COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 43 (2019) e12734

ISSN: 1551-6709 online
DOI: 10.1111/cogs.12734

The Aesthetic Preference for Nature Sounds Depends on Sound Object Recognition

Stephen C. Van Hedger,[a,b] Howard C. Nusbaum,[a,b]
Shannon L. M. Heald,[a,b] Alex Huang,[a,b] Hiroki P. Kotabe,[c]
Marc G. Berman[a,b]

[a]*Department of Psychology, The University of Chicago*
[b]*Center for Practical Wisdom, The University of Chicago*
[c]*SKK Graduate School of Business, Sungkyunkwan University*

Received 23 August 2018; received in revised form 13 March 2019; accepted 2 April 2019

Abstract

People across the world seek out beautiful sounds in nature, such as a babbling brook or a nightingale song, for positive human experiences. However, it is unclear whether this positive aesthetic response is driven by a preference for the perceptual features typical of nature sounds versus a higher-order association of nature with beauty. To test these hypotheses, participants provided aesthetic judgments for nature and urban soundscapes that varied on ease of recognition. Results demonstrated that the aesthetic preference for nature soundscapes was eliminated for the sounds hardest to recognize, and moreover the relationship between aesthetic ratings and several measured acoustic features significantly changed as a function of recognition. In a follow-up experiment, requiring participants to classify these difficult-to-identify sounds into nature or urban categories resulted in a robust preference for nature sounds and a relationship between aesthetic ratings and our measured acoustic features that was more typical of easy-to-identify sounds. This pattern of results was replicated with computer-generated artificial noises, which acoustically shared properties with the nature and urban soundscapes but by definition did not come from these environments. Taken together, these results support the conclusion that the recognition of a sound as either natural or urban dynamically organizes the relationship between aesthetic preference and perceptual features and that these preferences are not inherent to the acoustic features. Implications for nature's role in cognitive and affective restoration are discussed.

Keywords: Audition; Aesthetics; Nature; Categorization; Perception; Concepts; Attention restoration theory; Stress reduction theory

Correspondence should be sent to Stephen Van Hedger, Department of Psychology, The Brain and Mind Institute, Western University, London, ON, Canada N6A 3K7. E-mail: svanhedg@uwo.ca

1. Introduction

Compared to our experiences with humanmade structures and urban environments, interacting with nature provides a range of cognitive, affective, and health benefits. Experiences with nature can improve directed attention (Berman, Jonides, & Kaplan, 2008; Berman et al., 2012), reduce perceived pain (Kline, 2009) and stress (Alvarsson, Wiens, & Nilsson, 2010), and generally increase well-being (Marselle, Irvine, & Warber, 2014). Benefits of nature have also been documented in studies that use greenspace as a quasi-independent variable, suggesting that one's ability to engage with nature on a day-to-day basis may have critical implications for academic and professional performance (cf. Berto, 2005; Raanaas, Evensen, Rich, Sjostrom, & Patil, 2011), cognitive development (Dadvand et al., 2015), and general health (Kardan, Gozdyra, et al., 2015).

An important consideration in any discussion of the mechanisms through which nature confers these benefits is the aesthetic preference of naturalistic experiences. Aesthetic preference has been defined as a "like-dislike" affective response (Zajonc, 1980), and the affective restorative potential of nature has been clearly associated with a strong aesthetic preference for nature stimuli (Han, 2010; Hartig & Staats, 2006; Purcell, Peron, & Berto, 2001; van den Berg, Koole, & van der Wulp, 2003). Yet, despite a clear association between aesthetic preferences for nature and restoration, the origin of our aesthetic preferences for nature is debated. Understanding the level at which this association forms is a fundamental scientific question—which is also tied to discussions of the psychological role of genetic endowment and evolution versus the effect of experience—that has broad implications for how we should conceptualize nature as a restorative environment.

One view, first voiced in the "Biophilia Hypothesis" (Gullone, 2002; Martín-López, Montes, & Benayas, 2007; Ulrich, 1993; Wilson, 1984; Wilson & Kellert, 1993) is that the preference for nature developed on an evolutionary time scale through natural selection. According to this account, our tendency to prefer natural environments is an inherited trait that was selected through the enhanced fitness of our ancestors who demonstrated a similar affinity. In nature, we are at our best both affectively and cognitively due to our species' extensive history with natural environments. Urban or humanmade environments, in contrast, have only made up a small fraction of our evolutionary history, and thus may not represent a particularly aesthetic environment that is optimized for human cognitive and affective processing. While there is some disagreement as to whether the innate affinity for nature is granted by genetic traits or by biologically prepared learning, the evolution-based perspectives on the origin of nature preferences argue that humans have been innately programmed to process, attend to, and ultimately prefer visual and auditory features present in natural scenes. Throughout this paper, we refer to this innate and evolutionary perspective as the *inherent-organization hypothesis*.

An alternative view to this evolution-based *inherent-organization hypothesis*, however, is that the aesthetic preference for nature is a product of cultural transmission, born out of historical and social treatments of natural environments and learned through practice (Simaika & Samways, 2010). Such *constructivist* accounts suggest that the aesthetic preference of a given natural environment depends less on how the visual and auditory features that comprise the environment may be innately processed due to evolutionary constraints. Rather, aesthetic preference for a given natural environment may critically depend on the extent to which the visual and auditory features that comprise the environment can be recognized and associated with the *concept* of nature, which then confers aesthetic and restorative value. Put another way, under a constructivist perspective, the salubrious effects of nature—including positive aesthetic evaluations—are garnered through the semantic idea of nature and its restorative abilities (Carlson, 1981). Throughout this paper, we refer to this constructivist perspective as the *conceptual-organization hypothesis*, as it suggests that conceptual knowledge (e.g., knowing that an environment is natural versus urban) will dynamically organize how individuals aesthetically value specific features of the environment.

Evolution-based accounts and constructivist accounts on the origin of nature preferences can thus be viewed as differing in the import they place on perceptual features in grounding aesthetic preference for a given stimulus. While there has been some discussion of the features that differentiate natural and urban visual scenes (Berman et al., 2014), including some discussions of whether preference for visual nature scenes can be thought of as perceptually driven (Kardan, Demiralp, et al., 2015), in audition it is presently unclear which perceptual features differentiate nature from urban soundscapes, as well as how these features relate to aesthetic preference. As such, the present set of experiments aimed to quantify the relative evidence in favor of the *inherent-organization* versus the *conceptual–organization* hypotheses in the context of aesthetic judgments of natural and urban sounds.

One possible explanation for why nature sounds may be aesthetically preferred, which may be unique to the auditory modality, relates to perceived loudness—that is, nature sounds are aesthetically valued because they are thought to provide a quiet respite from typically loud urban noises (Mace, Bell, & Loomis, 2004). Indeed, urban noises in the context of noise pollution (e.g., Cohen & Weinstein, 1981) are reported as highly annoying (Basner et al., 2014), which may have subsequent consequences for measures of health and well-being. The association between loudness and preference, however, may not completely explain individuals' aesthetic preference for nature sounds. When aesthetically rating nature and urban soundscapes presented at the same normalized amplitude, listeners still exhibit a robust preference for nature soundscapes (Van Hedger et al., 2018). Furthermore, many individuals seek out the sounds of nature in situations where loudness can be rapidly adjusted (e.g., through websites or apps), and recent research has demonstrated that such presentations of nature sounds can lead to improvements in mood and workplace productivity (e.g., DeLoach, Carter, & Braasch, 2015). This, in turn, suggests that the aesthetic preference for such stimuli may exist independently of loudness cues and may not be entirely tied to the experience of these sounds in their real-world environments.

Why might nature be preferred in situations where loudness is controlled? One possibility, which would be predicted under the *inherent–organization hypothesis*, is that listeners inherently prefer the acoustic features (such as spectral centroid, or entropy) that may differentiate nature from urban sound environments, perhaps due to an evolutionary preference for the features of natural environments. This kind of explanation suggests that even in situations where a sound is not identified as coming from nature, listeners may still show an aesthetic preference if the sound is acoustically reminiscent of nature. A second possibility, which would be predicted under the *conceptual-organization hypothesis*, is that the aesthetic preference for nature depends on the extent to which sounds can be recognized and conceptually associated with nature (e.g., birdsong). Under the *conceptual–organization hypothesis*, the relationship between listeners' aesthetic preferences and acoustic features should critically depend on the extent to which a sound can be recognized and associated with nature. Furthermore, listeners may not aesthetically prefer sounds that are, acoustically speaking, reminiscent of nature but not recognized as coming from nature.

To adequately test these hypotheses, one must vary the degree to which a sound can be recognized while also preserving the general perceptual features within each sound. This is accomplished in this paper using two complementary approaches. The first approach (Experiments 1–2) is to manipulate recordings of nature and urban sounds in a manner that obscures the conceptual identification of the sound-object (e.g., bird, car) but largely preserves the acoustic features of each sound. Experiment 1 serves as a "pre-study" in that it confirms that nature and urban sounds can be manipulated in this manner. Experiment 2 then uses these nature and urban sounds to directly test the *inherent–organization* and *conceptual–organization* hypotheses. The second approach (Experiment 3) is to generate noises that acoustically overlap with nature and urban sounds but, by definition, do not contain any objective conceptual ties to either nature or urban sound-objects.

Assessing the degree to which the aesthetic preference for nature sounds is acoustically versus conceptually based will inform our understanding of several benefits associated with nature sounds, including perceived attention restoration and stress recovery (Ratcliffe, Gatersleben, & Sowden, 2013), improved post-operation recovery (Bauer et al., 2011), and neural connectivity changes reflecting an increase in parasympathetic activity (Gould van Praag et al., 2017), as all of these benefits could be argued to have an aesthetic component. Beyond providing a frame for interpreting the myriad of restorative elements associated with nature, the present approach will also inform our understanding of the extent to which aesthetic preferences for nature may be innately or evolutionarily grounded as opposed to constructed by a particular historical context.

2. Experiment 1: Development of stimulus materials

Experiment 1 assessed the degree to which participants could identify the sound-objects contained within stereotypic nature and urban environments. The sounds were either unaltered in presentation, time-domain scrambled, or presented as a thinslice (100 ms). These two manipulations were chosen because they were hypothesized to impair the recognition of sound-objects. Scrambling has been used previously in visual paradigms to obscure the recognition of objects but maintain lower-level perceptual features (Berman et al., 2014). Thinslicing has been used previously in audition (e.g., Krumhansl, 2010) to demonstrate that object-level information is greatly attenuated (e.g., identifying the title or artist from 300 ms of a popular recording), though thin slicing sound "textures" such as typical nature and urban soundscapes should not, in theory, substantially change the spectral profiles of these sounds (cf. McDermott & Simoncelli, 2011).

The primary goal of Experiment 1 was to assess whether sound–object recognition—that is, identifying whether a sound came from a bird, waterfall, car, and so on. could be dissociated from the spectro-temporal profiles of each sound, as this is a necessary condition for the first approach we outline in testing the *inherent-organization* and *conceptual–organization hypotheses*. We predicted that participants would have greater difficulty identifying sound objects for manipulated soundscapes, conceptually replicating prior research (Berman et al., 2014; Kotabe, Kardan, & Berman, 2017). Moreover, we predicted that the spectral properties of the manipulated and non-manipulated sounds would largely overlap. Such findings would suggest that spectral properties of a sound are, to an extent, dissociable from recognizing the object that produced a sound.

