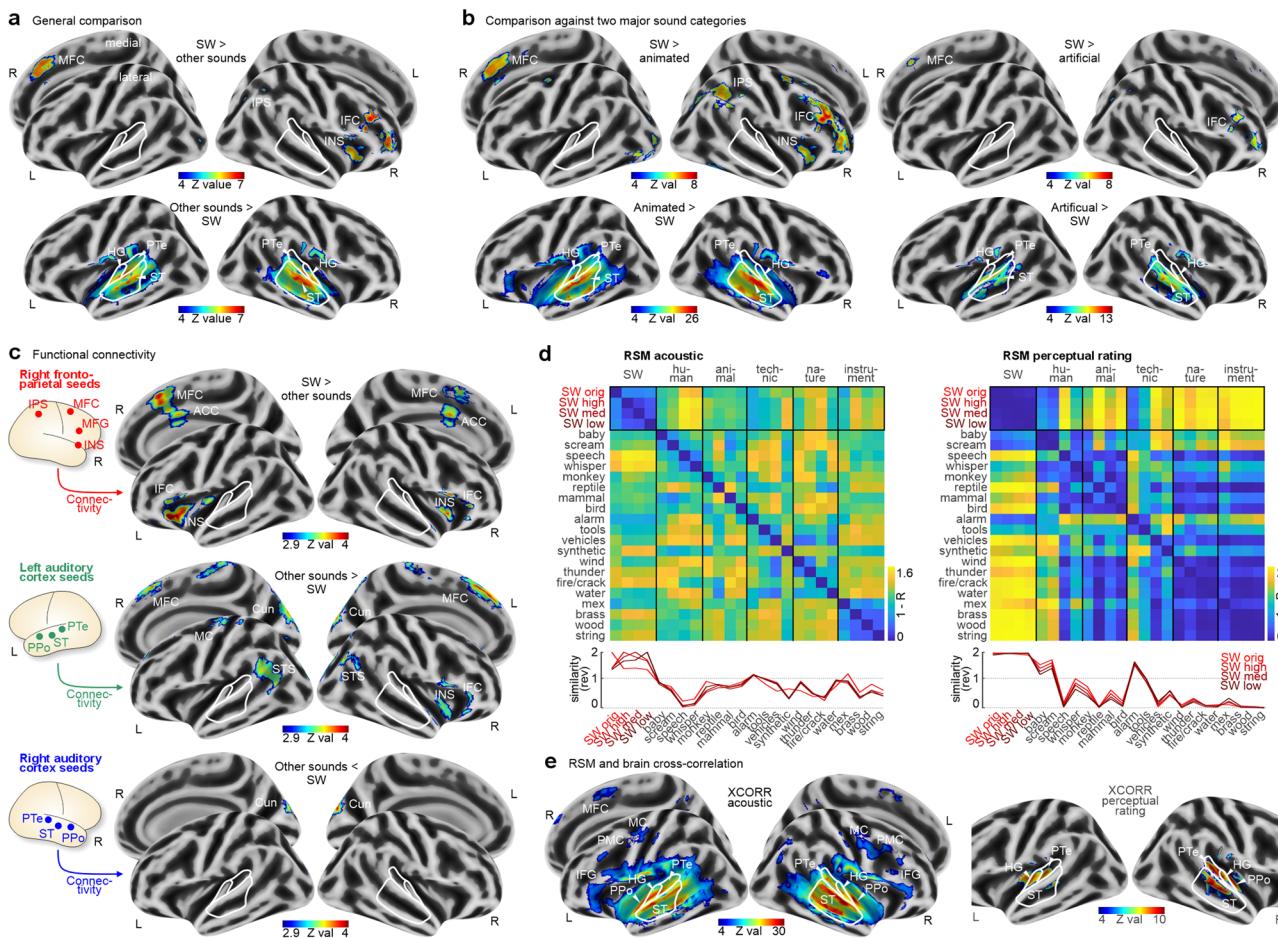
While the neural effort for decoding similarities of acoustic patterns between skull whistle and other sounds thus seems to involve a multi-level

process, pointing to an ambiguous mixture of familiar and unfamiliar sound components as well as a multi-layered symbolism, the neural effort in terms of matching perceived affective similarities seems rather unambiguous.

