Journal of Environmental Psychology 72 (2020) 101498

Contents lists available at ScienceDirect

Journal of Environmental Psychology

journal homepage: www.elsevier.com/locate/jep

The affective benefits of nature exposure: What's nature got to do with it?

Kimberly L. Meidenbauer [a,*], Cecilia U.D. Stenfors [a,b,c], Gregory N. Bratman [d], James J. Gross [e], Kathryn E. Schertz [a], Kyoung Whan Choe [a], Marc G. Berman [a,**]

[a] The Environmental Neuroscience Lab, Department of Psychology, University of Chicago, Chicago, IL, USA
[b] Department of Psychology, Stockholm University, Stockholm, Sweden
[c] Aging Research Center, Department of Neurobiology, Care Science & Society, Karolinska Institute, Solna, Sweden
[d] School of Environmental and Forest Sciences, University of Washington, Seattle, WA, USA
[e] Department of Psychology, Stanford University, Stanford, CA, USA

ARTICLEINFO

Keywords:
Nature
Emotion
Aesthetics
Environmental preferences
Mood induction

ABSTRACT

Nature interactions have been demonstrated to produce reliable affective benefits. While adults demonstrate strong preferences for natural environments over urban ones, it is not clear whether these affective benefits result from exposure to nature stimuli per se, or result from viewing a highly preferred stimulus. In one set of studies (Study 1 and 2), state affect before and after image viewing was examined as a function of both preference level (high, low, very high, or very low aesthetic value) and environment type (nature or urban). When aesthetic value was matched, no differences in affect change were found between environments. However, affect change was predicted by individual participants' ratings for the images. The largest affective benefits occurred after viewing very high aesthetic nature images, but Study 2 lacked an equivalently preferred urban image set. In a second set of studies (Study 3 and 4), new sets of very highly preferred images in categories other than nature scenes (urban scenes and animals) were employed. As before, individual differences in preference for the images (but not image category) was predictive of changes in affect. In Study 5, the nature and urban images from Study 1 were rated on beauty to assess whether the stimuli's preference ratings were capturing anything other than simple aesthetics. Results showed that beauty/aesthetics and preference ('liking') were nearly identical. Lastly, a replication of Study 2 (Study 6) was conducted to test whether priming preference accounted for these benefits, but this was not the case. Together, these results suggest that nature improves affective state because it is such a highly preferred environment.

In the field of environmental psychology, a substantial body of research documents the relationships between the physical environment and human psychological functioning. One particular focus has been the link between nature exposure and emotional well-being. This link has been demonstrated using a variety of study designs, exposure types, and outcome measures (Bowler et al., 2010; McMahan & Estes, 2015). Frequently, such studies have employed controlled experimental designs in which the effects of nature interventions are compared with that of control interventions (often urban environment exposures).

Short-term effects of nature exposure have been examined using nature interventions that have varied in type and duration, but the observed benefits are consistent. Brief walks in natural settings have been shown to increase positive affect and decrease negative affect compared to urban walks, and the effects have been found in both healthy and clinical populations (Bratman et al., 2015; Fuegen & Breitenbecher, 2018; Hartig et al., 2003; Johansson et al., 2011; Mayer et al., 2009). Passively viewing a natural environment (e.g. sitting in a forested area) improves self-reported and physiological measures of affect compared to spending the same time viewing a built environment (e.g. sitting in a parking lot) (Tsunetsugu et al., 2013; Lee et al., 2009). Simulated nature in the form of videos, image slideshows, and virtual reality (VR) elicit improvements in emotional state as well (Beute & de Kort, 2014; Hartig et al., 1996; Valtchanov et al., 2010; van den Berg et al., 2003), though the effects are somewhat smaller than those of actual nature exposure (McMahan & Estes, 2015).

Benefits of nature exposure have also been examined longitudinally in epidemiological (e.g. White et al., 2013) and experience sampling studies. The latter combine regular assessments of emotional state (and

* Corresponding author. 5848 S University Avenue, Chicago, IL, 60637, USA.
** Corresponding author.
E-mail address: meidenbauer@uchicago.edu (K.L. Meidenbauer).

https://doi.org/10.1016/j.jenvp.2020.101498
Received 16 December 2019; Received in revised form 11 September 2020; Accepted 11 September 2020
Available online 18 September 2020
0272-4944/

other measures) with GPS location data from a mobile device. These studies find that being in an outdoor natural environment is related to more positive affect than being in other types of environments (e.g., indoors at home, outdoor urban environments, in transit, etc.) even when accounting for other relevant variables such as weather, daylight, or physical activity (Ryan et al., 2010; Beute & de Kort, 2018; Glasgow et al., 2019; MacKerron & Mourato, 2013).

While studies documenting benefits of nature exposure on emotional well-being are abundant, it is unclear whether these effects are the result of exposure to nature per se, or whether these effects are the result of viewing preferred stimuli. Research on adult environmental preferences has demonstrated that overwhelmingly, nature scenes are preferred over their urban counterparts (Kaplan & Herbert, 1987; Kaplan & Kaplan, 1989; Kaplan et al., 1972; Ulrich, 1979; 1983). Despite work identifying preference as an influence in the emotional benefits of nature exposure (Mayer et al., 2009; Beute & de Kort, 2014; White et al., 2010; Browning et al., 2020), many studies linking nature contact to psychological well-being do not assess preference for these environments, or examine how affect change relates to preference (Lee et al., 2009; Tsunetsugu et al., 2013; Valtchanov et al., 2010).

Those studies which have directly linked preference for natural environments to affective benefits have approached the question in different ways. Beute and de Kort (2014) showed participants an image slideshow and measured state affect and preference for the slideshows. The nature slideshow was preferred over the urban one, and a subsequent mediation analysis showed that positive affect change (operationalized as "hedonic tone") was mediated by preference for the slideshow. A different approach was taken by White et al. (2010) who did not employ the standard pre/post study design. Instead, they had participants rate a series of scenes on aesthetics (i.e "how attractive is the scene?"), behavioral preference (i.e. "how willing would you be to visit this scene?"), and their affective response to each image (i.e. "how does this photo make you feel?"), and then examined how these attributes related to each other and to the image types (i.e. fully built environments, green nature, nature with water, built environments with water, etc.). They found that images which were rated most highly on preference were also highly rated on positive affective responses. Additionally, they found that by incorporating bodies of water in images of built environments, these environments were also rated more favorably on both preference and affect measures, compared to urban images without water. Another approach measured affective restoration using videos of a built environments with water elements (dockland) and contrasting it with a video of a nature reserve (Karmanov & Hamel, 2008). Here, the nature video was rated as more attractive than the urban one, and the nature video caused relatively larger affective restoration, again suggesting that there is a potentially important role of aesthetic preferences. Interestingly, a recent study also found that while the beauty of a VR nature scene was related to change in affect, ratings of scene disgust were not (Browning et al., 2020).

These findings suggest an important question, namely whether there is something unique or special about the "naturalness" of visual exposure to nature itself (other than that it is preferred over other environments) which leads to affect change. Though preference has, at times, been conceptualized as a type of positive affective response (Zajonc, 1980), there is still reason to treat these constructs as separable, and question whether an aesthetic preference for nature is fully responsible for a change in affect. A general change in affective state (typically varying along valence and arousal dimensions) is believed to be one precursor to the experience of an emotion (Barrett, 2006; Lindquist et al., 2016). While the experience of an emotion may be construed based on contextual, social, and dispositional factors, a change in affective state is a more general, physiologically-based response which can be triggered by an external stimulus and/or result from interoceptive sensations (Barrett et al., 2007). In contrast, the notion of preference or affinity is necessarily directed at something (i.e., there must be an object or thing that is preferred, there is no self-contained state of "preference"). Therefore, being exposed to a preferred thing can cause a change in affective state, but a change in affective state does not need to involve a preferred or non-preferred stimulus. As such, preference and affect are not one in the same.