2.1. Method

2.1.1. Participants

We recruited 150 participants on Amazon's Mechanical Turk (MTurk), and 145 participants completed the task. Participants were randomly assigned to condition in a between-subjects design: 49 participants (28 male, 20 female, 1 no response) heard soundscapes that had not been altered (Unaltered Group), 48 participants (29 male, 19 female) heard soundscapes that had been time-domain scrambled (Scrambled Group), and 48 participants (29 male, 19 female) heard soundscapes that had been thin sliced (Thin-Sliced Group). Participants had to have a minimum 90% positive rating from prior MTurk assignments, had to have completed a minimum of 10 previous assignments, and had to be residing in the United States to qualify. All participants consented to voluntary participation via the guidelines established by the Institutional Review Board at The University of Chicago and self-identified as having normal or diversely corrected hearing. All participants were compensated for their participation in the experiment.

2.1.2. Materials

We tested 80 soundscapes, which were downloaded from an online video sharing website. Half of the sounds were selected by the authors to represent typical soundscapes of nature, while the remaining sounds were selected to represent typical soundscapes of urban environments. The full database of sounds can be accessed through Open Science Framework. It should be noted that both the nature and urban soundscapes were not necessarily limited to a single sound-object (e.g., a single nature soundscape could contain sounds of insects, running water, and wind). We sampled multiple sources of nature and urban sounds, as we wanted to increase the potential for our analyses to generalize beyond specific sound classes (e.g., just birdsong). To verify that the soundscapes selected by the authors were indeed considered by others to be typical of nature and urban environments, we recruited an additional 50 participants from MTurk to provide "naturalness" ratings for the soundscapes on a 7-point scale (1: very urban, 7: very natural). The results of this validation test, which are available on Open Science Framework, show clear separation of nature and urban soundscapes. In fact, the naturalness ratings for these two classes of sounds were completely non-overlapping; that is, the highest-rated urban sound was still rated lower than the lowest-rated natural sound (Fig. S2). We interpret these findings as clear evidence that the selected soundscapes were representative of nature and urban environments.

The time-domain scrambled sounds were manipulated using the "scramble" toolbox in Matlab (Mathworks: Natick, MA). Each sound file was chopped into a set of short (25 ms) windows with a 50% overlap and was tapered with a raised cosine window. The 25 ms windows were then shuffled and re-overlapped within a 250 ms radius. This process preserved the long-term spectral features of each sound but effectively destroyed the spectro-temporal transitions within each soundscape. The thin-sliced sounds were taken from the beginning of each full 5-s sound clip with the exception of some birdsong stimuli, in which there was no audible birdsong at the beginning of the clip. All sounds were root mean square (RMS) normalized and digitized at a 44.1 kHz sampling rate with a 16-bit depth. The experiment was coded in jsPsych (De Leeuw, 2015).

2.1.3. Acoustic analyses

For each sound, we calculated five acoustic features: (a) spectral entropy, (b) spectral flatness, (c) spectral skewness, (d) spectral centroid, and (e) dominant frequency. These features were calculated with the *seewave* package in R (Sueur, Aubin, & Simonis, 2008). These features were not specifically selected with respect to the hypotheses of the experiment. Rather, the features were meant to represent the spectral profiles of our sounds in a non-redundant manner.

The spectral centroid represents the spectral center of gravity and is perceptually related to the "brightness" of a sound. Spectral flatness is obtained by comparing the geometrical mean and the arithmetical mean of the spectrum and can be thought to represent the extent to which a sound is heard as "noisy" versus "tonal." Spectral entropy is also a "noisiness" measure ranging from zero to one but is calculated differently (assessing Shannon entropy) and thus is not redundant with spectral flatness. Spectral skewness is a measure of asymmetry, with negative values indicating left skew (lower frequencies), values of zero indicating no skew, and positive values indicating right skew (higher frequencies). Dominant frequency represents the frequency with the greatest power at the particular sampled time window. We averaged across all time windows to obtain a single value for each sound.

2.1.4. Procedure

After providing informed consent, participants were presented with a 10-s sample of pink noise (RMS normalized to the same level as the soundscapes) and were instructed to adjust their computer volume to a comfortable listening level. Once this calibration had been completed, participants received instructions that they would hear 80 five-second sounds (Unaltered and Time-Domain Scrambled Groups) or 80 one-hundred-millisecond sounds (Thin-Sliced Group). Participants were additionally instructed that they would be asked to type, in a few words, what they thought made each sound. Thus, the wording of the task encouraged participants to respond with specific sound-objects (e.g., “bird”) rather than the broad distinction of “nature” or “urban” or descriptive terms such as “buzzing.” The 80 sounds were presented in a random order. The experiment was conducted online.

2.2. Results

2.2.1. Sound–object identification accuracy

To assess sound–object identification accuracy, we used participant responses in the Unaltered Group to establish correct responses for each sound file. The main stipulation was that each response had to appear more than once to be included as a correct response. Typos were manually corrected (e.g., “birdss” would be corrected to “birds”). We adopted this method of determining correct answers because our stimuli were recorded from real-world environments, and thus we did not want to limit artificially the number of acceptable answers. Rather, through looking at consistency of word use among participants, we were able to obtain a comprehensive yet consistent set of answers.

Using this method of determining correct answers, participants in the Unaltered Group correctly identified 83.6% (SD = 10.0%) of objects within soundscapes. Participants were more accurate at identifying unaltered nature soundscapes (M = 89.3%, SD = 10.4%) compared to unaltered urban soundscapes (M = 77.9%, SD = 12.0%; $t(48) = 7.75$, $p < .001$). As hypothesized, sound–object identification in the Scrambled Group (M = 54.3%, SD = 8.7%) was significantly worse than identification in the Unaltered Group ($t(95) = 15.33$, $p < .001$). Even though time-domain scrambling significantly impaired sound-object identification for both nature and urban soundscapes independently, we found an unexpected asymmetry in that scrambling appeared to impair urban sound-object identification more than nature sound-object identification. Compared to unaltered soundscape identification, accuracy for scrambled nature soundscapes dropped 17.1 percentage points to 72.2% ($t(95) = 7.78$, $p < .001$) while accuracy for scrambled urban soundscapes dropped 41.6 percentage points to 36.3% ($t(95) = 18.05$, $p < .001$). Finally, we observed poorest sound–object identification performance in the Thin-Sliced Group, with participants only accurately identifying 31.5% (SD = 14.8%) of sound-objects. This overall level of performance was significantly worse than sound–object identification in both the Unaltered ($t(95) = 20.32$, $p < .001$) and Scrambled ($t(94) = 9.18$, $p < .001$) groups. The pattern of nature soundscapes being easier to identify than urban soundscapes was present in the thin-sliced sounds but was only marginally supported (nature: M = 33.3%, SD = 15.8%; urban: M = 29.6%, SD = 16.6%; $t(47) = 1.92$, $p = .06$).

2.2.2. Acoustic differences between nature and urban soundscapes

For unaltered soundscapes, we found that our nature soundscapes had higher spectral centroids, greater spectral entropy, higher dominant frequencies, greater spectral flatness, and less spectral skew compared to urban soundscapes. These findings were generally consistent across both scrambling and thin slicing (see Table 1), with the sole exception being spectral skew, which did not significantly differ between scrambled nature and urban soundscapes.

To more formally assess the relationship between acoustic features and nature-urban category membership, we used a linear discriminant analysis (LDA) to predict whether a soundscape would be classified as natural or urban based on our five acoustic measures. The LDA was calculated using the {MASS} package in R, using leave-one-out cross validation. For unaltered soundscapes, the LDA correctly classified 72.5% of natural sounds and 92.5% of urban sounds (82.5% overall accuracy). For scrambled soundscapes, the LDA correctly classified 67.5% of natural sounds and 87.5% of urban sounds (77.5% overall accuracy). For thin-sliced soundscapes, the LDA correctly classified 80% of natural sounds and 92.5% of urban sounds (86.25% overall accuracy). These results demonstrate that scrambled and thin-sliced soundscapes could be classified into natural and urban categories based on their acoustic features in a comparable manner (within 5% overall accuracy) as the unaltered soundscapes. As such, the scrambled and thin-sliced sounds were, acoustically speaking, similarly predictive of category membership, even if

Table 1
Summary and comparison of acoustic features for nature and urban soundscapes

Feature	$M \pm SE$ (Nature)	$M \pm SE$ (Urban)	t	Cohen's d
Unaltered				
Centroid	4.16 ± 0.24	2.08 ± 0.17	7.10***	1.59
Entropy	0.74 ± 0.02	0.64 ± 0.02	4.17***	0.93
Skew	9.62 ± 1.52	28.29 ± 5.49	−3.28**	−0.73
Flatness	0.16 ± 0.01	0.09 ± 0.01	3.70***	0.83
DomHz	1.47 ± 0.26	0.32 ± 0.05	4.42***	0.99
Scrambled				
Centroid	4.16 ± 0.24	1.92 ± 0.16	7.86***	1.76
Entropy	0.75 ± 0.02	0.64 ± 0.02	4.24***	0.95
Skew	7.62 ± 5.64	11.10 ± 1.94	−1.62	−0.36
Flatness	0.18 ± 0.02	0.09 ± 0.01	4.05***	0.91
DomHz	1.49 ± 0.26	0.32 ± 0.05	4.41***	0.99
Thin-Sliced				
Centroid	3.97 ± 0.24	1.81 ± 0.15	7.74***	1.73
Entropy	0.73 ± 0.02	0.63 ± 0.02	3.54***	0.79
Skew	8.52 ± 0.97	11.37 ± 0.89	−2.16*	−0.48
Flatness	0.15 ± 0.02	0.08 ± 0.01	4.38***	0.98
DomHz	1.27 ± 0.23	0.28 ± 0.04	4.31***	0.96

Note. ***p < .001, **p < .01, *p < .05.

participants exhibited difficulty in labeling the sound objects contained within these soundscapes.

2.3. Discussion

There are two primary conclusions to draw from Experiment 1. First, the manipulations significantly impaired sound–object identification. Second, despite this decrease in sound–object identification, the measured acoustic features of these soundscapes remained relatively stable across all conditions (i.e., the features that differentiated unaltered nature and urban soundscapes also differentiated scrambled and thin-sliced soundscapes). These points, taken together, suggest (a) that our nature sounds were acoustically differentiable from our urban sounds, and (b) it is generally possible to preserve the spectral properties of a sound while increasing the difficulty of explicitly recognizing the object that produced the sound.

One unanticipated finding from the present experiment was the asymmetry in performance for scrambled sounds. Participants successfully classified 72.2% of scrambled nature sounds but only 36.3% of scrambled urban sounds. This asymmetry clearly shows that time-domain scrambling was more detrimental to urban versus nature sounds; however, it is unclear why this might be the case. Below, we offer two potential explanations.

First, there was some evidence that urban sounds with human voices and background music (e.g., café ambiance) were particularly difficult to identify when scrambled (Fig. S1). This could be because voice and music sounds are much more dependent on spectro-temporal transitions to be identified; indeed, nature sounds such as the ones used in this experiment are more likely to fit the definition of a "sound texture"—distinguished by relative temporal homogeneity—which, by definition, would be less susceptible to time-domain scrambling (see McDermott & Simoncelli, 2011). Second, it is possible that the acoustic features that differentiated our nature and urban sounds non-linearly influenced classification. For example, spectral centroid strongly predicted identification accuracy across our stimulus set, as well as within the nature sound set, but not within the urban sound set. As such, it is possible that sounds with certain acoustic profiles (e.g., high spectral centroids, high dominant frequencies) may be less susceptible to time-domain scrambling, perhaps because there are fewer categories for which the sound can be confused.