## Discussion

We carefully assume a certain level of comparability of how modern humans and humans in Aztec communities many centuries ago responded to skull whistle sounds on a basic affective level. This assumption might be valid given the very basic and salient psychoacoustic effects that rely on biological and neural principles of sound recognition that are even shared between humans and animals<sup>36,48</sup>. Human listeners in our experiments rated skull whistle sounds as very negative and specifically labeled them largely as scary and aversive, which potentially also trigger urgent response tendencies and interfere with ongoing mental processes<sup>37</sup>. This immediacy effect is further highlighted by very specific brain activity in low-order AC that is correlated with the affective similarity of SWs to other sounds. In case of an aversive and alerting sound quality, low-order AC detects the aversive sound quality<sup>44</sup> and tunes the neural system for in-depth sound analysis<sup>33,49</sup>.

We of course have to highlight the notion that our experimental data were acquired with samples from a modern European population. Our data thus first provide basic evidence of how modern humans respond to the acoustic quality of a historically very important sound tool with a unique acoustic profile. All listeners in our experiments were naïve about the presence of skull whistle sounds on the acoustic samples presented in the experiments, and thus all data are largely unbiased and basic assessments of the psychoacoustic effects of skull whistles. On this basic level of psychoacoustic processing, we might assume some similarity to the acoustic processing in humans from previous Aztec cultures. Aztec communities might have capitalized on this aversive and scary nature of skull whistle sounds in specific contexts, which would support the introduced warfare and ritual symbolism hypothesis, but rather not the deity symbolism hypothesis. However, given also the



**Fig. 7 | Brain responses to skull whistle sounds.** **a** Functional brain activity in human listeners ( $n = 32$ ) when comparing activity for skull whistles with activity for all other 5 general sound categories. **b** Brain activity when contrasting skull whistle with animated sounds (human, animal; left panel) and with artificial sounds (technical, instrument; right panel). **c** Functional connectivity patterns from seed regions in the right fronto-parietal network that were significant for processing skull whistle sounds (upper panel), and for the left (mid panel) and right auditory cortex regions (lower panel) that were significant for processing other sounds. Voxel-level

threshold of  $p = 0.005$  combined with an FWE-corrected cluster-level threshold of  $p = 0.05$ . **d** Representative similarity matrices (RSM) for acoustic and perceptual patterns of the sounds included in the neuroimaging experiment. Bottom plots show the RSM for the skull whistle sounds only (rev, reversed coding), corresponding to the black box in the upper panels. **e** Brain areas with significant cross-correlation between the acoustic and perceptual similarity patterns and measures of neural similarity between sound categories. All brain activity from contrasts includes a voxel threshold  $p < 0.05$  (FWE corrected), cluster threshold  $k = 10$ .

strong associative/symbolic sound impressions requiring higher-order cognition, warfare usage seemed rather unlikely. Given both the aversive/scary and associative/symbolic sound nature as well as currently known excavation locations at ritual burial sites with human sacrifices, usage in ritual contexts seems very likely, especially in sacrificial rites and ceremonies related to the dead. Skull whistles might have been used to scare the human sacrifice or the ceremonial audience, but further cross-documentation is needed here. The symbolic and associative meaning might be specifically related to the dangerous and scary travel of the dead to the underworld. This potential double nature of a biological and symbolic affective significance might also be based on the hybrid natural-artificial status of skull whistle sounds, as SWs sounds received ambiguous associations with natural and artificial sound sources across different experiments and analyses in our study.