In general, aesthetic judgments typically fall along the spectra of ugly-to-beautiful and disliked-to-liked. Preference ('liking') and beauty are also highly related, though they need not always overlap. That is, while things that are beautiful are often preferred, there are plenty of things many people have an affinity for that are not canonically beautiful (i.e., pet cockroaches, a child's artwork). In the case of nature preferences, the aesthetics have been defined by both measures of affinity ('liking') and beauty (Ulrich, 1983; van den Berg et al., 2003).

Another reason this question necessitates empirical investigation is due to theoretical disagreements over whether the affective and cognitive benefits of nature exposure are linked, and what role aesthetic preferences play for each. Attention Restoration Theory proposes a mechanism of cognitive restoration that is independent of affect change and unrelated to the aesthetic preference for such environments(Kaplan, 1995; Kaplan & Berman, 2010) In contrast, Stress Reduction Theory proposes that the cognitive effects arise from improvements in affect and a reduction in stress, which then frees up cognitive resources (Ulrich, 1983; Ulrich et al., 1991). In this theory, the aesthetic preference is what drives the affective benefits and also the cognitive benefits. The Perceptual Fluency Account also proposes a causal link between improved affect and resulting cognitive restoration, but proposes that the affective benefits arise from fluently processing features (e.g., fractalness) that are more prevalent in natural environments, though not necessarily beauty (Joye et al., 2016; Van den Berg et al., 2016). Other researchers have suggested that nature may be endogenously visually rewarding due to the distribution of visual spatial frequency information in nature scenes (Valtchanov & Ellard, 2015). Prospect-Refuge theory suggests that the extent to which a scene corresponds to an evolutionarily beneficial natural environment (i.e., one offering safety and a good prospect for acquiring resources) is what will engender a preference for these environments, leading to the cognitive and affective benefits. Lastly, Kuo (2015) suggested a variety of "active ingredients" which can explain nature's health benefits, such as negative air ions, phytoncides, and biodiversity, factors which may also explain changes in emotional processing and reduced stress responses.

Thus, some of these theories suggest that the cognitive benefits arise specifically from the affective ones, which result from either 1) an evolutionarily based affinity for these environments (Stress Reduction Theory, Prospect-Refuge) or 2) the low-level visual features of natural environments (Perceptual Fluency Account, endogenous visual reward). Attention Restoration Theory suggests that neither affect nor preference play a role. Ultimately, since every framework differs to a certain extent on whether or how preference plays a role, but all theories reference the effects of nature (not simply the effects of a "preferred stimulus"), it is important to test whether there is something about the naturalness of nature scenery that is important for mood. As the semantic category of nature has been shown to predict preference above and beyond the low-level features (Kotabe et al., 2017; van Hedger et al., 2019), it is entirely possible that the category of nature may provide some additional affective benefit beyond preference. Based on this, we hypothesized in our pre-registration that nature images may generate a larger benefit to state affect than preference-equated urban images. Ultimately, though nature is highly preferred, it remains unknown whether unthreatening but aesthetically ugly nature would have any of the mood boosting effects that typical natural scenery would, or if particularly beautiful nature improves affect beyond particularly beautiful cities. The current set of studies were designed to answer this outstanding question.

The aim of this research was to test the effects of environment type (nature vs. urban) and aesthetic value (preference level) on changes in affect from pre to post viewing of image sets. It should be noted here that in each of the studies, nature is operationalized as scenes or environments with primarily natural (non-built) elements. To keep the focus on environmental scenes we do not have stimuli that include animals (both our more natural and our more urban scenes were devoid of human and non-human animals). We excluded animals in our scene stimuli because the presence of animals in more urban scenes may have complicated their interpretation as natural or not. In addition, we did not have stimuli that represented singular organic or inorganic material (e.g., a single flower or a piece of wood, both of which are natural, but do not constitute an environment; or a single inorganic object, like a cell-phone). Additionally, the natural environments included are highly varied in landscape type, and include examples of complete wilderness, deserts, open fields, beaches, mountains, large bodies of water, forest trails, sand dunes, etc. Study 1 was conducted to generate the preference-equated sets of natural and urban scenes. We reasoned that if nature has a positive effect on affective state that is not simply due to its status as a preferred environment, then exposure to natural environments should elicit larger positive affective changes than preference-equated urban environments. Conversely, if environment-type is less important than aesthetic preferences, then differences in affect should be observed primarily between image sets that vary on aesthetic value, but not on environment-type. Alternatively, both preference and environment type could have interactive effects on affect. All of these possibilities were tested in Study 2.

1. Study 1: Original image validation

1.1. Study Intro

Study 1 was conducted to find preference-equated sets of natural and urban scene images to be used in Study 2, which then examined how exposure to images of different environments across multiple levels of aesthetic value (preference) altered state affect.

1.2. Method

1.2.1. Participants

401 US-based adults (195 male, 201 female, 4 other, 1 no response) were recruited from Amazon Mechanical Turk. The age of participants ranged from 19 to 73 years (M 36.0, SD=11.0). The full study procedures were expected to take approximately 15 minutes and participants were compensated $1.50 for participating. Informed consent was administered by the University of Chicago Institutional Review Board.

1.2.2. Stimuli & procedure

Participants were randomly assigned to rate 100 images out of a total of 375 potential images on a 1–7 Likert scale evaluating either their preference for the images (anchors of 1 = “strongly dislike” and 7 = “strongly like”; n = 200) or the naturalness of the image (anchors of 1 = “very man-made” and 7 = “very natural”; n = 201). The 375 images were taken from the SUN database (Xiao et al., 2010) as well as an online image search for non-copyrighted scene images. Criteria for image inclusion were that the images: 1) Do not contain people or animals, 2) Have sufficiently high resolution for clear viewing, 3) Have minimal trees/natural elements in the urban images and minimal man-made elements in the nature images, and 4) Have minimal text, signs, or graffiti. Additionally, images were selected containing a variety of viewpoints (horizons, slanted towards the ground, slanting upwards, etc.) and varied urban or natural forms (different types of nature, varied buildings or cities). The 100 images that participants saw were selected pseudorandomly, to show a relatively equal number of natural and urban images as well as images that varied on aesthetic value. Based on this pseudorandomization each image received a minimum of 33 preference ratings, with an average of roughly 51 ratings per image, and a minimum of 37 naturalness ratings, with an average of 53 ratings per image.

1.3. Results

From the 375 images examined, six sets of 45 images each were able to be created. Aesthetic preference ratings for the 45 images in each condition are presented in Table 1. Two pairs of preference-matched nature and urban image sets were created - High Aesthetic Value Nature (HA-Nat) and Urban (HA-Urb), and Low Aesthetic Value Nature (LA-Nat) and Urban (LA-Urb) conditions. The High and Low Aesthetic Value image sets had an average preference rating of 4.6 and 3.8, respectively. Unfortunately, there were very few strongly liked urban images or strongly disliked nature images. As such, there were not enough urban images to match the Very High Aesthetic Value Nature (VHA-Nat) images, nor were there enough disliked nature images to match those in the Very Low Aesthetic Value Urban (VLA-Urb) condition. Though we were not able to equate these stimuli, we did not want to throw them out as we could still examine differences in preference within environment type.

Across all Aesthetic Value levels, images in the Nature conditions and Urban conditions were rated very differently on naturalness. Average naturalness ratings for VHA, HA, and LA-Nat were 6.72, 6.66, and 6.64, respectively. Average naturalness ratings for HA, LA, and VLA-Urb were 1.42, 1.44, and 1.39, respectively. A subset of the images validated in this study and used in studies 2, 5, and 6 were also used in (Meidenbauer et al., 2019).

2. Study 2: Examining affect change as a function of environment and aesthetic value

2.1. Study Intro

The primary question for Study 2 was whether changes in affect are due to naturalness, whether they are purely due to preference, or result from a combination of the two. All three of these possibilities were examined, using the image sets obtained from Study 1. A visual depiction of the hypotheses for Study 2, as well as the analyses chosen to address each hypothesis, are presented in Fig. 1.