Overall, Experiment 1 demonstrates that these sounds are well suited for testing the *inherent-organization* and *conceptual–organization hypotheses*. By reducing sound-object recognition but maintaining the general spectral properties of each sound, the *inherent-organization hypothesis* would predict that listeners should display an aesthetic preference for nature sounds across all sound manipulations (e.g., listeners should exhibit positive relationships between aesthetic ratings and spectral entropy, as nature sounds were consistently higher on entropy across all sound manipulations). In contrast, the *conceptual–organization hypothesis* would predict that the relationship between aesthetic ratings and acoustic features should change as a function of sound-object identification—that is, the overall aesthetic valuing of nature (and the association between aesthetic ratings and acoustic properties) should change across sound manipulations.

3. Experiment 2: Aesthetic preference for nature and urban soundscapes

Experiment 2—which is divided into two subexperiments (2A and 2B)—tests whether listeners' aesthetic valuing of the nature and urban sounds from Experiment 1 can be better explained under the *inherent-organization hypothesis* versus the *conceptual–organization hypothesis*. In Experiment 2A, participants provided aesthetic ratings for the nature and urban sounds described in Experiment 1 (unaltered, scrambled, and thin-sliced, between participants). In Experiment 2B, participants provided aesthetic ratings for the thin-sliced sounds but were asked to categorize each sound as either natural or urban prior to their aesthetic rating.

Under the *inherent–organization hypothesis*, listeners should aesthetically prefer nature sounds even when the sounds are scrambled and thin-sliced (i.e., when the sounds are more difficult to recognize). Moreover, the relationship between aesthetic ratings and acoustic features should remain invariant regardless of sound condition, as this would suggest that the acoustic features that differentiate nature and urban sounds are driving this aesthetic preference. Finally, having participants categorize each sound as either natural or urban prior to providing an aesthetic rating should not substantially influence those aesthetic judgments, assuming they are driven by inherent acoustic differences that exist between nature and urban sounds.

Conversely, under the *conceptual–organization hypothesis*, we would predict that listeners' aesthetic preference for nature sounds might be attenuated for scrambled and thin-sliced sounds, as these sounds are more difficult to recognize and associate with nature and urban environments. Moreover, the relationship between aesthetic ratings and acoustic features should vary dynamically across sound manipulation (i.e., acoustic features may not invariantly relate to aesthetic preference). Finally, under the *conceptual–organization hypothesis* we would predict that the process of categorizing difficult-to-identify (thin-sliced) sounds as natural or urban should significantly change aesthetic preference, as participants may be able to bring conceptual information online to guide their aesthetic judgments.

3.1. Experiment 2A

3.1.1. Method

3.1.1.1. Participants: We recruited 150 individuals from MTurk to participate in the experiment, and 149 completed the task. Of the 149 participants, 50 (24 male, 26 female) listened to Unaltered Sounds, 50 (28 male, 22 female) listened to Scrambled Sounds, and 49 (30 male, 17 female, 2 no response) listened to Thin-Sliced Sounds. Similar to Experiment 1, participants had to have a minimum 90% positive rating from prior MTurk assignments, had to have completed a minimum of 10 previous assignments, and had to be residing in the United States to qualify. All participants consented to voluntary participation via the guidelines established by the Institutional Review Board at The University of Chicago and self-identified as having normal or diversely corrected hearing. All participants were compensated for their participation in the experiment.

3.1.1.2. Materials: The sounds were identical to those used in Experiment 1. The experiment was coded in jsPsych (De Leeuw, 2015).

3.1.1.3. Procedure: The consent procedure and volume calibration were identical to Experiment 1. After volume calibration, participants received instructions that they would hear 80 five-second sounds (Unaltered and Scrambled Sounds) or 80 one-hundred-millisecond sounds (thin-sliced sounds). Participants were also instructed that they would rate how much they liked each sound. The 80 sounds were presented in a random order. Participants provided their aesthetic judgments on a 7-point scale, with “1” representing “strongly dislike,” “4” representing “neither dislike nor like,” and “7” representing “strongly like.” Once participants heard and rated all 80 sounds, they were asked to provide basic demographic information.

3.1.2. Results

3.1.2.1. Aesthetic ratings: To assess how the aesthetic ratings for nature and urban soundscapes changed as a function of sound manipulation, we constructed a 2 (category: nature, urban) × 3 (sound manipulation: unaltered, scrambled, thin-sliced) mixed ANOVA. Mean aesthetic rating was the dependent variable, while category was a within-participant factor and manipulation was a between-participant factor.

In this analysis, we found a significant main effect of manipulation ($F(2, 146) = 7.50$, $p < .001$), in which overall aesthetic ratings decreased as a function of sound identification (i.e., unaltered soundscapes were higher than scrambled soundscapes, which in turn were higher than thin-sliced soundscapes). We also found a significant main effect of category ($F\ (1,\ 146) = 290.1$, $p < .001$), in which participants rated nature soundscapes higher than urban soundscapes. While this finding seems to support the *inherent-organization hypothesis*, we also found a significant interaction between category and manipulation ($F(2, 146) = 115.8$, $p < .001$), whereby the nature-related aesthetic preference was robustly observed in the Unaltered Group (Nature $M \pm SE$: 5.40 ± 0.11; Urban $M \pm SE$: 3.18 ± 0.11), attenuated in the Scrambled Group (Nature $M \pm SE$: 4.34 ± 0.10; Urban $M \pm SE$: 3.49 ± 0.09), and not present in the Thin-Sliced Group (Nature $M \pm SE$: 3.82 ± 0.11; Urban $M \pm SE$: 3.81 ± 0.11). This interaction (see Fig. 1) strongly supports the *conceptual–organization hypothesis*, as the relative aesthetic preference for nature soundscapes was reduced as sound identification was reduced.

From these analyses, however, it is difficult to tell whether the observed attenuation in aesthetic ratings between nature and urban soundscapes is truly a result of increasing the difficulty of sound identification, as opposed to an artifact of the acoustic manipulations themselves. Put another way, time-domain scrambling and thin-slicing recordings may render aesthetic judgments meaningless, with participants responding arbitrarily (which,

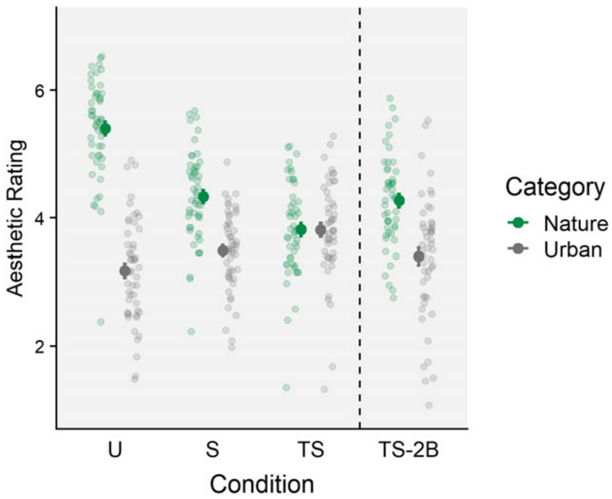

Fig. 1. Mean plots of aesthetic ratings across soundscape type (nature or urban) for each of the sound conditions. While participants had clear aesthetic preferences for unaltered (U) nature soundscapes compared to urban soundscapes, this preference was attenuated for scrambled (S) soundscapes and not present for thin-sliced (TS) soundscapes. Given the graded conceptual identification from Experiment 1, this pattern of results is what would be expected under the *conceptual-organization hypothesis*. In further support of this hypothesis, encouraging participants to think about the thin-sliced soundscapes in terms of nature or urban categories restored the nature-related aesthetic preference (TS-2B).

when averaged, would result in mean ratings close to the mid-point of the scale) because aesthetic judgments of such sounds might be difficult. If this were the case, then we would predict an attenuation of any relationship between our acoustic features and aesthetic judgments as a function of soundscape identification. This is assessed in the next section.

3.1.2.2. Correlations of acoustic features with aesthetic ratings: Based on the *inherent-organization hypothesis*, we would expect positive relationships between aesthetic ratings and (a) spectral centroid, (b) spectral entropy, (c) dominant frequency, and (d) spectral flatness, as well as a negative relationship between aesthetic ratings and spectral skew. Critically, these relationships would need to persist across all sound conditions, as this would support the argument that inherent aesthetic preferences are tied to these acoustic properties, regardless of the ease with which a sound can be identified as coming from nature.

Unaltered Sounds: As depicted in Fig. 2A, we found significant positive correlations between aesthetic ratings and (a) median spectral centroid (r = .53, p < .001), (b) spectral entropy (r = .31, p = .005), (c) dominant frequency (r = .45, p < .001), and (d) spectral flatness (r = .37, p < .001). Spectral skew was significantly negatively correlated with aesthetic ratings (r = −.31, p = .006). Scrambled Sounds: As depicted in Fig. 2B, we did not find a significant correlation between aesthetic ratings and any of our acoustic features. Thin-Sliced Sounds: As depicted in Fig. 2C, we found *negative* correlations between aesthetic ratings and (a) spectral entropy ($r = -.72$, $p < .001$), (b) spectral centroid ($r = -.30$, $p = .008$), and (c) spectral flatness ($r = -.27$, $p = .016$). Spectral skew was positively associated with aesthetic ratings ($r = .47$, $p < .001$).

These correlations suggest that participants did not merely respond randomly—at least not in the thin-sliced conditions. Rather, it appears that participants underwent a strategic shift when determining aesthetic qualities for nature and urban soundscapes, based on the degree to which the objects within a soundscape could be identified. In fact, the strongest relationship between aesthetic ratings and an acoustic feature was observed in the thin-sliced condition, and moreover the direction of these relationships was the *opposite* of