Listeners in our experiments were naïve to the presented sounds, especially about the fact that the broader sound set also included skull whistle sounds. Modern humans perceived skull whistle sounds as being of a hybrid natural-artificial nature and a hybrid familiar-unfamiliar nature. Unlike naïve listeners in our experiments, human listeners in Aztec communities were probably aware that the sound originated from the skull whistles as a musical sound tool, especially when used in ritual contexts and being a common part of Aztec musical culture. Being aware that the skull whistle sound was produced by a technical tool, it still seems to elicit impressions of originating or

mimicking an animated sound source. This might greatly contribute to the symbolic and potentially fictional nature of skull whistle sounds<sup>50</sup> and highlights the frequent intention of ancient cultures to capture and represent mythical entities/contexts in musical tools<sup>50–52</sup>.

In terms of the hybrid, symbolic, and potentially fictional nature of skull whistle sounds mentioned above, this is also reflected in our neuroimaging brain data, which makes the ritual hypothesis more likely than the warfare hypothesis. Listening to skull whistle sounds elicited significantly higher brain activity in regions and neural networks associated with higher-order auditory cognition, sound evaluation, and associative processing. Sounds with affective significance commonly trigger neural processes to evaluate the emotional nature and the specific source of the sound because this is relevant to preparing appropriate behavioral responses<sup>37</sup>. Skull whistle sounds are unambiguous in their affective nature but are rather ambiguous in the determination of their sound origin, which intensifies higher-order brain processing. The psychoacoustic ambiguity and hybridity however open many options for symbolism in an archeoacoustic perspective.

### Limitations

As an important limitation, we have first to highlight the notion that we only tested the psychoacoustic nature of skull whistles in human samples from modern European populations, which imposes some limitations in terms of

the anthropological and cultural comparability to Aztec cultures many centuries ago. This unfortunately is a common limitation for all experimental historical studies. However, we are confident that some basic neurocognitive mechanisms are shared between modern humans and humans from Aztec cultures. Many affective and auditory cognitive mechanisms are shared between humans and their closest monkey species<sup>35,36</sup>. The shared neurocognitive mechanisms between modern humans and Aztec humans thus seem very reasonable, and we tested for the basic neurocognitive mechanisms of decoding skull whistle sounds in the last experiment in our study.

A related topic concerns the pattern of sound categories that were used in some experiments. These categories to classify sounds were derived from modern classifications of sounds as used in acoustic, machine-learning, and psychological studies<sup>15,37</sup>. Some of these sound categories might not match the soundscape of Aztec cultures. Unfortunately, a detailed taxonomy of sound classes described by Aztec cultures is missing today, but major sound classes (human, animal, nature, music, tool sounds) might be similar to modern humans. We have to note that all participants in our study were naive and unbiased regarding the sounds used, and the sounds were not presented as belonging to a certain sound category. The taxonomy of sound classes was only used posthoc during data analysis for the purpose of data grouping.

An additional common limitation for experimental historical studies is that evidence is often not ultimate in favor of one of several competing hypotheses. New evidence might support certain hypotheses to some degree while disconfirming other hypotheses to another degree. This also concerns the data reported here, which are relatively but not absolutely in favor of one of the current hypotheses concerning skull whistles (ritual symbolism hypothesis). To obtain a relatively broad, consistent, and relatively unbiased assessment of SW sounds, we here performed several physical and perceptual experiments on skull whistle sounds using diverse approaches. This allowed a multi-level and critical weighing of evidence in support of certain hypotheses.

## Data availability

The conditions of ethics approval and consent procedures do not permit the public archiving of the participants' anonymized study data. These data are available upon request to the corresponding author in consultation with the cantonal ethics committee of the Swiss canton Zurich. Preprocessed numerical data are available at <https://github.com/caneuro/skullwhistle>, and functional brain maps are available at <https://identifiers.org/neurovault.collection:18295>. Examples of skull whistle sounds can be found here: <https://caneuro.github.io/blog/2024/study-skullwhistle/>.