2.2. Method

Study 2 was preregistered on OSF prior to data collection: https://osf.io/tuezg. The data for Study 2 are publicly available at: https://osf.io/ehtk9.

2.2.1. Participants

615 US-based adults (287 male, 324 female, 4 other) were recruited from Amazon Mechanical Turk. Age of the participants ranged from 20 to 76 years (M=37.0, SD =10.9). The full study procedures (including additional tasks after collecting the data for this study) were expected to take approximately 30 min and participants were compensated $3.00 for participating. Informed consent was administered by the University

Table 1
Study 1 aesthetic value ratings of images in each condition.
Ratings of all images in each condition from a previous validation study.
Aesthetic value ratings are on a 1–7 scale (1 = strongly dislike, 7 = strongly like).

	Nature Conditions		Urban Conditions	
Aesthetic Value Condition	*M (SD)*	*Range* [Min, Max]	*M (SD)*	*Range* [Min, Max]
Very High Aesthetic Value	6.34 (0.13)	[6.19, 6.62]		
High Aesthetic Value	4.59 (0.17)	[4.28, 4.86]	4.58 (0.29)	[4.15, 5.29]
Low Aesthetic Value	3.78 (0.35)	[2.90, 4.21]	3.78 (0.23)	[3.30, 4.12]
Very Low Aesthetic Value			2.64 (0.31)	[1.77, 3.08]

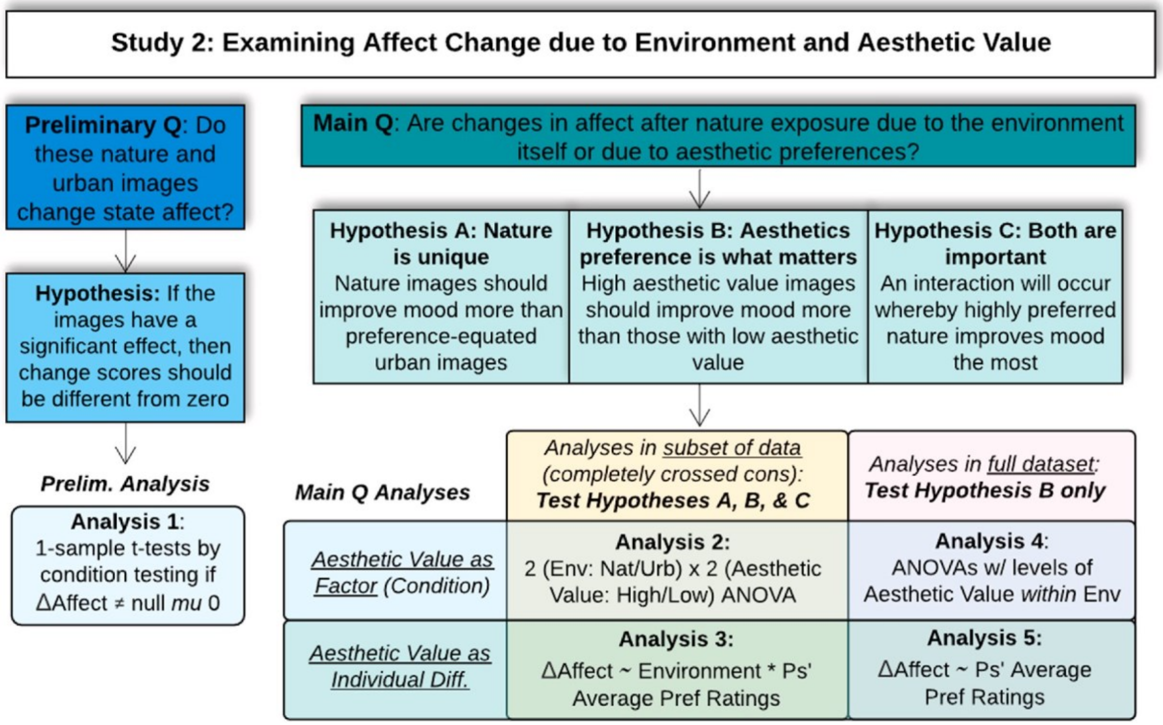

Fig. 1. Hypotheses & analyses overview for study 2.

of Chicago Institutional Review Board. Sample size was decided prior to data collection and specified in the pre-registration. Each of the image conditions presented below was also broken into two groups in subsequent tasks (unreported), so the sample was based on power for these subsequent tasks.

2.2.2. Experimental conditions & stimuli

Participants were randomly assigned to one of 6 image conditions based on Study 1: Very High Aesthetic Value Nature (VHA-Nat; n = 103), High Aesthetic Value Nature (HA-Nat; n = 103), High Aesthetic Value Urban (HA-Urb; n = 104), Low Aesthetic Value Nature (LA-Nat; n = 103), Low Aesthetic Value Urban (LA-Urb; n=100), and Very Low Aesthetic Value Urban (VLA-Urb; n=102). A between-subjects design was chosen here over a within-subjects design (and in subsequent studies) for three primary reasons. First, in a within-subject design, image condition order may have affected how participants anchored their preference judgments in subsequent sessions and would have likely led to large order effects. Second, in a within-subject design, it is possible that having participants rate their state affect 12 times could lead to anchoring and adjustment of their self-reported affect, which may have produced a great deal of noise in the affect measurements. Third, in a within-subject design, it would likely not be possible to bring Mturk participants back for an additional 5 sessions. The full image sets and ratings (validated in Study 1) can be accessed at https://osf.io/ehtk9/.

2.2.3. State affect measures

To assess changes in state affect, two primary measures were used. The first consisted of a 6-item version of the State Trait Anxiety Inventory (STAI) (Marteau & Bekker, 1992) with 3 negative items (upset, tense, worried), and 3 positive low-arousal items (calm, relaxed, content). The presentation order of the 6 items in this scale was randomized across time points and participants. The second was a visual analog scale (VAS) for 4 emotion labels: happy, sad, inspired, and angry. The scale spanned from 1 to 100 at 1 unit intervals, and the order of the emotion labels was also randomized across time points and participants.

The composite STAI measures (STAI-Pos and STAI-Neg) were used in all reported analyses. A principal component analysis was also employed which allowed inclusion of all affect measures, and demonstrated very similar effects to the results of the STAI results reported. However, the data required standardization prior to performing the PCA, which affected the interpretability of changes in principal component values, so these analyses and results are detailed in the supplementary materials.

2.2.4. Procedure

Participants were randomly assigned to one of the 6 conditions (i.e., VHA-Nature, HA-Nature, LA-Nature, HA-Urban, LA-Urban, and VLA-Urban). After providing informed consent, participants filled out baseline measures of their affective state (T1). Though not analyzed in the current study, participants also completed a brief measure of nature connectedness before the questions measuring affective state. Subsequently, they viewed a series of 45 images and gave ratings of their aesthetic preference for the images on a 1–7 scale (1= strongly dislike, 7 = strongly like). Each image was on the screen for a minimum of 7 seconds, and the next image would appear after a rating for the image was selected. The image intervention took a minimum of 5 minutes and 25 seconds, and the average time taken by participants was 7 minutes and 20 seconds. Directly after the image rating task, participants completed the same measures of their affective state a second time (T2).

2.2.5. Analysis

A visual depiction of the hypotheses for this study as well as the analyses chosen to address each hypothesis are presented in Fig. 1. Before comparing affect change as a function of condition, preliminary analyses were conducted to examine whether the image sets used in the study elicited significant changes in emotional state. To this end, one sample t-tests were conducted on change scores for STAI-Pos and STAI-Neg, testing the null hypothesis is that there was no change between T1 (pre) and T2 (post). This was performed separately for each condition (Analysis 1).