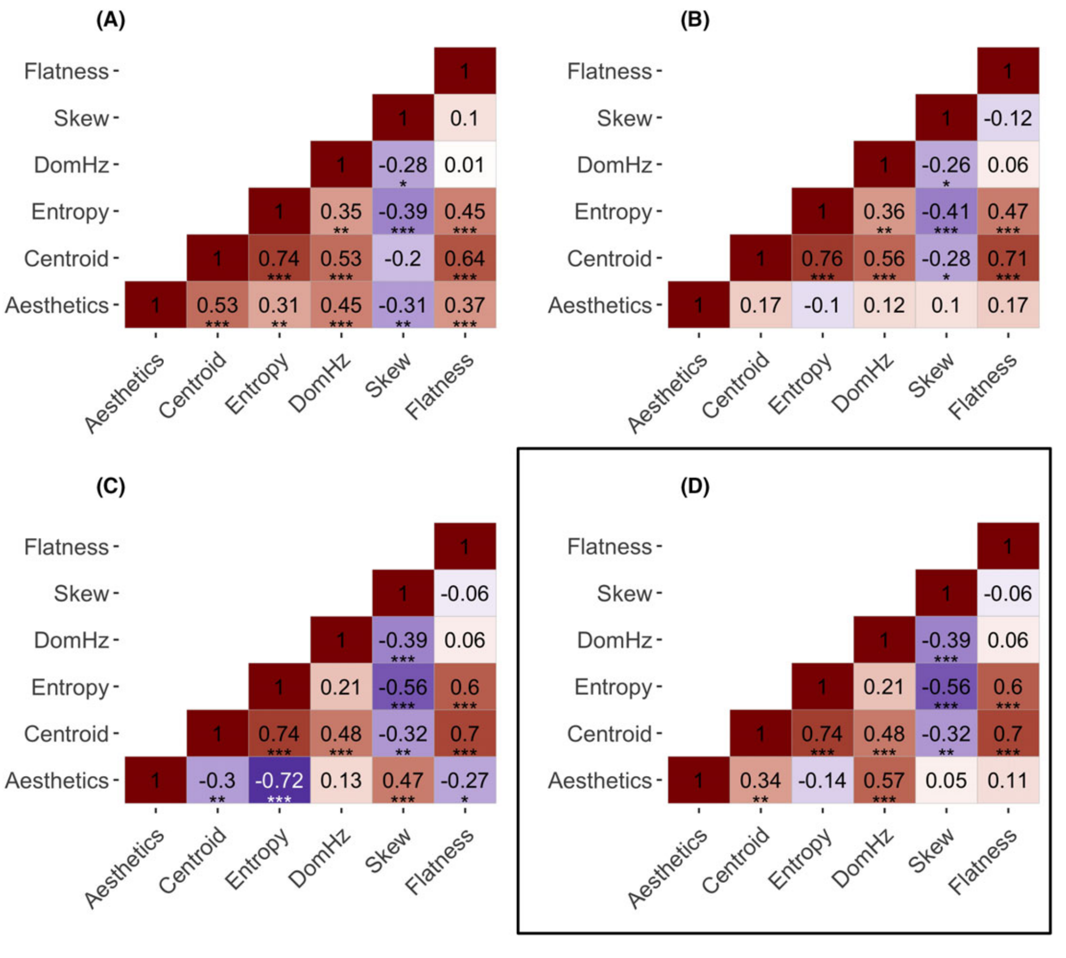

Fig. 2. Correlation matrices for aesthetic ratings and acoustic properties. Panels A-C represent unaltered, scrambled, and thin-sliced sounds, respectively (Experiment 2A). Panel D represents the same thin-sliced sounds as those plotted in C, with the only difference being that the aesthetic ratings were explicitly framed in terms of nature and urban categories (Experiment 2B). ***$p < .001$, **$p < .01$, *$p < .05$.

what we would expect under the *inherent-organization hypothesis*. For example, the relationship between entropy and aesthetics reverses in direction for difficult-to-identify thin slices, providing evidence that features alone are not predictive of aesthetic preference, but contribute differentially based on the ability to recognize and associate a sound with either a nature or urban environment.

3.2. Experiment 2B

3.2.1. Method

3.2.1.1. Participants: We recruited 49 individuals (34 male, 15 female) on MTurk to participate in the experiment, as this was our sample size for the Thin-Sliced Group from Experiment 2A. Similar to the previous experiments, participants had to have a minimum 90% positive rating from prior MTurk assignments, had to have completed a minimum of 10 previous assignments, and had to be residing in the United States to qualify. All participants consented to voluntary participation via the guidelines established by the Institutional Review Board at The University of Chicago, and all participants were compensated for their participation in the experiment.

3.2.1.2. Materials: The sounds were identical to those used in the Thin-Sliced Group in Experiment 2A. The experiment was coded in jsPsych (De Leeuw, 2015).

3.2.1.3. Procedure: The consent and volume calibration were identical to Experiments 1 and 2A. After volume calibration, participants received instructions that they would hear 80 one-hundred-millisecond sounds. Importantly, participants were told that these short sounds were sampled from longer soundscapes that either came from nature or urban environments. Additionally, prior to making an aesthetic rating, participants made a forced-choice judgment without feedback as to whether they believed each thin slice came from a nature or urban environment (by pressing the “n” or “u” key). Participants rated how much they liked each sound on the same 7-point scale used in Experiment 2A. Once participants heard and rated all 80 sounds, which were presented in a random order, they were asked to provide basic demographic information.

3.2.2. Results

3.2.2.1. Aesthetic ratings: Participants displayed a robust aesthetic preference for nature soundscapes compared to urban soundscapes ($t(48) = 7.44$, $p < .001$, $d = 1.06$), which is notable given that the same sounds did not exhibit a nature-related aesthetic preference in Experiment 2A. The mean rating for nature soundscapes was 4.27 (*SE:* 0.10) while the mean rating for urban soundscapes was 3.40 (*SE*: 0.14). The aesthetic ratings for nature and urban sounds are plotted in Fig. 1 (far right).

3.2.2.2. Category accuracy: Participants were significantly above chance at determining whether a thinslice came from a nature or urban environment ($M = 70.6\%$, $SD = 11.0\%$, $t(48) = 13.09$, $p < .001$). This result is not surprising given the results from Experiment

1, which demonstrated that the recognition of specific sound-objects was greatly attenuated but not entirely eliminated for thin-sliced sounds. However, it is important to note that this level of accuracy is also well below ceiling performance. As such, these results lend support to the conjecture that participants in the Thin-Sliced Group from Experiment 2A—who were not told anything about the nature or urban origin of the sounds—were likely making aesthetic judgments without explicit awareness that the sounds came from nature or urban environments.

We additionally assessed how category accuracy—split by nature and urban sounds—related to aesthetic preference ratings. Under the *conceptual–organization hypothesis*, we would predict that category accuracy should positively relate to aesthetic ratings for nature sounds (as the less ambiguous sounds should be more aesthetically preferred), whereas category accuracy should negatively relate to aesthetic ratings for urban sounds (as the less ambiguous sounds should be less aesthetically preferred). These predictions were strongly confirmed. Category accuracy was strongly and positively related to aesthetic ratings for nature sounds ($r(38) = .91$, $p < .001$), whereas category accuracy was strongly and negatively related to aesthetic ratings for urban sounds ($r(38) = -.67$, $p < .001$). This relationship is plotted in Fig. 3A.

This strong relationship between category accuracy and aesthetic ratings, however, may have been driven by response bias—that is, participants who just labeled a sound as coming from nature may provide a high aesthetic rating simply because of this categorization outcome, not necessarily because the nature/urban framework provided participants with an insight into the identity of the sound (e.g., bird song) which then made the nature stimuli actually *sound* more aesthetically pleasing. This is a subtle but important distinction. In order to test whether this relationship could be explained by response bias, we examined the relationship between aesthetic ratings in this experiment and aesthetic

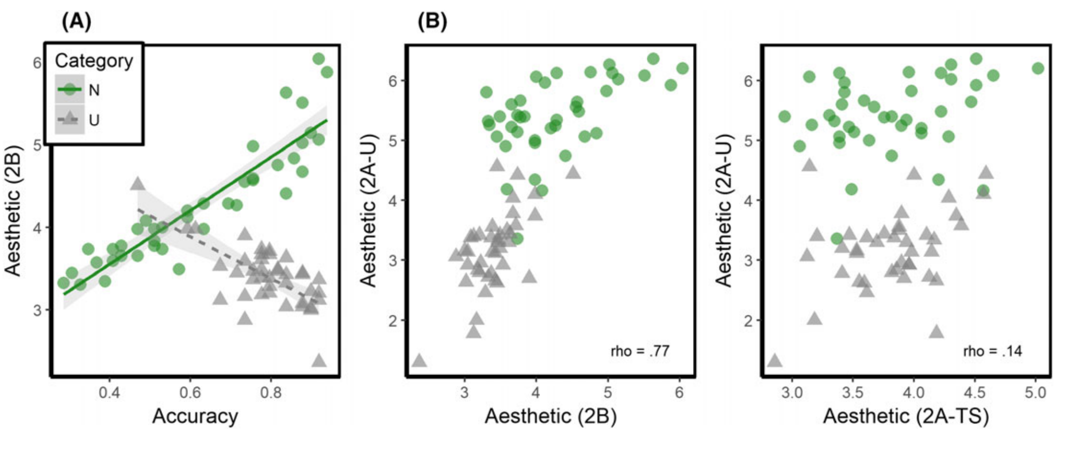

Fig. 3. Relationship between category identification accuracy and aesthetic ratings in this experiment, split by nature (N) and urban (U) sounds (Panel A). Panel B (left) demonstrates that the aesthetic ratings from this experiment are highly related to the aesthetic ratings from the unaltered soundscapes from Experiment 2A. This relationship does not hold when using the thin-sliced sounds from Experiment 2A (Panel B, right).

ratings from the unaltered soundscapes from Experiment 2A, which by definition had high levels of recognition. If the rank-ordering of these two sound sets in terms of aesthetic ratings is similar, this would suggest that participants were not simply responding with an arbitrarily higher or lower aesthetic rating after labeling a sound as nature or urban, respectively; rather, such a relationship would suggest that the provided nature/urban framework allowed participants to better understand the identity of the sounds, which was then used make an aesthetic judgment.

We used a Spearman's Rank-Order Correlation to assess the relationship between aesthetic ratings from Experiment 2A and this experiment. The correlation was positive and significant ($\rho(38) = .77$, $p < .001$), suggesting that the nature/urban framing of this experiment resulted in aesthetic ratings that were rank-ordered in a highly similar manner as the easy-to-identify, unaltered soundscapes. In contrast, the aesthetic ratings for the thin-sliced sounds from Experiment 2A (which received no nature/urban framing) were not significantly related to the unaltered soundscape ratings ($\rho(38) = .14$, $p = .213$). As such, these results suggest that introducing the nature/urban framework in this experiment allowed participants to accurately interpret the sound-objects in the thin slices, which then largely determined participants' aesthetic ratings. This in turn suggests that the aesthetic preference for nature sounds in this experiment was likely driven by recognizing and associating particular sounds with nature and urban objects, not simply rating a sound higher aesthetically because of the categorization process.

3.2.2.3. Correlations of acoustic features with aesthetic ratings: The framing our thin-sliced soundscapes in terms of nature or urban category membership significantly influenced the relationship between our acoustic features and aesthetic ratings. Most notably, the relationship between spectral entropy and aesthetic ratings—which was strongly negative in Experiment 2A ($r = -.71$, $p < .001$)—was attenuated in this experiment such that the relationship was no longer significant ($r = -.14$, $p = .21$). Moreover, this change in correlation strength between experiments was significant ($z = -4.76$, $p < .001$).

We additionally found evidence that aesthetic ratings were positively related to both median spectral centroid ($r = .34$, $p = .002$) and dominant frequency ($r = .57$, $p < .001$), which is notable for two reasons. First, these positive relationships were not significantly present for thin-sliced sounds in Experiment 2A (centroid: $r = -.30$; dominant frequency: $r = .13$), and the correlations in this experiment marked a significant increase from these values (centroid: $z = 4.21$, $p < .001$; dominant frequency: $z = 3.21$, $p = .001$). Second, these acoustic features—which differentiated our nature and urban soundscapes—were significantly and positively related to aesthetic ratings for our unaltered soundscapes in Experiment 2A, which by definition had high conceptual identification.

3.3. Discussion

The results from Experiment 2 strongly support the *conceptual-organization hypothesis*. For thin-sliced sounds, which were the hardest to recognize as containing nature or urban sound-objects (Experiment 1), participants did not display an aesthetic preference for nature sounds in Experiment 2A. This elimination of aesthetic preference, moreover, was accompanied by a shift in the relationship between acoustic features and aesthetic ratings. More specifically, aesthetic ratings were negatively associated with spectral centroid, entropy, and flatness, although our nature sounds were significantly *higher* on these features compared to urban sounds, and aesthetic ratings were positively associated with spectral skew, although our nature sounds were significantly *lower* on skew compared to urban sounds.