## Code availability

Study-specific analysis codes and data are deposited at <https://github.com/caneuro/skullwhistle>, and the code can be used to reproduce the statistical analysis as reported in the manuscript, which can be used together with the Several codes for the data analysis are taken from existing toolboxes: MPS analysis (<https://github.com/theunissenlab/soundsig>), RSA analysis ([https://github.com/rsagroup/rsatoolbox\\_matlab](https://github.com/rsagroup/rsatoolbox_matlab)), CA analysis (<http://www.sthda.com/english/articles/31-principal-component-methods-in-r-practical-guide/113-ca-correspondence-analysis-in-r-essentials/#r-packages>), brain data analysis with SPM12 brain connectivity analysis with the CONN toolbox (<https://web.conn-toolbox.org/>), and MVPA analysis (<https://sites.google.com/site/ttddecodingtoolbox/>).

Received: 24 May 2024; Accepted: 31 October 2024;

Published online: 11 November 2024

## References

- McDermott, J. The evolution of music. *Nature* **453**, 287–288 (2008).
- Both, A. A. Skull whistles of the Aztecs (Totenkopfpfeifen der Azteken). *Baessler-Arch.* **53**, 43–54 (2005).
- Velázquez Cabrera, R. Silbato de la muerte. *Arqueología* **42**, 184–202 (2009).
- von Hornbostel, E. M. & Sachs, C. Systematik der Musikinstrumente. Ein Versuch. *Z. Ethnol.* **553**, 590 (1914).
- MIMO Consortium. Revision of the Hornbostel-Sachs Classification of Musical Instruments by the MIMO Consortium. <http://www.mimo-international.com/documents/hornbostel%20sachs.pdf> (2008).
- Beristain, S., Menchaca, R. & Velazquez, R. Ancient noise generators. *J. Acoust. Soc. Am.* **112**, 2368–2368 (2002).
- Both, A. A. On the context of imitative and associative processes in Prehispanic music. *Stud. zur. Musikarchäologie* **5**, 319–332 (2006).
- Kollmann, J. Flöten und Pfeifen aus Alt-Mexiko. in *Adolf Bastian als Fest-Gruss zu seinem 70. Geburtstage* 559–574 (Berlin, 1896).
- Trevor, C., Arnal, L. H. & Fröhholz, S. Terrifying film music mimics alarming acoustic feature of human screams. *J. Acoust. Soc. Am.* **147**, EL540–EL545 (2020).
- Velazquez, R. & Rivas Castro, F. Death whistle. *J. Acoust. Soc. Am.* **128**, 2389–2389 (2010).
- Both, A. A. Aztec music culture. *World Music* **52**, 14–28 (2010).
- Fröhholz, S., Dietziker, J., Staib, M. & Trost, W. Neurocognitive processing efficiency for discriminating human non-alarm rather than alarm scream calls. *PLoS Biol.* **19**, e3000751 (2021).
- Fröhholz, S., Trost, W. & Grandjean, D. Whispering—The hidden side of auditory communication. *Neuroimage* **142**, 602–612 (2016).