To test whether changes in affect are due to naturalness (Hypothesis A: *Nature is unique*), whether they are purely due to preference

(Hypothesis B: *Aesthetic preference is what matters*), or a combination of the two (Hypothesis C: *Both are important*), four analytical approaches were employed. In those conditions which were completely crossed (i.e., HA-Nat and HA-Urb, LA-Nat and LA-Urb), both an ANOVA (Analysis 2: testing aesthetic preference conditions and environment type) and a linear regression (Analysis 3: using individual differences in preference and environment type) were conducted. In these completely crossed conditions, both main effects of environment and aesthetic preference could be examined, as could the interaction of the two. To accommodate the conditions that were not completely crossed, ANOVAs examining different aesthetic value conditions within an environment type were conducted (Analysis 4). Additionally, to analyze the role of individual differences in preference for the images, a linear regression predicting change in affect by participants' average preference rating (ignoring condition) was conducted (Analysis 5).

All statistical analyses were conducted in R v. 3.5.1 (R Core Team, 2019). ANOVAs were conducted using package 'ez' (v4.4-0, Lawrence, 2016), and post hoc comparisons were conducted using Tukey HSD multiple comparisons correction with a 95% family-wise confidence level. All between-subjects ANOVAs were first tested for significant heteroscedasticity using the Breush-Pagan test (function bptest in package 'lmtest';(Zeileis and Hothorn, 2002)). If identified, the ezANOVA option "white.adjust" was set to true, which uses a heteroscedasticity-corrected coefficient covariance matrix ('hccm' in package 'car'; Fox & Weisberg, 2019). Partial eta-squared effect size and 95% CIs were calculated using 'eta_sq' in package 'sjstats' (v0.17.6, Lüdecke, 2019). Cohen's d effect size and 95% CIs for comparison of group means were calculated using the 'apa.d.table' in package 'apaTables' (v.2.0.5, Stanley, 2018). Linear regressions were conducted using the 'lm' command in the 'stats' package (R Core Team, 2019).

Participants' preference for the images they viewed were determined by taking the average of their ratings for all 45 images they viewed. The average and standard deviation for each picture condition across all participants in that condition are presented in Table 2. As in Study 1, the average preference ratings for the equated image sets (HA-Nat and HA-Urb, LA-Nat and LA-Urb) were not significantly different from one another.

2.3. Results

2.3.1. Analysis 1: preliminary tests for significant affect change by condition

The results of these preliminary tests can be found in Table 3. Overall, the VHA-Nat, HA-Nat, HA-Urb, and VLA-Urb elicited a significant change in positive and negative state affect. However, this was not generally true for the LA images, with the exception of a modest increase in STAI-Pos in LA-Nat. All changes were in a positive direction (affect improvement) except for the VLA-Urban images, which had a deleterious effect on affective state.

2.3.2. Analysis 2: affect change in completely crossed data subset (aesthetic value as factor)

This analysis was conducted using only those conditions which were completely crossed (HA-Nat, HA-Urb, LA-Nat, LA-Urb), excluding the VHA-Nat and VLA-Urb conditions. To test the effect of environment and aesthetic value on affect change, 2 (Nat vs. Urb) x 2 (High vs. Low Aesthetic Value) factorial ANOVAs were conducted on change scores (T2 minus T1) for STAI positive and STAI negative scores. Results of these analyses did not yield significant effects of environment, aesthetic value condition, or the interaction for either STAI-Pos or STAI-Neg (all $p > 0.1$) [Fig. 2].

Average change in STAI-Pos (Left Panel) and STAI-Neg (Right Panel) between baseline (pre/T1) and after image viewing (post/T2) for each of the 6 image conditions. Error bars represent SEM.

Table 2
Study 2 Aesthetic Value Ratings of Images in Each Condition.
Ratings of all images in each condition from participants in Study 2. Aesthetic value ratings are on a 1–7 scale (1 = strongly dislike, 7 = strongly like).

Aesthetic Value Condition	Nature Conditions	Urban Conditions
	M (SD)	*M (SD)*
Very High Aesthetic Value	5.65 (0.84)	
High Aesthetic Value	4.78 (0.96)	4.55 (0.90)
Low Aesthetic Value	4.14 (1.18)	4.07 (0.79)
Very Low Aesthetic Value		3.26 (0.92)

2.3.3. Analysis 3: affect change in completely crossed data subset (aesthetic value as individual difference measure)

Analysis 3 was also conducted using only the data with completely crossed conditions, but using participants' own average ratings for the images they viewed, rather than treating aesthetic value as a factor. To do this, two multiple regressions were performed predicting STAI-Pos and STAI-Neg by environment type, average image rating, and their interaction.

For STAI-Pos the overall model was not significant ($R^2 = 0.015$, $F(3,406) = 2.07$, $ps = 0.10$), though participants' average preference ratings were significantly predictive of change in STAI-Pos ($B = 0.07$, 95% CI [0.01, 0.14], $p=0.03$). However, neither environment type nor the interaction of environment and individual preference ratings were significant (all $p > 0.41$). For STAI-Neg, the overall model was also not significant ($R^2 = 0.005$, $F(3,406) = 0.70$, $ps = 0.55$) and none of the predictors had a significant effect on the outcome variable (all $p > 0.49$).

2.3.4. Analysis 4: affect change between aesthetic value conditions within an environment

To handle the conditions which were not completely crossed in our design (VHA-Nat and VLA-Urb), separate factorial ANOVAs were conducted on each of the three aesthetic value levels for nature and urban images.

2.3.4.1. Nature conditions.

A one-way ANOVA with the 3 aesthetic value levels (Very High, High, Low) in the nature condition was performed for change in STAI-Pos and STAI-Neg. Results of this ANOVA for STAI-Pos yielded a trending effect of aesthetic value level $F(2, 306) = 2.88$, $p = 0.057$, $\eta_p = 0.018$, 95% CI [0.0, 0.054]. The partial eta-squared indicates a small effect size. Post hoc comparisons were conducted and family-wise error corrected using Tukey's HSD, which showed a significant difference between the VHA-Nat ($M = 0.29$, $SD = 0.54$) and LA-Nat conditions ($M = 0.12$, $SD = 0.52$, $p = 0.045$, $d = 0.34$, 95% CI [0.06, 0.61]), indicating a greater increase in STAI-Pos for those in the VHA-Nat condition relative to those in the LA-Nat condition. However, there were no differences between VHA-Nat and HA-Nat ($M = 0.22$, $SD = 0.48$, $ps = 0.55$) or between HA-Nat and LA-Nat ($ps = 0.37$).

Results of this analysis for STAI-Neg showed a significant effect of aesthetic value, $F(2,306) = 3.27$, $p = 0.039$, $\eta_p^2 = 0.021$, 95% CI [0.0, 0.058]. The partial eta-squared indicates a small effect size. Post hoc comparisons showed a significant difference between VHA-Nat ($M = -0.16$, $SD = 0.38$) and LA-Nat ($M = -0.03$, $SD = 0.42$, $p = 0.036$, $d = 0.32$, 95% CI [0.05, 0.60]). This difference indicates a greater reduction in STAI-Neg for participants in the VHA-Nat condition relative to LA-Nat. No significant difference was found between VHA-Nat and HA-Nat ($M = -0.12$, $SD = 0.32$, $ps = 0.77$) or between HA-Nat and LA-Nat ($ps = 0.17$) [Fig. 2].

2.3.4.2. Urban conditions.

For the urban images, a one-way ANOVA with the 3 aesthetic values (High, Low, Very Low) was also performed for change in STAI-Pos and STAI-Neg. Results of the first ANOVA indicated a significant effect of aesthetic value level on STAI-Pos, $F(2, 303) = 6.08$, $p = 0.003$, $\eta_p = 0.039$, 95% CI [0.005, 0.086]. Here, the value of the partial eta-squared indicated a small-to-medium effect size. Post hoc

Table 3
Study 2 STAI Results: Change in affective state relative to baseline.
Results of 1-sample t-tests comparing STAI positive and negative change to zero in each of the 6 conditions. **Significant p-value with Bonferroni family-wise multiple comparisons correction (α = 0.008) *Significant p-value uncorrected (α = 0.05).