This dynamic relationship between acoustic features and aesthetic ratings suggests that the conceptual identification of a sound as either natural or urban fundamentally alters the way in which a listener uses acoustic cues to arrive at an aesthetic judgment. Indeed, this was causally tested and supported by Experiment 2B, in which simply categorizing thin-sliced sounds as either natural or urban prior to providing an aesthetic judgment led to an aesthetic preference for nature as well as shifted the relationship between aesthetic ratings and acoustic features. These findings also did not appear to be driven by a response bias (e.g., participants giving a sound a high aesthetic rating solely because they had just provided a "nature" label), as rank-ordering the sounds of this experiment on their aesthetic rating strongly mapped onto the aesthetic ratings of the unaltered soundscapes from Experiment 2A. As such, the most parsimonious explanation of the results is that categorizing the thin-sliced sounds into nature or urban categories provided listeners with a meaningful interpretative framework for organizing their aesthetic judgments, ultimately yielding aesthetic judgments that were highly related to those provided for unaltered nature and urban sounds.

One concern with the results thus far is that the support for the *conceptual-organization hypothesis* rests on a stimulus set that, despite the manipulations of scrambling and thin-slicing, have objective ties to nature or urban environments. As such, a stronger test of the *conceptual-organization* and *inherent-organization hypotheses* is to use a novel sound set with no objective ties to nature. If aesthetic ratings of these sounds correlate with acoustic features in a manner similar to the unaltered soundscapes from Experiment 2A, this would support the *inherent-organization hypothesis*. If, however, aesthetic ratings of these sounds correlate with acoustic features in a manner similar to the thin-sliced soundscapes from Experiment 2A, this would support the *conceptual-organization hypothesis*.

4. Experiment 3: Aesthetic preference for generated noises

In Experiment 3—which is divided into two subexperiments (3A and 3B)—listeners aesthetically rated a set of computer-generated aperiodic noises. These noises were not generated by sound-objects typically associated with nature or urban environments, and as such, allow for a more construct-valid test of the *conceptual–organization* and *inherent–organization hypotheses*. In Experiment 3A, participants provided aesthetic ratings for these generated noises. In Experiment 3B, participants were asked to categorize each noise as either "natural" or "urban" prior to providing an aesthetic rating.

If the acoustic features that differentiated the nature and urban sounds from Experiments 1–2 (e.g., higher spectral centroid, higher spectral entropy) positively relate to aesthetic ratings, this would suggest that these features may relate to aesthetic preference independently of a nature/urban framework. If this were the case, we would additionally predict in Experiment 3B that classifying each sound as "natural" or "urban" should not significantly influence aesthetic ratings. Such a pattern of results, taken together, would be consistent with the *inherent-organization hypothesis*. If, however, the acoustic features that differentiate our nature and urban sounds negatively relate (or do not relate at all) to aesthetic ratings, this would suggest that the acoustic features of nature are not preferred in a neutral context (i.e., with no objective associations to nature or urban sound-objects). If this were the case, we would additionally predict in Experiment 3B that sounds that are classified as "natural" should display a significantly different relationship between aesthetic ratings and acoustic features compared to sounds that are classified as "urban," despite neither classification being objectively correct. Such a pattern of results, taken together, would be consistent with the *conceptual–organization hypothesis*.

4.1. Experiment 3A

4.1.1. Method

4.1.1.1. Participants: We recruited 50 participants from MTurk for the experiment, and 48 individuals (24 male, 23 female, 1 no response) completed the task. The recruitment and treatment of participants was identical to Experiments 1–2, in that participants had to have a minimum 90% positive rating from prior MTurk assignments, had to have completed a minimum of 10 previous assignments, and had to be residing in the United States to qualify. All participants consented to voluntary participation via the guidelines established by the Institutional Review Board at The University of Chicago, and all participants were compensated for their participation in the experiment.

4.1.1.2. Materials: We generated 75 variants of white, pink, and brown noise. White noise is characterized by a flat frequency spectrum (i.e., equal power in any frequency band of a given bandwidth). Pink noise is characterized by equal amounts of energy per octave (e.g., the energy from 100 to 200 Hz will equal the energy from 1,000 to 2,000 Hz). Pink noise is also referred to as $1/f$ noise because the power spectrum density (PSD), has power that is inversely proportional to the frequency (f) of the signal. Pink noise has a 3-dB attenuation per octave. Brownian (brown) noise is characterized by random walk motion and has a more extreme attenuation of higher frequency components compared to pink noise (with a PSD of $1/f^2$). Brown noise has a 6-dB attenuation per octave. Each noise was one second in duration.

We chose brown, pink, and white noises, as opposed to another class of sounds (e.g., complex tones) for two reasons. First, these noises by definition include variability on the acoustic dimensions of interest (see Fig. 4 for a comparison of noise and soundscape stimuli). Second, aperiodic noises are found in both nature and urban settings (e.g., wind rustling through trees, television static). Of the 75 noise stimuli, 54 were generated in

Adobe Audition (Adobe Systems: San Jose, CA) using the "generate noise" function (18 brown, 18 pink, and 18 white). We applied six kinds of permutations to these noise stimuli (unaltered, high-pass filter of frequencies >500 Hz, low-pass filter of frequencies <500 Hz, band-pass filter of frequencies between 500 and 5,000 Hz, band-stop filter of frequencies between 500 and 5,000 Hz, and a comb filter with peaks at octaves of 50 Hz up to 9,500 Hz). This resulted in six versions of each of the noise types. We then frequency modulated these sounds by 0 Hz, 0.5 Hz, or 0.2 Hz, resulting in 54 total sounds (three noise types × six filters × three modulations). These permutations were chosen to generate acoustic diversity among the classes of noise; they were not hypothesized to make the noises sound particularly natural or urban.

The remaining 21 noises were generated using Java source code that created a continuum of pink noise. Given that canonical pink noise can be described as containing a PSD of $1/f$, multiplying f by an exponent allows for different variations of pink noise (i.e., different PSDs). Specifically, we generated 9 intermediary noises between pink and white noise (in which the exponent ranged from 0.1 to 0.9 in increments of 0.1), 9 intermediary noises between pink and brown noise (in which the exponent ranged from 1.1 to 1.9 in increments of 0.1), and the 3 canonical versions of each noise (white, pink, and brown). These 21 noise stimuli were not altered in any way (i.e., using any of the filters or modulations that were applied to the Adobe Audition noises). All stimuli were RMS normalized. Additionally, all noise stimuli were presented in stereo at a 44.1 kHz sampling rate and 16-bit depth. The experiment was coded in jsPsych (De Leeuw, 2015).

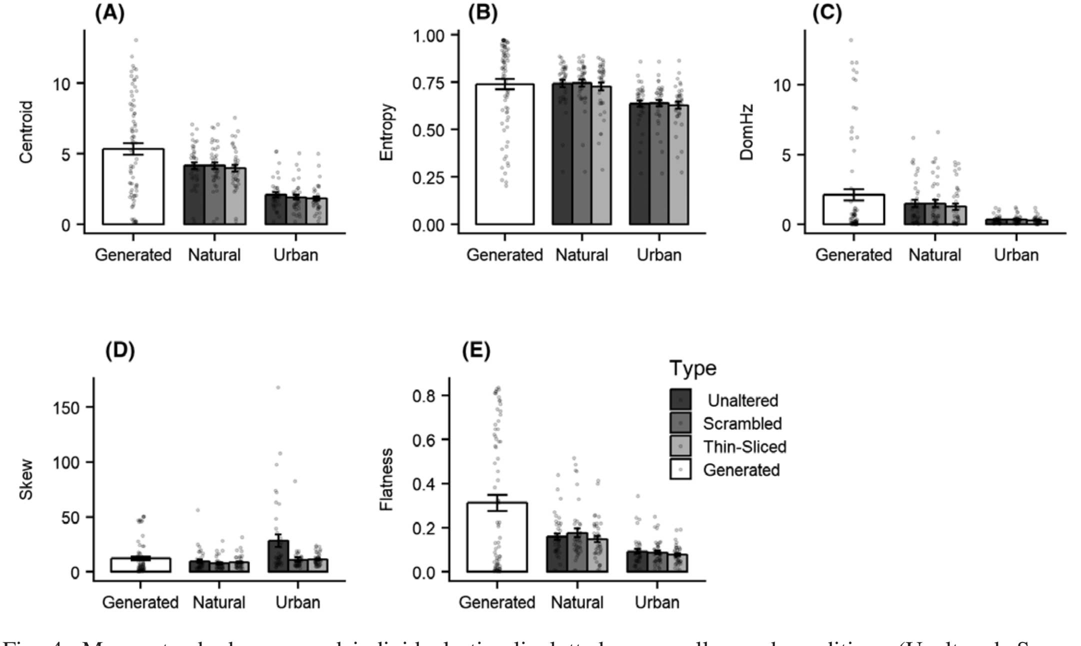

Fig. 4. Mean, standard error, and individual stimuli plotted across all sound conditions (Unaltered, Scrambled, and Thin-Sliced of Experiments 1–2, and Generated from Experiment 3) for (A) spectral centroid, (B) spectral entropy, (C) dominant frequency, (D) spectral skew, and (E) spectral flatness.

4.1.1.3. Acoustic analyses: We analyzed each of the 75 sounds using the same procedures as described in Experiment 1. Thus, each sound had the following five attributes: spectral centroid, dominant frequency, spectral entropy, spectral flatness, and spectral skew.

4.1.1.4. Procedure: The procedure was identical to Experiment 2A, with the exception that participants were told that they would be hearing 75 one-second sounds.

4.1.2. Results

As displayed in Fig. 5A, there was a negative relationship between aesthetic ratings and (a) spectral entropy ($r = -.84$, $p < .001$), (b) spectral centroid ($r = -.67$, $p < .001$), spectral flatness ($r = -.66$, $p < .001$), and dominant frequency ($r = -.30$, $p = .009$). There was a positive relationship between aesthetic ratings and spectral skew ($r = .65$, $p < .001$). For comparison, the direction of these correlations was identical to the thin-sliced soundscapes without a conceptual label from Experiment 2A (Fig. 2C), except for dominant frequency which was negative in this experiment and weakly positive in Experiment 2A.

4.2. Experiment 3B

4.2.1. Method

4.2.1.1. Participants: We recruited 48 individuals (23 male, 24 female, 1 no response) on MTurk to participate in the experiment, as this was our sample size for Experiment 3A. Participant recruitment and treatment was identical to the previous experiments, in

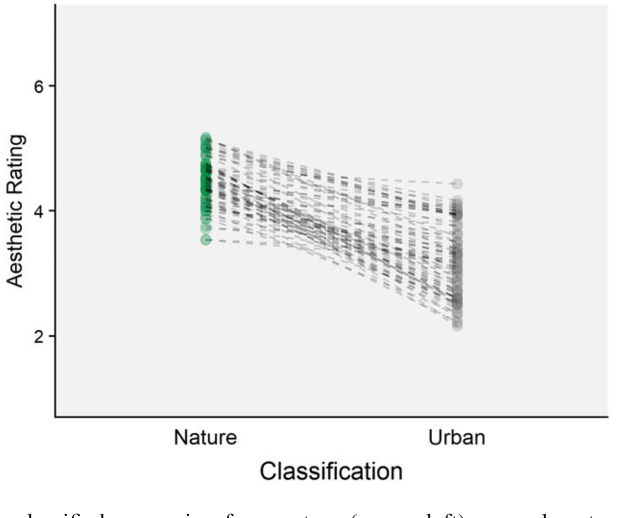

Fig. 5. Noises that were classified as coming from nature (green, left) were almost exclusively given higher aesthetic ratings compared to when the same noise was classified as urban (grey, left). The dashed lines link each individual noise sample.

that participants had to have a minimum 90% positive rating from prior MTurk assignments, had to have completed a minimum of 10 previous assignments, and had to be residing in the United States to qualify. All participants consented to voluntary participation via the guidelines established by the Institutional Review Board at The University of Chicago, and all participants were compensated for their participation in the experiment.