- Fröhholz, S., Klaas, H. S., Patel, S. & Grandjean, D. Talking in fury: the cortico-subcortical network underlying angry vocalizations. *Cereb. Cortex* **25**, 2752–2762 (2015).
- Picczak, K. J. ESC: Dataset for environmental sound classification. in *MM 2015—Proceedings of the 2015 ACM Multimedia Conference* 1015–1018 (Association for Computing Machinery, Inc, 2015). <https://doi.org/10.1145/2733373.2806390>.
- Belin, P., Fillion-Bilodeau, S. & Gosselin, F. The Montreal affective voices: a validated set of nonverbal affect bursts for research on auditory affective processing. *Behav. Res Methods* **40**, 531–539 (2008).
- Philharmonia Orchester. Philharmonia Database.
- Eyben, F., Weninger, F., Gross, F. & Schuller, B. Recent developments in openSMILE, the Munich open-source multimedia feature extractor. in *Proceedings of the 21st ACM International Conference on Multimedia—MM '13* 835–838 (ACM Press, New York, New York, USA, 2013). <https://doi.org/10.1145/2502081.2502224>.
- Eyben, F. et al. The Geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing. *IEEE Trans. Affect Comput* **7**, 190–202 (2016).
- Colace, F. & Casaburi, L. An approach for sentiment classification of music. in *ICEIS 2016—Proceedings of the 18th International Conference on Enterprise Information Systems* vol. 2 421–426 (SciTePress, 2016).
- Elliott, T. M. & Theunissen, F. E. The modulation transfer function for speech intelligibility. *PLoS Comput. Biol.* **5**, e1000302 (2009).
- Nili, H. et al. A toolbox for representational similarity analysis. *PLoS Comput. Biol.* **10**, (2014).
- Bones, O., Cox, T. J. & Davies, W. J. Sound categories: category formation and evidence-based taxonomies. *Front. Psychol.* **9**, 1277 (2018).
- Plutchik, R. A psychoevolutionary theory of emotions. *Soc. Sci. Inf.* **21**, 529–553 (1982).
- Ashburner, J. & Friston, K. J. Diffeomorphic registration using geodesic shooting and Gauss–Newton optimisation. *Neuroimage* **55**, 954–967 (2011).
- Whitfield-Gabrieli, S. & Nieto-Castañon, A. Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect* **2**, 125–141 (2012).
- Hebart, M. N., Görgen, K. & Haynes, J.-D. The Decoding Toolbox (TDT): a versatile software package for multivariate analyses of functional imaging data. *Front. Neuroinform.* **8**, 88 (2015).
- Turner, B. O., Mumford, J. A., Poldrack, R. A. & Ashby, F. G. Spatiotemporal activity estimation for multivoxel pattern analysis with rapid event-related designs. *Neuroimage* **62**, 1429–1438 (2012).