		STAI-Positive			STAI-Negative		
Condition	n	*t-statistic*	*p-value*	*Cohen's d*	*t-statistic*	*p-value*	*Cohen's d*
VHA-Nat	103	5.46**	<0.001	0.54	—4.28**	<0.001	0.42
HA-Nat	103	4.22**	<0.001	0.41	—3.90**	<0.001	0.38
HA-Urb	104	2.42*	0.017	0.24	—2.83**	0.006	0.28
LA-Nat	103	2.54*	0.012	0.25	—0.70	0.486	0.07
LA-Urb	100	1.86	0.065	0.19	—1.84	0.069	0.18
VLA-Urb	102	—2.14*	0.034	0.21	2.41*	0.018	0.24

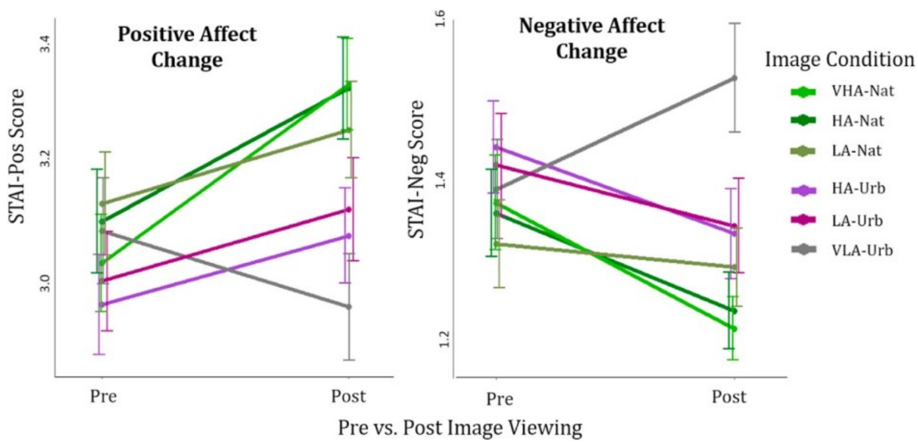

Fig. 2. Study 2 change in affect by image condition.

comparisons were conducted and family-wise error corrected using Tukey's HSD, which showed a significant difference between the VLA-Urb (M 0.12, SD 0.59) and LA-Urb conditions (M 0.12, SD = 0.63, p = 0.007, d = 0.40, 95% CI [0.12, 0.68]) as well as between VLA-Urb and HA-Urb (M = 0.11, SD = 0.47, p = 0.008, d = 0.44, 95% CI [0.17, 0.72]), indicating that participants in the VLA-Urb condition showed less of an improvement in STAI-Pos compared to HA-Urb or LA-Urb. No difference was found for STAI-Pos change between HA-Urb and LA-Urb (ps = 0.99).

Results of this analysis for STAI-Neg also showed a significant effect of aesthetic value, $F(2,303) = 8.37$, $p < 0.001$, $\eta_p^2 = 0.052$, 95% CI [0.012, 0.104]. The partial eta-squared indicates a medium effect size. Post hoc comparisons yielded a significant difference between VLA-Urb (M = 0.14, SD = 0.59) and LA-Urb (M = —0.08, SD = 0.42, p = 0.003, d = 0.42, 95% CI [0.15, 0.7]), as well as between VLA-Urb and HA-Urb (M = —0.11, SD = 0.39, $p < 0.001$, d = 0.5, 95% CI [0.22, 0.78]), but not between LA-Urb and HA-Urb (ps = 0.88). These results suggest a larger reduction in STAI-Neg for participants in the HA-Urb and LA-Urb conditions compared to those in the VLA-Urb condition [Fig. 2].

2.3.5. Analysis 5: affect change as predicted by individual preference ratings in full dataset

To examine whether individual differences in participants' preference ratings for the images were related to changes in affect, two linear regressions were conducted to predict change in STAI-Pos and STAI-Neg by individuals' average preference rating.

Analyses conducted on the full dataset (not only on the completely-crossed conditions) showed that average image preference rating explained 5.5% of the variance in STAI-Pos change (R^2 = 0.055, F(1,613) = 35.92, $p < 0.001$). In this case, a higher average preference rating for the images viewed significantly predicted a greater increase in STAI-Pos β = 0.24, 95% CI [0.16, 0.31], $p < 0.001$). Additionally, average image preference explained 2.4% of the variance in STAI-Neg change (R^2 = 0.024, F(1,613) = 14.76, $p < 0.001$). Here, higher preference ratings significantly predicted a greater decrease in STAI-Neg β = —0.15, 95% CI [-0.23, —0.07], $p < 0.001$).

2.4. Discussion

Study 2 failed to demonstrate a significant effect of environment type on changes in either positive (STAI-Pos) or negative (STAI-Neg) affect for stimuli where preference was equated. That is, although affect change differed slightly between HA-Nat and LA-Urb, and between LA-Nat and LA-Urb, the HA-Nat and HA-Urb were not significantly different from one another, nor were the LA-Nat and LA-Urb (Analysis 2). There was modest evidence for individuals' own preference ratings as a predictor of positive but not negative affect change in these completely crossed conditions (Analysis 3). However, the overall experimental design included conditions which were not completely crossed (i.e. VHA-Nat and VLA-Urb). When examining differences in aesthetic value within an environment type, these more extreme aesthetic value conditions (VHA and VLA) yielded significantly larger changes in affective state compared to the HA or LA conditions. Specifically, VHA-Nat lead to greater improvements in both positive and negative affect relative to LA-Nat, and VLA-Urb lead to worsened positive and negative affect relative to both LA-Urb and HA-Urb (Analysis 4). Furthermore, participants' own preference ratings of the images were significantly predictive of change in both the positive and negative affect change when analyzed in the full dataset (Analysis 5).

While the greatest positive affect changes due to image condition were found in the VHA-Nat condition, Study 1 did not yield a sample of urban images that were as highly preferred as these nature stimuli to create a comparable condition in Study 2. Studies 3 and 4 were designed to address this problem by finding other image types that were as preferred as the VHA-Nat condition, but qualitatively different in context from the nature scenes.

3. Study 3: Very high aesthetic images validation

3.1. Study Intro

To examine whether the significant changes in state affect found for the Very High Aesthetic value Nature condition (VHA-Nat) were due to the images being very highly preferred or due to something else specific to nature (over and above aesthetic preference), additional image sets were required that would include multiple categories of images also rated very highly. Study 3 was conducted to generate the stimulus sets needed to compare against the VHA-Nat condition. Another environmental condition (i.e. very high aesthetic urban) was the most relevant category, but highly preferred images in other, qualitatively different contexts were also examined in Study 3: Animals, Space, and Abstract Art.

3.2. Method

3.2.1. Participants

409 US-based adults (202 male, 206 female, 1 not reported) were recruited from Amazon Mechanical Turk. Age of the participants ranged from 19 to 74 years (*M*=36.9, *SD* =10.7). The full study procedures were expected to take approximately 15 min and participants were compensated $1.50 for participating. Informed consent was administered by the University of Chicago Institutional Review Board.

3.2.2. Procedure

Participants were randomly assigned to rate 75 images on a 1–7 Likert scale evaluating each photo on preference (anchors of 1 = "strongly dislike" and 7 = "strongly like"; n = 206) or naturalness (anchors of 1 = "very man-made" and 7 = "very natural"; n =203). The images in Study 3 were also taken from the SUN database (Xiao et al., 2010) as well as an online image search for non-copyrighted scene images. As in Study 1 and 2, all scenes were required to have sufficiently high resolution for clear viewing and do not contain people or text. In addition, for the animals category, selected images were gathered from a variety of animal types (not simply "canonical pets") and contained minimal background natural scenery. For the space and abstract art images, the main goal was to find aesthetically pleasing images which were somewhat varied (i.e., space images that were not all just the Milky-way galaxy or a night sky, and abstract images that had varied patterns and color profiles). Each participant saw 15 images across five categories of images: animals, space, abstract patterns, natural environments, and urban environments. Images were pulled randomly, and each image received a minimum of 25 ratings on each attribute with an average of roughly 40 ratings per image.