4.2.1.2. Materials: The sounds were identical to those used in Experiment 3A. The experiment was coded in jsPsych (De Leeuw, 2015).

4.2.1.3. Procedure: The consent and volume calibration were identical to the previous experiments. After volume calibration, participants received instructions that they would hear 75 one-second sounds. Similar to Experiment 2A, participants were told that these sounds were sampled from longer soundscapes that either came from nature or urban environments. Additionally, prior to making an aesthetic rating, participants made a forced-choice judgment without feedback as to whether they believed each sound came from a nature or urban environment (by pressing the “n” or “u” key). Participants rated how much they liked each sound on the same 7-point scale as used previously. Once participants heard and rated all 75 sounds, which were presented in a random order, they were asked to provide basic demographic information.

4.2.2. Results

4.2.2.1. Categorization of noises as natural versus urban: Given that the computer-generated noises did not have objective origins in natural or urban sound-objects, we first assessed the general distribution of category responses across our sounds. In particular, we measured the overall frequency of responding with a “nature” versus “urban” category label, as well as the distribution of these category judgments across our stimuli.

Overall, participants were approximately equally likely to classify sounds into nature (47.6%) and urban (52.4%) categories. These category judgments, moreover, appeared to be normally distributed across our stimuli (Shapiro-Wilk: $W = 0.987$, $p = .623$), with the 25th percentile corresponding to labeling a sound as natural 37.5% of the time and the 75th percentile corresponding to labeling a sound as natural 58.3% of the time. These results highlight the variability in classification judgments (i.e., relatively few sounds were consistently classified as nature or urban), which is perhaps not surprising considering that our stimuli did not actually originate from nature or urban environments.

4.2.2.2. Aesthetic ratings: The approach to analyzing aesthetic ratings was informed by the ambiguity of the sound set. Specifically, given that the tested sounds were not consistently heard as natural or urban, we assessed how aesthetic ratings *for identical sounds* changed as a function of classification. For each sound, we calculated separate mean aesthetic ratings for when it was heard as natural (as opposed to urban). Noises that were classified as natural received an average aesthetic rating of 4.47 (*SD*: 0.39), whereas the same noises—classified as urban—received an average aesthetic rating of 3.19 (*SD*: 0.59), which represented a highly significant difference ($t(74) = 17.27$, $p < .001$). In fact, as represented in Fig. 5, virtually all the noises (74 of 75) were rated aesthetically higher when classified as natural versus urban.

4.2.2.3. Relationship to acoustic features: In assessing how aesthetic ratings related to our measured acoustic features, we used the same nature-averaged and urban-averaged aesthetic response for each sound as described in the previous section. This approach allows for a causal test of the hypothesis that classifying a sound as nature will attenuate the strong negative relationship between spectral entropy, spectral centroid, and spectral flatness, which was found in Experiment 2.

When the noises were classified as urban (Table 2), the relationship between aesthetic ratings and acoustic features was virtually identical to what was observed in Experiment 3A (i.e., when no nature/urban framework was provided). Specifically, we found negative relationships between aesthetic ratings and spectral entropy ($r = -.72$, $p < .001$), spectral centroid ($r = -.56$, $p < .001$), spectral flatness ($r = -.60$, $p < .001$), and dominant frequency ($r = -.26$, $p = .024$). Spectral skew was positively related to aesthetic ratings ($r = .59$, $p < .001$).

When the noises were classified as nature (Table 2), these negative relationships were attenuated. While we still observed a significant negative relationship between aesthetic ratings and spectral entropy ($r = -.37$, $p = .001$), this relationship was significantly attenuated compared to when the noises were labeled as urban ($z = 3.12$, $p = .002$). We did not find a significant relationship between aesthetic ratings and spectral centroid ($r = -.11$, $p = .347$), spectral flatness ($r = .03$, $p = .798$), or dominant frequency ($r = -.02$, $p = .865$). This attenuated relationship, moreover, was significant for spectral centroid ($z = 3.13$, $p = .002$) and spectral flatness ($z = 4.34$, $p < .001$) but not for dominant frequency ($z = 1.48$, $p = .139$). Spectral skew remained significantly positively related to aesthetic ratings ($r = .41$, $p < .001$), and the correlation strength was not significantly different from what was observed when the noises were heard as urban ($z = 1.45$, $p = .147$).

Table 2
Correlations between aesthetic ratings and acoustic features for generated noises in Experiments 3A and 3B

Feature	3A	3B (classified as urban)	3B (classified as nature)
Centroid	$-.67^{***}$	$-.56^{***}$	$-.11$
Entropy	$-.84^{***}$	$-.72^{***}$	$-.37^{***}$
DomHz	$-.30^{**}$	$-.26^{*}$	$-.02$
Skew	$.65^{***}$	$.59^{***}$	$.41^{***}$
Flatness	$-.66^{***}$	$-.60^{***}$	$.03$

Note. $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$.

4.3. Discussion

Are novel sounds, not specifically modeled after nature or urban sound environments, aesthetically judged based on similar acoustic features as actual sounds from nature and urban environments? The results from Experiment 3A suggest that the acoustic features found to differentiate actual nature and urban sounds were *inversely* related to aesthetic ratings for generated noises. Put in another way, this means that the generated noises that were more acoustically like our nature sounds from Experiments 1–2 were given *lower* aesthetic ratings. This pattern of results is similar to what was observed for thin-sliced sounds in Experiment 2A, which were also the hardest sounds to recognize. Furthermore, Experiment 3B demonstrates a causal effect of sound identification on aesthetic preference. Specifically, a simple manipulation, meant to encourage listeners to frame these ambiguous sounds in a natural/urban framework, significantly influenced aesthetic ratings. The *same* sounds were aesthetically preferred when they were classified as natural versus urban. Moreover, the relationship between aesthetic ratings and acoustic features (in particular, spectral centroid, spectral entropy, and spectral flatness) was significantly altered depending on whether the sounds were classified as natural or urban. As such, the results from Experiment 3 strongly support the *conceptual–organization hypothesis* and refute the *inherent-organization hypothesis*.

5. General discussion

The aesthetic preference for natural over urban environments appears ever-present, manifesting cross-culturally (e.g., Yang & Kaplan, 1990) and across age ranges (Balling & Falk, 1982; Moore, 1986, though see Meidenbauer et al., 2018). This aesthetic preference can be large enough that the distributions of aesthetic judgments for nature and urban stimuli sometimes barely overlap (Kaplan et al., 1972). Indeed, this was the case with the non-manipulated nature and urban soundscapes used in this study; the mean aesthetic rating for nature sounds was over 3 *SD* higher than the mean aesthetic rating for urban sounds. The consistently large effects found in empirical work, combined with over a century of philosophical treatments of the aesthetic value of nature (e.g., Muir, 1894), lend weight to the notion that the preference for natural environments is innate and part of our evolutionary history (e.g., Wilson, 1984).

Situated within this framework, the goal of this work was to systematically investigate the level of processing at which this aesthetic preference for nature emerges. More specifically, we sought to address whether aesthetic preferences for nature could be explained by the perceptual features that differentiate nature from urban environments (i.e., a relatively low level of processing), or whether observers needed to interpret perceptual input in a nature/urban framework in order to display a nature-related aesthetic preference (i.e., a relatively high level of processing). Answering this question has important implications for how we conceptualize the role of nature as a restorative environment, as well as for how we conceptualize the relationship between perceptual and aesthetic judgments more generally.

The present results demonstrate that nature sounds are aesthetically preferred over urban sounds only when they can be recognized and associated with nature—not because the acoustic features that differentiate nature and urban sounds are inherently preferred. In fact, when difficult-to-identify or completely ambiguous sounds were not given a nature/urban framework for interpretation, the relationship between identified acoustic features and aesthetic preference was the opposite of what one would expect under the *inherent–organization hypothesis*—that is, the acoustic features that were higher among nature sounds were negatively related to aesthetic ratings (and vice versa). This observation supports a dynamic, *conceptual organization* of the relationship between nature and aesthetic preferences and also draws a notable parallel to a conceptually related study using nature images (Kotabe et al., 2017).

Importantly for the *conceptual–organization hypothesis*, relatively simple changes to the experiment, in which participants had to classify these difficult-to-identify and ambiguous sounds as natural or urban prior to aesthetically rating the sounds, was sufficient to shift the way in which aesthetic responses related to acoustic features of the sounds. These findings are conceptually similar to a recent study (Haga, Halin, Holmgren, & Sörqvist, 2016) demonstrating that an ambiguous noise was rated higher on restorative potential if participants were told it came from a waterfall (as opposed to urban machinery or no label). While Haga and colleagues demonstrate that an identical sound can be heard as either natural or urban, which then has implications for how the sound is evaluated, the present set of experiments expand upon these findings in at least two critical ways.

First, by using a sound set that had objective ties to natural and urban environments (Experiments 1–2), these experiments provide compelling evidence that acoustically *unambiguous* nature sounds (i.e., clearly differentiated from urban soundscapes on several acoustic features) that are difficult to conceptually identify are not aesthetically preferred over urban sounds. This observation strengthens the argument that the preference for nature sounds is not driven by the acoustic features that differentiate nature and urban sounds on broad spectral properties. Second, through testing more than one auditory stimulus, we were able to assess how category knowledge shifted the relative weighting between acoustic features and aesthetic preferences. These findings are perhaps most critical as we found, somewhat paradoxically, that the sounds that were acoustically more "nature-like" were *less* aesthetically preferred when the sound could not be appropriately framed as either natural or urban.

Overall, then, the present results suggest that the aesthetic preference for nature sounds emerges at a higherlevel of processing: one in which observers have access to the relevant framework for identifying specific sound objects and associating them with natural or urban environments. In the next several paragraphs, we explore how these results inform our treatment of nature as a restorative environment, as well as our understanding of the relationship between perceptual and aesthetic judgments more broadly.

5.1. Implications for nature as a restorative environment

Nature is thought to be a restorative environment, in that experiences with nature have been shown to provide recovery from stress and mental fatigue, allowing individuals to meet the demands of their lives (e.g., Berto, 2014). However, the underlying processes that ground these restorative benefits—in addition to the importance of aesthetic experiences in this restorative process—are debated. In the next several paragraphs, we explore how the present findings may provide some testable predictions that could inform our understanding of the mechanisms through which nature acts as a restorative environment.

Generally speaking, aesthetic preference has been strongly associated with an environment's perceived restorative potential. For example, Purcell and colleagues found a highly positive correlation between preference for natural scenes and their perceived restorative potential (Purcell et al., 2001). The authors interpret these findings as evidence that restorative potential may be used as a frame of reference for making a preference judgment. However, given that restoration is a broad construct, it is possible that the relationship between aesthetic preference and restoration may depend on *what* is being restored by nature.