29. Nichols, T., Brett, M., Andersson, J., Wager, T. & Poline, J. B. Valid conjunction inference with the minimum statistic. *Neuroimage* **25**, 653–660 (2005).
30. Woolley, S. M. N., Fremouw, T. E., Hsu, A. & Theunissen, F. E. Tuning for spectro-temporal modulations as a mechanism for auditory discrimination of natural sounds. *Nat. Neurosci.* **8**, 1371–1379 (2005).
31. Fukushima, M., Doyle, A. M., Mullarkey, M. P., Mishkin, M. & Averbeck, B. B. Distributed acoustic cues for caller identity in macaque vocalization. *R. Soc. Open Sci.* **2**, 150432 (2015).
32. Hechavarria, J. C., Jerome Beetz, M., Garcia-Rosales, F. & Kössl, M. Bats distress vocalizations carry fast amplitude modulations that could represent an acoustic correlate of roughness. *Sci. Rep.* **10**, 7332 (2020).
33. Arnal, L. H., Kleinschmidt, A., Spinelli, L., Giraud, A.-L. & Mégevand, P. The rough sound of salience enhances aversion through neural synchronisation. *Nat. Commun.* **10**, 3671 (2019).
34. Guastavino, C. et al. Sound categories: category formation and evidence-based taxonomies. (2018) <https://doi.org/10.3389/fpsyg.2018.01277>.
35. Bodin, C. et al. Functionally homologous representation of vocalizations in the auditory cortex of humans and macaques. *Curr. Biol.* **31**, 4839–4844.e4 (2021).
36. Fröhholz, S. & Schweinberger, S. R. Nonverbal auditory communication—evidence for integrated neural systems for voice signal production and perception. *Prog. Neurobiol.* (2020) <https://doi.org/10.1016/j.pneurobio.2020.101948>.
37. Fröhholz, S., Trost, W. & Kotz, S. A. The sound of emotions—towards a unifying neural network perspective of affective sound processing. *Neurosci. Biobehav. Rev.* **68**, 1–15 (2016).
38. Fröhholz, S. & Grandjean, D. Processing of emotional vocalizations in bilateral inferior frontal cortex. *Neurosci. Biobehav. Rev.* **37**, 2847–2855 (2013).
39. Bizley, J. K. & Cohen, Y. E. The what, where and how of auditory-object perception. *Nat. Rev. Neurosci.* **14**, 693–707 (2013).
40. Rauschecker, J. P. & Scott, S. K. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat. Neurosci.* **12**, 718–724 (2009).
41. Yao, J. D., Gimoto, J., Constantinople, C. M. & Sanes, D. H. Parietal cortex is required for the integration of acoustic evidence. *Curr. Biol.* **30**, 3293–3303.e4 (2020).
42. Etkin, A., Egner, T. & Kalisch, R. Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends Cogn. Sci.* **15**, 85–93 (2011).
43. Giordano, B. L., McAdams, S., Zatorre, R. J., Kriegeskorte, N. & Belin, P. Abstract encoding of auditory objects in cortical activity patterns. *Cereb. Cortex* **23**, 2025–2037 (2013).
44. Concina, G., Renna, A., Grossi, A. & Sacchetti, B. The auditory cortex and the emotional valence of sounds. *Neurosci. Biobehav. Rev.* **98**, 256–264 (2019).
45. Staib, M., Abivardi, A. & Bach, D. R. Primary auditory cortex representation of fear-conditioned musical sounds. *Hum. Brain Mapp.* **41**, 882–891 (2020).
46. Archakov, D. et al. Auditory representation of learned sound sequences in motor regions of the macaque brain. *Proc. Natl Acad. Sci.* **117**, 15242–15252 (2020).
47. Morán, I., Perez-Orive, J., Melchor, J., Figueroa, T. & Lemus, L. Auditory decisions in the supplementary motor area. *Prog. Neurobiol.* **202**, 102053 (2021).
48. LeDoux, J. Rethinking the emotional brain. *Neuron* **73**, 653–676 (2012).
49. Kumar, S., von Kriegstein, K., Friston, K. & Griffiths, T. D. Features versus feelings: dissociable representations of the acoustic features and valence of aversive sounds. *J. Neurosci.* **32**, 14184–14192 (2012).
50. Trevor, C. & Fröhholz, S. Music as an evolved tool for socio-affective fiction. *Emot. Rev.* **16**, 180–194 (2024).
51. Mehr, S. A., Krasnow, M. M., Bryant, G. A. & Hagen, E. H. Origins of music in credible signaling. *Behav. Brain Sci.* **44**, e60 (2021).
52. Barber, S. B. & Sánchez, M. O. A divine wind: the arts of death and music in terminal formative Oaxaca. *Anc. Mesoam.* **23**, 9–24 (2012).

## Acknowledgements

We are very thankful to *Arnd Adje Both* for helpful advice during the study from a music archeological perspective and for guiding and supervising the artisanal reproduction of original skull whistles. We thank *Osvaldo Padrón Pérez* for manufacturing two of the skull whistle replicas. We thank the *Ethnological Museum of Berlin (EBM)* for access to original skull whistle artifacts. This study was supported by the Swiss National Science Foundation (SNSF PP00P1\_157409/1 and 100014\_182135/1). The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

## Author contributions

S.F., P.R. contributed to designing the experiments, data acquisition, data analysis, and writing the manuscript; M. Bobin, M. Bonard, F.S. contributed to data acquisition and data analysis.

## Competing interest

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s44271-024-00157-7>.

**Correspondence** and requests for materials should be addressed to Sascha Fröhholz.

**Peer review information** *Communications psychology* thanks Etienne Thoret and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Marike Schiffer. A peer review file is available.

**Reprints and permissions information** is available at <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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

© The Author(s) 2024