3.3. Results

Average preference ratings for the 75 images in each category are plotted in Fig. 3 (Top Panel). Based on these ratings, only two categories

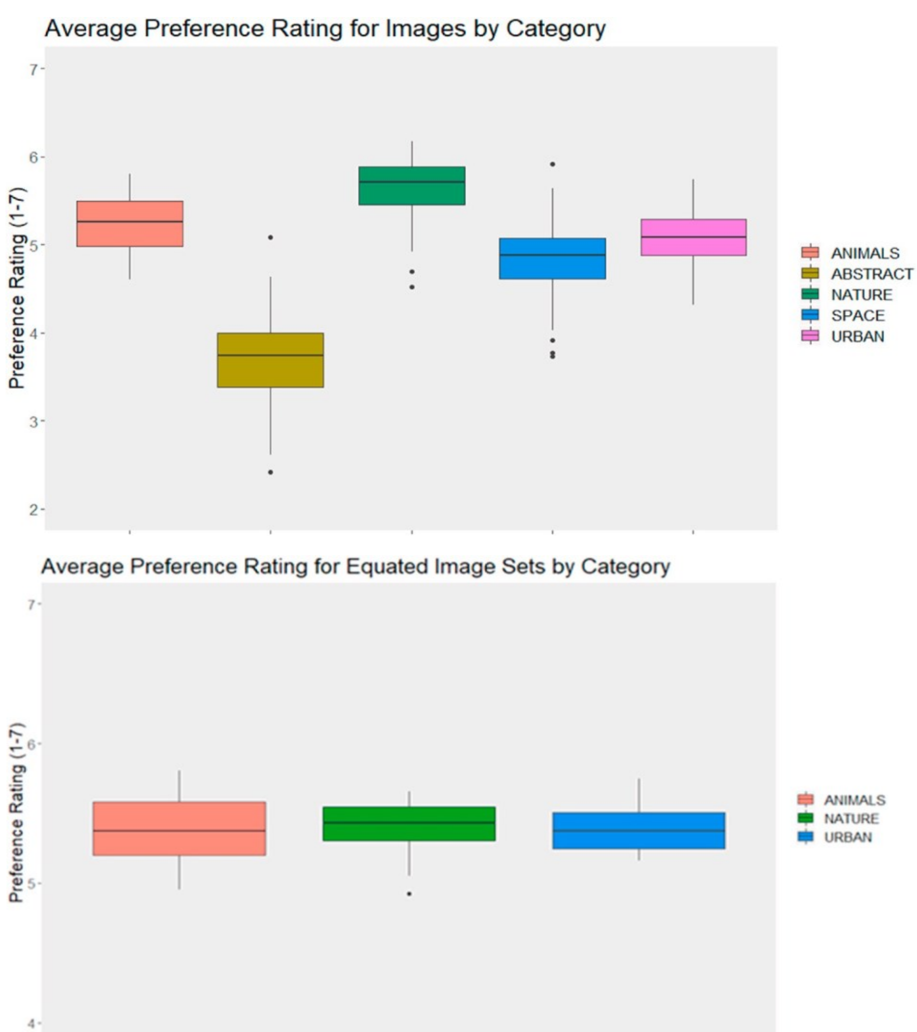

Fig. 3. Image Preference results from Study 3.
Top Panel: Boxplots of preference ratings for all 75 images in each category examined.
Bottom Panel: Boxplots of preference-equated image sets (30 images each) in the Nature, Urban, and Animal image categories. Abstract images and images of space were significantly lower in preference and did not yield enough preference-equated images to create another stimulus set.

of images (Animals and Urban) had sufficiently overlapping distributions to extract preference-matched image sets with the VHA-Nat images. The images sets used for Study 4 involved 30 images each of nature, urban, and animal images with similar average preference ratings and standard deviations (Fig. 3 Bottom Panel; Table 4).

4. Study 4: Examining affect change in very highly preferred stimuli

4.1. Study Intro

Study 4 sought to determine whether very highly preferred nature images would cause the same improvement in state affect as equally preferred images from other categories. That is, Study 4 tested whether the scenic nature category itself was an additional source of affective benefit above and beyond aesthetic preference. It's worth clarifying that although animals are certainly part of the 'natural' world, the previous literature examining nature's cognitive and affective benefits focuses primarily on natural environments in terms of the physical spaces/ scenes rather than the living organisms which might inhabit or visit such spaces. Thus, we are making a distinction here between scenic nature (VHA-Nature) and domestic nature (VHA-Animals).

In addition, Study 4 employed a negative mood induction procedure (MIP) in half of the participants to examine if baseline mood might impact how effective the VHA images are at improving affective state. For example, it might have been the case that as participants did not have high baseline negative affect in Study 2, they may have been less affected by the stimuli. In Study 2, the average STAI-Neg rating at baseline was 1.3 (on a 1–4 scale), the baseline average score for VAS Sad was 11, and for VAS Angry it was 6 (both on a 1–100 scale). Positive affect also started out relatively high in Study 2 (average STAI-Pos was 3.0, average VAS Happy was 62, and average VAS Inspired was 45). As such, in Study 4, the Negative MIP was included to ensure that any effects (or lack thereof) were not simply due to ceiling/floor effects of baseline affect.

Two empirical questions were examined in Study 4. The main question was whether the improvement in affect found for VHA-Nat in Study 2 was due to nature itself or simply due to preference. The secondary question was whether or not baseline affect influenced whether scenic nature had an additional emotional benefit above and beyond preference. A visual depiction of the hypotheses for this study as well as the analyses chosen to address each hypothesis are presented in Fig. 5.

4.2. Method

Study 4 was pre-registered on OSF prior to data collection: https://osf.io/u5r4c. The pre-registration included a PCA for data reduction (results of which can be found in supplementary materials), but due to lessened interpretability of this approach (as in Study 2), the analyses presented diverge from the pre-registered ones. The data for Study 4 are publicly available at: https://osf.io/ehtk9/.

4.2.1. Participants

602 US-based adults (271 male, 327 female, 1 other, 3 not reported) were recruited from Amazon Mechanical Turk. Age of the participants ranged from 19 to 69 years (M = 37.5, SD = 11.0). The full study procedures were expected to take approximately 15 min and participants were compensated $1.50 for participating. Informed consent was administered by the University of Chicago Institutional Review Board. The sample size was determined prior to data collection (see pre-registration) and was partially based on recreating the conditions of Study 2.

4.2.2. Experimental conditions

Participants were randomly assigned to one of 3 image conditions based on the results of Study 3 (Anim, Nat, Urb) and one of 2 mood induction procedure groups (Negative, Neutral). The study design was fully between subjects with 6 conditions: Anim-Negative (n= 99), Anim-Neutral (n=102), Nat-Negative (n=102), Nat-Neutral (n= 103), Urb-Negative (n=97), Urb-Neutral (n = 99). The full image sets and ratings (validated in Study 3) can be accessed at: https://osf.io/ehtk9/.

4.2.3. Procedure

After informed consent was obtained, participants first completed the mood induction procedure (MIP). Participants were provided with a brief description prior to reading a short story: "In this study we would like to examine how different types of storytelling influence, thoughts, feelings, and judgments. On the next page we have a short story for you to read. The story will be on screen for a minimum of 2 minutes. Please try to minimize distraction and take your time reading the story completely before continuing." The negative story was taken from a report of a young woman whose father died after suffering from Alzheimer's dementia, and the neutral story was a short excerpt from the introductory chapter of *A Brief History of Time* (Hawking, 1988). The two stories were validated in a previous study evaluating the efficacy of several text-based online mood inductions (Verheyen & Göritz, 2009).