Two prominent theories have emerged to explain how nature may help to restore cognitive and affective functioning, respectively. Cognitive restoration is generally discussed within the framework of *attention restoration theory* (ART), which posits that nature provides an ideal environment within which bottom-up involuntary attention is softly captured, thereby allowing top-down directed attention a chance to rest and replenish (Kaplan & Kaplan, 1989). While there are several proposed reasons why nature is thought to provide an ideal restorative environment for directed attention, including richness, coherence, and providing an opportunity for an individual to experience a state of "being away" (Kaplan, 1995), aesthetic responses are not discussed as integral to this process. In fact, empirical work in support of ART has shown that individuals need not enjoy their experience with nature in order to display cognitive benefits, with variables such as self-reported affect and aesthetic preference of the nature exposure not relating to the observed changes in attentional performance (Berman et al., 2008, 2012; Stenfors et al., 2018; Van Hedger et al., 2018). These findings, in turn, have led to empirical work that aims to quantify the low-level perceptual differences between nature and urban scenes (Berman et al., 2014), with possible applications of incorporating perceptual elements of nature into humanmade structures in order to improve restorative potential.

In contrast, *stress reduction theory* (SRT) operates on an affective dimension, positing that nature provides restoration from stress through improving positive affect and reducing negative affect and anxiety (Ulrich, 1983). SRT is framed in a psycho–evolutionary perspective; that is, nature confers these affective benefits because it represents the environment in which humans have spent the clear majority of their evolutionary history. SRT has been conceptually supported through a diverse set of empirical findings, including nature's ability to improve post-operative recovery (Ulrich, 1984), lower self-reported and physiological measures of stress (Ulrich et al., 1991), and improve positive affect and lower negative affect (McMahan & Estes, 2015).

Thus, the relationship between aesthetic preference and restoration may depend on whether the restoration is operating at a cognitive or affective level. Specifically, the aesthetic preference for nature may be more integral to nature's restorative potential under SRT (reducing stress) compared to ART (improving directed attention) based on the main tenets of these theories and associated empirical findings. Critically, the present set of experiments laid the groundwork for empirically testing this possibility of dissociating aesthetic responses from cognitive and affective restoration. Using stimuli that have acoustic signatures of nature but are not aesthetically preferred in a restoration paradigm (e.g., measuring improvement on cognitive or affective measures after engendering cognitive fatigue or a stressor), we might expect cognitive restoration but not affective restoration, as the former may be more independent of aesthetic preference than the latter. Overall, then, the results from the present set of experiments provide a theoretical wedge to potentially dissociate aesthetic preference from specific kinds of restorative experiences that nature may provide, such as cognitive restoration.

5.2. *Implications for perceptual and aesthetic judgments beyond nature*

The dynamic interplay between perceptual and conceptual information in forming aesthetic preferences, observed in the present experiments, is paralleled in other aesthetic domains such as musical and visual art. While there are particular, sensory-based foundations for aesthetic reactions to musical events (e.g., Terhardt, 1974) as well as conceptual foundations, the interplay between perceptual and conceptual factors can lead to "paradoxical" situations, in which, for example, the aesthetic appraisal of music does not match the sensory pleasurableness (Brattico, 2015). Moreover, sounds that are consistently judged as unpleasant, presumably due to acoustic features (e.g., fingernails on a chalkboard), are judged as less unpleasant when given an alternative conceptual framing, such as being a part of contemporary music (Reuter & Oehler, 2011). These findings similarly suggest that the framing of a particular sound can substantially alter the way in which it is aesthetically valued.

In visual art, understanding the context in which a particular work was produced can substantially change how it is valued; philosophical treatments of aesthetics in visual art suggest that "appropriate" aesthetic reactions must be informed by factors such as the historical and social context in which a work is produced, as well as an appropriate categorization of a work into a school or style (Walton, 1970). One particularly illuminating example of this point is the philosophical treatment of forgeries—for example, two paintings that are perceptually indiscriminable, yet one is a forgery of the other (e.g., Hoaglund, 1976; Kulka, 1982). While Kulka (1982) considers both paintings in such a scenario to be aesthetically identical—essentially treating aesthetics as an invariant judgment to perceptual features—our results suggest that aesthetic judgments cannot be dissociated from contextual frames that are used for interpretation. This point of view is nicely illustrated by Berger, Blomberg, Fox, Dibb, and Hollis (1972), in which readers are presented with a painting of a cornfield with birds flying out of it and asked to take a moment to observe. Then, when readers turn the page, the painting is reprinted with a new framework—the last painting Vincent Van Gogh painted before he killed himself. Berger writes, “It is hard to define exactly how the words have changed the image but undoubtedly they have. The image now illustrates the sentence” (p. 28). This example demonstrates that context changes aesthetic reactions at multiple levels—that is, the new framework does not necessarily change the recognition of the objects in the painting, yet it presumably changes the interpretation of the meaning of those objects and consequently the aesthetic response.

Beyond music and visual art, the present results fit within existing frameworks from the psychology of aesthetics and preference formation more generally. Bergeron and Lopes (2012) argue that aesthetic experiences require the evaluation of an object, possess an affective dimension (i.e., it is subjectively felt), and *contain a semantic dimension* (i.e., are not captured by mere sensation). This final point is particularly important for understanding the present results, as it suggests that aesthetic reactions, as a construct, may not exist independently of a semantic framework for interpretation. This influence of frames on preference extends beyond the psychology of aesthetics to general aspects of decision making. In economics, “reference levels” have been shown to significantly influence decision making (Tversky & Kahneman, 1991), in that changes of reference level can fundamentally change (and even reverse) preference. For example, the valuing of a choice is defined by the framework of gains and losses relative to a specific reference level, with losses demonstrating a steeper function than gains (i.e., loss aversion). In this sense, the present results suggest that the processes that underlie aesthetic preferences of nature stimuli may reflect a fundamental aspect of how humans assign aesthetic value and preference across domains.

5.3. *Conclusion*

Aesthetic reactions to nature have been discussed in science, philosophy, and urban planning for well over a century (e.g., Muir, 1894; Olmsted, 1865/1952). Understanding how nature can elicit aesthetic reactions from individuals makes up an important part of the contemporary study of empirical aesthetics (e.g., Leder & Nadal, 2014), and a better understanding of the processes that contribute to aesthetic preferences for nature have both theoretical and practical applications. The present results situate aesthetic responses to nature in a broader context of preference formation and decision-making by highlighting that listeners’ aesthetic responses to nature sounds cannot be effectively modeled by the acoustic properties of nature sounds alone. Rather, it is essential to incorporate the relevant framework (i.e., that the sounds originate from nature), as this framework dynamically reorganizes the relationship between aesthetic ratings and acoustic properties, as well as drives the aesthetic preference for nature sounds. These findings have important implications for our understanding of nature as a restorative environment, as well as more broadly inform our understanding of aesthetic preference formation. Consistent with recent theoretical accounts of empirical aesthetics, the present results strongly support the idea that aesthetic judgments cannot be understood via mere sensations, but rather require a semantic understanding. The dynamic interplay between acoustic features and aesthetic judgments in this paper help to provide testable hypothesis that will further our understanding of the mechanisms through which nature may provide affective and cognitive benefits.

Acknowledgments

This work was supported by a grant from the TKF Foundation to Marc G. Berman, two grants from the John Templeton Foundation (the University of Chicago Center for Practical Wisdom) to Howard Nusbaum (PI) and Marc G. Berman (co-PI) and the Virtue, Happiness, and Meaning of Life Scholars Group, and a grant from the National Science Foundation (BCS-1632445) to Marc G. Berman, as well as an internal grant from the University of Chicago to Marc G. Berman.

Authors' contributions

SCVH and MGB designed the experiments. AH coded the experiments as well as assisted in data processing. SCVH analyzed the data and prepared a first draft, after which all authors provided comments until the manuscript reached its final form.

Data availability

The data from all experiments, in addition to the experimental materials, are available on the Open Science Framework (https://osf.io/4wb27/).

References

Alvarsson, J. J., Wiens, S., & Nilsson, M. E. (2010). Stress recovery during exposure to nature sound and environmental noise. *International Journal of Environmental Research and Public Health*, *7*(3), 1036–1046. https://doi.org/10.3390/ijerph7031036.

Balling, J. D., & Falk, J. H. (1982). Development of visual preference for natural environments. *Environment and Behavior*, *14*(1), 5–28. https://doi.org/10.1177/0013916582141001.

Basner, M., Babisch, W., Davis, A., Brink, M., Clark, C., Janssen, S., & Stansfeld, S. (2014). Auditory and non-auditory effects of noise on health. *The Lancet*, *383*(9925), 1325–1332. https://doi.org/10.1016/S0140-6736(13)61613-X.

Bauer, B. a., Cutshall, S. a., Anderson, P. G., Prinsen, S. K., Wentworth, L. J., Olney, T. J., Messner, P. K., Brekke, K. M., Li, Z., & Sundt, T. M., & Kelly, R. F. (2011). Effect of the combination of music and nature sounds on pain and anxiety in cardiac surgical patients: A randomized study. *Alternative Therapies in Health and Medicine*, *17*(4), 16–23.

Berger, J., Blomberg, S., Fox, C., Dibb, M., & Hollis, R. (1972). *Ways of seeing*. London: British Broadcasting Corporation and Penguin Books.

Bergeron, V., & Lopes, D. M. (2012). Aesthetic theory and aesthetic science: Prospects for integration. In A. Shimamura & S. E. Palmer (Eds.), *Aesthetic science: Connecting minds, brains, and experience* (pp. 63–79). New York: Oxford University Press.

Berman, M. G., Hout, M. C., Kardan, O., Hunter, M. R., Yourganov, G., Henderson, J. M., Hanayik, T., Karimi, H., & Jonides, J. (2014). The perception of naturalness correlates with low-level visual Features of environmental scenes. *PLoS ONE*, *9*(12), e114572. https://doi.org/10.1371/journal.pone.0114572.

Berman, M. G., Jonides, J., & Kaplan, S. (2008). The cognitive benefits of interacting with nature. *Psychological Science*, *19*(12), 1207–1212. https://doi.org/10.1111/j.1467-9280.2008.02225.x.

Berman, M. G., Kross, E., Krpan, K. M., Askren, M. K., Burson, A., Deldin, P. J., Kaplan, S., Sherdell, L., Gotlib, I. H., & Jonides, J. (2012). Interacting with nature improves cognition and affect for individuals with depression. *Journal of Affective Disorders*, *140*(3), 300–305. https://doi.org/10.1016/j.jad.2012.03.012.Interacting.

Berto, R. (2005). Exposure to restorative environments helps restore attentional capacity. *Journal of Environmental Psychology*, *25*(3), 249–259. https://doi.org/10.1016/j.jenvp.2005.07.001.

Berto, R. (2014). The role of nature in coping with psycho-physiological stress: A literature review on restorativeness. *Behavioral Sciences*, *4*(4), 394–409. https://doi.org/10.3390/bs4040394.

Brattico, E. (2015). From pleasure to liking and back: Bottom-up and top-down neural routes to the aesthetic enjoyment of music. In J. P. Huston, M. Nadal, F. Mora, L. F. Agnati, & C. J. Cela-Conde (Eds.), *Art, aesthetics and the brain* (pp. 303–318). New York: Oxford University Press.

Carlson, A. (1981). Nature, aesthetic judgment, and objectivity. *The Journal of Aesthetics and Art Criticism*, *40*(1), 15–27. https://doi.org/10.2307/430349.

Cohen, S., & Weinstein, N. (1981). Nonauditory effects of noise on behavior and health. *Journal of Social Issues*, *37*(1), 36–70. https://doi.org/10.1111/j.1540-4560.1981.tb01057.x.