Following the MIP, participants filled out baseline measures of their affective state (T1). The measures used were identical to those collected in Study 2. Participants then viewed a series of 30 images and gave ratings of their aesthetic preference for the images on a 1–7 scale (1 = strongly dislike, 7 = strongly like). Each image was on the screen for a minimum of 7 seconds, and the next image would appear after a rating for the image was selected. The image intervention took a minimum of 3 minutes and 30 seconds. After the image rating task, participants completed the same measures of their affective state a second time (T2), [See Fig. 4 for study design].

4.2.4. Analysis

A visual depiction of the analytic approach for this study is presented in Fig. 5. The main question of Study 4 was whether the improvement in affect found for VHA-Nat in Study 2 was due to scenic nature itself (Hypothesis A: *Nature is unique*) or simply due to preference (Hypothesis B: *Aesthetic Preference is what matters*). The secondary question was whether baseline affect influenced whether nature has an additional emotional benefit above and beyond preference (Hypothesis A: *Baseline mood matters*) or isn't important (Hypothesis B: *Baseline mood is irrelevant*). The secondary question examined an interaction between preference-equated categories and mood induction. However, in Study 4, participants' ratings diverged somewhat from Study 3, resulting in significantly different preference ratings between categories (described below, see Table 5 and Fig. 6). This issue prevented a strong test of the main effect of preference-equated category, as well as the interaction between MIP and image category in the ANOVA used in Analysis 1. Therefore, linear regression (Analysis 2) which tested the independent predictive value of mood induction procedure, image category, and participants' own preference ratings was conducted as well. The same analysis tools in 'R' and relevant statistical procedures used in Study 2 were also employed in Study 4.

Participants' preferences for the images they viewed were determined by taking the average of their ratings for all 30 images they viewed. The average and standard deviation for each picture condition across all participants in that condition are presented in Table 5. Ratings

Table 4
Study 3 Ratings for the Preference-equated Image Sets.
Based on the participants' ratings in Study 3, three categories of preference-equated images were used in Study 4. Average preference and naturalness ratings for these image sets (30 images each), presented below.

	Nature	Urban	Animals
Preference M (SD)	5.41 (0.18)	5.39 (0.16)	5.39 (0.26)
Naturalness M (SD)	6.52 (0.21)	2.73 (0.36)	6.47 (0.23)

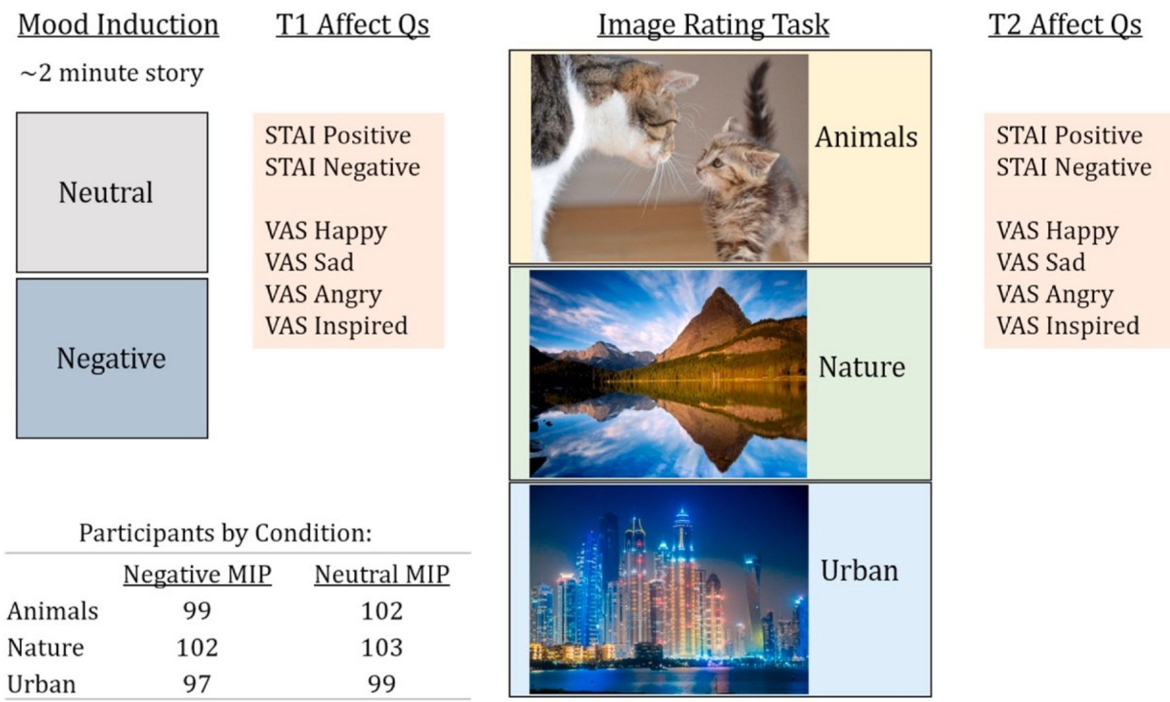

Fig. 4. Study 4 design.

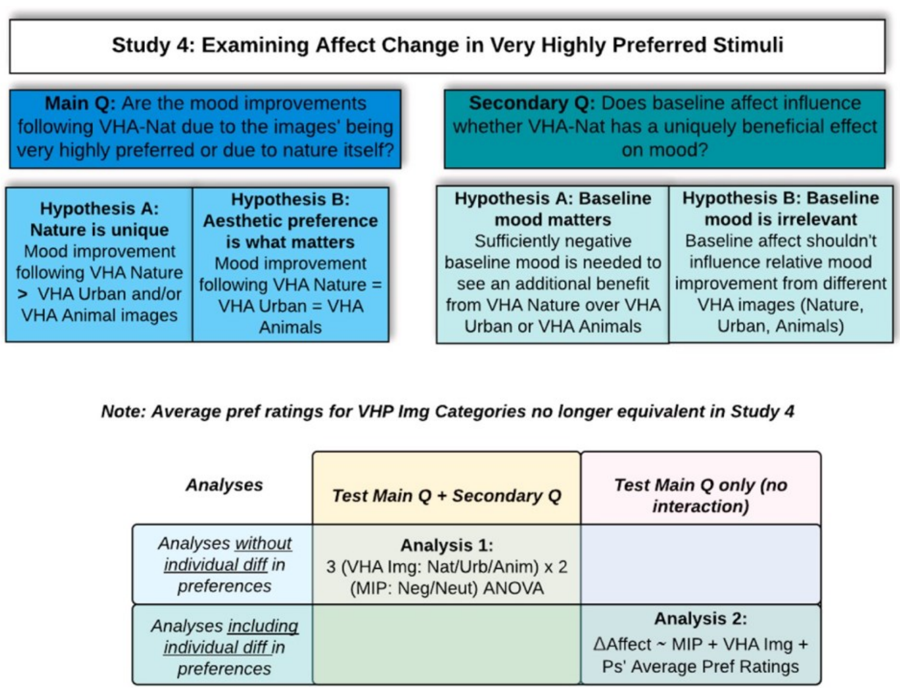

Fig. 5. Hypotheses & analyses overview for study 4.

Table 5
Study 4 participants' preference ratings.

Mean & SD by picture condition, collapsed across MIP conditions.

	Nature	Urban	Animals
Preference M (SD)	5.69 (0.85)	5.16 (0.96)	5.89 (0.88)

from participants in this study (Study 4) were somewhat different from those in image validation Study 3, which may have been due to rating a single image category rather than a variety of image categories [Fig. 6]. As a result our image categories were no longer fully equated on preference.

4.3. Results

4.3.1. Analysis 1: affect change by image condition and mood induction

To test the effect of image category and mood induction on affect change, a factorial ANOVA with Very High Aesthetic Value Image Category (Nature, Urban, Animal) and Mood Induction (Negative, Neutral) as between-subjects variable were conducted on change (T2 minus T1) for STAI-Neg and STAI-Pos.