Dadvand, P., Nieuwenhuijsen, M. J., Esnaola, M., Forns, J., Basagaña, X., Alvarez-Pedrerol, M., Rivas, I., López-Vicente, M., De Castro Pascual, M., Su, J., Jerrett, M., Querol, X., & Sunyer, J. (2015). Green spaces and cognitive development in primary schoolchildren. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(26), 7937–7942. https://doi.org/10.1073/pnas.1503402112.

De Leeuw, J. R. (2015). jsPsych : A JavaScript library for creating behavioral experiments in a Web browser. *Behavioral Research*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y.

DeLoach, A. G., Carter, J. P., & Braasch, J. (2015). Tuning the cognitive environment: Sound masking with "natural" sounds in open-plan offices. *The Journal of the Acoustical Society of America*, *137*, 2291. https://doi.org/doi:10.1121/1.4920363.

Gould van Praag, C. D., Garfinkel, S. N., Sparasci, O., Mees, A., Philippides, A. O., Ware, M., Ottaviani, C., & Critchley, H. D. (2017). Mind-wandering and alterations to default mode network connectivity when listening to naturalistic versus artificial sounds. *Scientific Reports*, *7*(February), 45273. https://doi.org/10.1038/srep45273.

Gullone, E. (2002). The biophilia hypothesis and life in the 21st century: Increasing mental health or increasing pathology? *Journal of Happiness Studies*, *1*(3), 293–322. https://doi.org/10.1023/a:1010043827986.

Haga, A., Halin, N., Holmgren, M., & Sörqvist, P. (2016). Psychological restoration can depend on stimulus-source attribution: A challenge for the evolutionary account? *Frontiers in Psychology*, *7*(Nov), 1–8. https://doi.org/10.3389/fpsyg.2016.01831.

Han, K. (2010). An exploration of relationships among the responses to natural scenes. *Environment and Behavior*, *42*(2), 243–270. https://doi.org/10.1177/0013916509333875.

Hartig, T., & Staats, H. (2006). The need for psychological restoration as a determinant of environmental preferences. *Journal of Environmental Psychology*, *26*(3), 215–226. https://doi.org/10.1016/j.jenvp.2006.07.007.

Hoaglund, J. (1976). Originality and aesthetic value. *British Journal of Aesthetics*, *16*(1), 46–55. https://doi.org/10.1093/bjaesthetics/16.1.46.

Kaplan, S., Kaplan, R., & Wendt, J. S. (1972). Rated preference and complexity for natural and urban visual material. *Perception & Psychophysics*, *12*(4), 354–356. https://doi.org/10.3758/BF03207221.

Kaplan, R., & Kaplan, S. (1989). *The experience of nature: A psychological perspective*. New York, NY, US: Cambridge University Press.

Kaplan, S. (1995). The restorative benefits of nature: Toward an integrative framework. *Journal of Environmental Psychology*, *15*(3), 169–182. https://doi.org/10.1016/0272-4944(95)90001-2.

Kardan, O., Demiralp, E., Hout, M. C., Hunter, M. R., Karimi, H., Hanayik, T., Yourganov, G., Jonides, J., & Berman, M. G. (2015). Is the preference of natural versus man-made scenes driven by bottom-up processing of the visual features of nature? *Frontiers in Psychology*, *6*(Apr), 1–13. https://doi.org/10.3389/fpsyg.2015.00471.

Kardan, O., Gozdyra, P., Misic, B., Moola, F., Palmer, L. J., Paus, T., & Berman, M. G. (2015). Neighborhood greenspace and health in a large urban center. *Scientific Reports*, *5*, 11610. https://doi.org/10.1038/srep11610.

Kline, G. A. (2009). Does a view of nature promote relief from acute pain? *Journal of Holistic Nursing*, *27*(3), 159–166. https://doi.org/10.1177/0898010109336138.

Kotabe, H. P., Kardan, O., & Berman, M. G. (2017). The nature-disorder paradox: A perceptual study on how nature is disorderly yet aesthetically preferred. *Journal of Experimental Psychology: General*, *146*(8), 1126–1142. https://doi.org/10.1037/xge0000321.

Krumhansl, C. L. (2010). Plink: "Thin slices" of music. *Music Perception: An Interdisciplinary Journal*, *27*(5), 337–354. https://doi.org/10.1525/MP.2010.27.5.337.

Kulka, T. (1982). The artistic and aesthetic status of forgeries. *Leonardo*, *15*(2), 115–117.

Leder, H., & Nadal, M. (2014). Ten years of a model of aesthetic appreciation and aesthetic judgments: The aesthetic episode—Developments and challenges in empirical aesthetics. *British Journal of Psychology*, *105*(4), 443–446. https://doi.org/10.1111/bjop.12084.

Mace, B. L., Bell, P. A., & Loomis, R. J. (2004). Visibility and natural quiet in national parks and wilderness areas: Psychological considerations. *Environment and Behavior*, *36*(1), 5–31. https://doi.org/10.1177/0013916503254747.

Marselle, M. R., Irvine, K. N., & Warber, S. L. (2014). Examining group walks in nature and multiple aspects of well-being: A large-scale study. *Ecopsychology*, *6*(3), 134–147. https://doi.org/10.1089/eco.2014.0027.

Martín-López, B., Montes, C., & Benayas, J. (2007). The non-economic motives behind the willingness to pay for biodiversity conservation. *Biological Conservation*, *139*(1–2), 67–82. https://doi.org/10.1016/j.biocon.2007.06.005.

McDermott, J. H., & Simoncelli, E. P. (2011). Sound texture perception via statistics of the auditory periphery: Evidence from sound synthesis. *Neuron*, *71*(5), 926–940. https://doi.org/10.1016/j.neuron.2011.06.032.

McMahan, E. A., & Estes, D. (2015). The effect of contact with natural environments on positive and negative affect: A meta-analysis. *The Journal of Positive Psychology*, *9760*(November), 1–13. https://doi.org/10.1080/17439760.2014.994224.

Meidenbauer, K., Stenfors, C. U. D., Young, J., Layden, E., Schertz, K., Kardan, O., Decety, J., & Berman, M. G. (2018). The gradual development of the preference for natural environments. *PsyArXiv*. https://doi.org/10.31234/osf.io/7hw83

Moore, R. C. (1986). Childhood outdoors: Toward a social ecology of the landscape. *Children's Environments Quarterly*, *3*(3), 52–69.

Muir, J. (1894). The mountains of California. *John Muir: A reading bibliography by Kimes*. Available at https://scholarlycommons.pacific.edu/jmb/213

Olmsted, F. L. (1865/1952). The Yosemite Valley and the Mariposa big trees: A preliminary report: With an introductory note by Laura Wood Roper. *Landscape Architecture*, *43*, 12–25.

Purcell, T., Peron, E., & Berto, R. (2001). Why do preferences differ between scene types? *Environment and Behavior*, *33*(1), 93–106. https://doi.org/10.1177/00139160121972882.

Raanaas, R. K., Evensen, K. H., Rich, D., Sjostrom, G., & Patil, G. (2011). Benefits of indoor plants on attention capacity in an office setting. *Journal of Environmental Psychology*, *31*(1), 99–105. https://doi.org/10.1016/j.jenvp.2010.11.005.

Ratcliffe, E., Gatersleben, B., & Sowden, P. T. (2013). Bird sounds and their contributions to perceived attention restoration and stress recovery. *Journal of Environmental Psychology*, *36*, 221–228. https://doi.org/10.1016/j.jenvp.2013.08.004.

Reuter, C., & Oehler, M. (2011). Psychoacoustics of chalkboard squeaking. *The Journal of the Acoustical Society of America*, *130*, 2545. https://doi.org/10.1121/1.3655174.

Simaika, J. P., & Samways, M. J. (2010). Biophilia as a universal ethic for conserving biodiversity. *Conservation Biology*, *24*(3), 903–906.

Stenfors, C. U. D., Van Hedger, S. C., Schertz, K., Meyer, F., Smith, K., Norman, G. J., Bourrier, S., Enns, J., Jonides, J., & Berman, M. G. (2018). Cognitive restoration from nature across multiple exposure types —A meta-analysis delineating the effects of test order and affect. *PsyArXiv*. https://doi.org/10.31234/osf.io/w6n2f

Sueur, J., Aubin, T., & Simonis, C. (2008). Sound analysis and synthesis with the package Seewave. *Bioacoustics*, *18*, 213–226.

Terhardt, E. (1974). Pitch, consonance, and harmony. *Journal of the Acoustical Society of America*, *55*(5), 1061–1069. https://doi.org/10.1121/1.1914648.

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, *106*(4), 1039–1061. https://doi.org/10.2307/2937956.

Ulrich, R. S. (1983). Aesthetic and affective response to natural environment. *Human Behavior & Environment: Advances in Theory & Research*, *6*, 85–125.

Ulrich, R. S. (1984). View through a window may influence recovery from surgery. *Science*, *224*(4647), 420–421. https://doi.org/10.1126/science.6143402.

Ulrich, R. S. (1993). Biophilia, biophobia, and natural landscapes. In E. O. Wilson & S. R. Kellert (Eds.), *The biophilia hypothesis* (pp. 73–137). Washington, D.C.: Island Press.

Ulrich, R. S., Simons, R. F., Losito, B. D., Fiorito, E., Miles, M. A., & Zelson, M. (1991). Stress recovery during exposure to natural and urban environments. *Journal of Environmental Psychology*, *11*, 201–230. https://doi.org/0272-4944/91/030201 + 30503.00/0.

van den Berg, A. E., Koole, S. L., & van der Wulp, N. Y. (2003). Environmental preference and restoration: (How) are they related? *Journal of Environmental Psychology*, *23*(2), 135–146. https://doi.org/10.1016/S0272-4944(02)00111-1.

Van Hedger, S. C., Nusbaum, H. C., Clohisy, L., Jaeggi, S. M., Buschkuehl, M., & Berman, M. G. (2018). Of cricket chirps and car horns: The effect of nature sounds on cognitive performance. *Psychonomic Bulletin and Review*. https://doi.org/10.3758/s13423-018-1539-1.

Walton, K. L. (1970). Categories of art. *Philosophical Review*, *79*(3), 334–367. https://doi.org/10.2307/2183933.

Wilson, E. O. (1984). *Biophilia*. Cambridge, MA: Harvard University Press.

Wilson, E. O., & Kellert, S. R. (1993). *The biophilia hypothesis*. Washington, D.C.: Island Press.

Yang, B., & Kaplan, R. (1990). The perception of landscape style: A cross-cultural comparison. *Landscape and Urban Planning*, *19*(3), 251–262. https://doi.org/10.1016/0169-2046(90)90024-V.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, *35*(2), 151–175.

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Figure S1. Identification accuracy for scrambled sounds, broken into specific basic-level categories of nature and urban sounds. NB, nature birdsong, NN, nature noise, NO, nature other, NW, nature water, UN, urban noise, UT, urban traffic, UO, urban other, UV, urban voice.

Figure S2. Naturalness ratings of the 40 nature and 40 urban soundscapes from a separate MTurk sample (n = 50). Ratings were made on a 7-point scale with "1" corresponding to "very urban" and "7" corresponding to "very natural." Nature sounds ($M \pm SD$: 5.95 ± 0.62) were rated significantly higher on the naturalness scale compared to urban sounds ($M \pm SD$: 2.36 ± 0.74; $t(78) = 23.51$, $p < .001$, $d = 5.26$). The distribution of naturalness ratings for nature and urban sounds was completely non-overlapping (highlighted by the dashed horizontal line); the lowest rated nature sound (4.30) was rated higher than the highest-rated urban sound (4.24). These results provide strong support that the selected sounds indeed represented nature and urban categories.