Results of this ANOVA for STAI-Pos yielded a main effect of Mood Induction $F(1,596) = 152.6$ $p < 0.001$, $\eta^2_p = 0.21$, 95% CI [0.15, 0.26]. The partial eta-squared indicates large effect of MIP Condition. Participants who were first inducted into a negative mood showed a larger increase in STAI-Pos ($M = 0.85$, $SD = 0.81$) relative to those in the neutral MIP ($M = 0.16$, $SD = 0.51$). There was not a significant effect of VHA Image Category ($ps = 0.34$) or an interaction of VHA Image Category and MIP ($ps = 0.9$).

Similar results were found for STAI-Neg. A main effect of MIP was found ($F(1,596) = 246.0$, $p < 0.001$, $\eta^2_p = 0.30$, 95% CI [0.24, 0.35], where again, participants first induced into a negative affective state showed a greater reduction in negative affect ($M = -0.81$, $SD = 0.76$) relative to those in the neutral mood induction ($M = -0.04$, $SD = 0.37$). The partial eta-squared indicates a large effect of mood induction type on change in STAI-Neg. No significant effect of VHA Image Category ($ps = 0.87$) or interaction of VHA Image Category and MIP ($ps = 0.39$) was found.

4.3.2. Analysis 2: affect change as predicted by image condition, mood induction, & individual preference ratings

To see whether individual differences in participants' preference ratings influenced change in affect, multiple regressions were conducted. With STAI-Pos and STAI-Neg as outcome variables, the regression analyses examined the respective contributions of mood induction procedure, VHA image category, and individuals' average preference ratings.

Results of this analysis on STAI-Pos are presented in Table 6A. The overall model was significant and explained 23.6% of the variance in change in positive affect ($R^2 = 0.236$, $F(4,597) = 46.11$, $p < 0.001$). In this case, a higher average preference rating for the images viewed significantly predicted a greater increase in STAI-Pos ($B = 0.14$, 95% CI [0.08, 0.20], $p < 0.001$). Being induced into a negative mood was also a significant predictor ($B = -0.67$, 95% CI [-0.78, -0.56], $p < 0.001$), but image category did not significantly contribute to positive affect change [Fig. 7].

Results of this analysis on STAI-Neg are presented in Table 6B. As in positive affect, the overall model was significant and explained 30.8% of the variance in STAI-Neg change ($R^2 = 0.308$, $F(4,597) = 66.36$, $p < 0.001$). A higher average preference rating for the images viewed significantly predicted a greater reduction negative affect ($B = -0.09$, 95% CI [−0.14, −0.03], $p = 0.001$), as did being in the negative mood induction group ($B = 0.77$, 95% CI [0.67, 0.86], $p < 0.001$). Importantly, image category did not significantly contribute to change in STAI-Neg [Fig. 7].

Table 6A
Panel A. Study 4 Regression results using STAI-Pos Change as the criterion. For image type (Pic_Con) the baseline condition was Animals.

Predictor	*b*	*b* 95% CI [LL, UL]	sr^2	sr^2 95% CI [LL, UL]	Fit
(Intercept)	0.05	[-0.32, 0.42]			
Avg_Pic_Rate	0.14**	[0.08, 0.20]	.03	[.00, .05]	
Pic_Con_Nat	-0.02	[-0.15, 0.11]	.00	[-.00, .00]	
Pic_Con_Urb	0.01	[-0.13, 0.14]	.00	[-.00, .00]	
MIP_Con	-0.67**	[-0.78, -0.56]	.19	[.14, .25]	
					R^2 = .236** 95% CI [0.18, 0.29]

A significant b-weight indicates the semi-partial correlation is also significant. b represents unstandardized regression weights. sr2 represents the semi-partial correlation squared. LL and UL indicate the lower and upper limits of a confidence interval, respectively.

Table 6B
Panel B. Study 4 Regression results using STAI-Neg Change as the criterion For image type (Pic_Con) the baseline condition was Animals.

Predictor	*b*	*b* 95% CI [LL, UL]	sr^2	sr^2 95% CI [LL, UL]	Fit
(Intercept)	-0.29	[-0.62, 0.04]			
Avg_Pic_Rate	-0.09**	[-0.14, -0.03]	.01	[.00, .03]	
Pic_Con_Nat	-0.03	[-0.14, 0.09]	.00	[-.00, .00]	
Pic_Con_Urb	-0.09	[-0.22, 0.03]	.00	[-.00, .01]	
MIP_Con	0.77**	[0.67, 0.86]	.29	[.23, .35]	
					R^2 = .308** 95% CI [0.26, 0.36]

A significant b-weight indicates the semi-partial correlation is also significant. b represents unstandardized regression weights. sr2 represents the semi-partial correlation squared. LL and UL indicate the lower and upper limits of a confidence interval, respectively.

4.4. Discussion

The results of Study 4 were consistent with those of Study 2. Study 4 found that very highly preferred image category (in this case, scenic nature, urban, or animals) did not have a differential impact on affect change, i.e., they all improved affect to the same degree. Individual differences in how much participants liked the images they saw did predict improvement in both STAI-Pos and STAI-Neg which was also found in Study 2. Participants who were first induced into a negative mood showed larger changes in affect but this did not interact with VHA image category. Thus, the results of both studies provide support for a preference-based account of mood change rather than the effects being specific to scenic nature stimuli (i.e., nature-based mood effects do not differ from other very highly preferred stimulus categories).

5. Study 5: Examining the relationship between beauty and affinity

5.1. Study Intro

In Studies 2 and 4, affect change was driven by differences in preference rather than environment/category. In each of these studies, participants had evaluated preference on a scale that assesses affinity for the images (i.e., how much do you like/dislike the image). Up to this point, we have been assuming that preference (affinity) and aesthetics (beauty) are the same construct. Indeed, the terms aesthetics, affinity, and preference, are often used interchangeably (van den Berg et al., 2003; Staats et al., 2003; Ulrich, 1983). However, it is still possible that there is something special (i.e., rewarding, pleasing, or affinity-inducing) about natural environments above and beyond aesthetics (beauty) that causes them to be preferred (liked). For example, (Valtchanov and Ellard, 2015) propose that natural stimuli are endogenously visually rewarding. If this is the case, ratings of beauty and ratings of affinity (liking) may not be identical overall or may be different when examined in nature scenes versus in urban scenes. To test this, all 375 images rated on affinity (Study 1) were also rated on aesthetics/beauty in a new sample (Study 5). If the affinity and beauty ratings are not identical in these images, this would suggest that participants' affinity ratings in Study 1 are due to something other than aesthetics (i.e. endogenous visual reward not captured by perceived beauty). Further, if the nature images and urban images differ in how correlated affinity and beauty are, this might suggest something categorically different in how participants evaluate nature images and urban images. However, if the two are highly correlated across all images and within category (nature vs. urban) this would suggest that the preference ratings are primarily evaluations of aesthetics.

5.2. Method

Study 5 was pre-registered on OSF prior to data collection (https://osf.io/u2e6n), though the analyses reported in this paper were not initially detailed in this pre-registration. The data for Study 5 are publicly available at: https://osf.io/ehtk9.

5.2.1. Participants

194 US-based adults (94 male, 100 female) were recruited from Amazon Mechanical Turk. This sample does not include 9 workers who provided feedback to indicate they encountered technical problems or were distracted during the study, or showed no variation in responding (i.e., gave the same rating to every image). Age of the participants ranged from 19 to 72 years (M = 38.3, SD = 12.4). The full study procedures were expected to take approximately 15 min and participants were compensated $1.50 for participating. Informed consent was administered by the University of Chicago Institutional Review Board.

5.2.2. Procedure

Study 5 was designed to match the conditions of the original image rating procedure in Study 1 as closely as possible. Therefore, all 375 images from Study 1 were used in Study 5. Each image in Study 1 was rated on preference (affinity, i.e. "How much do you like or dislike this image?") by roughly 51 individuals, and each participant saw 100 images in a session. As in Study 1, participants in Study 5 saw 100 images pulled pseudorandomly, attempting to show a relatively equal number of natural and urban images as well as images that varied on aesthetic value. In this study, each image received a minimum of 35 beauty ratings (i.e., "How ugly or beautiful is this image?") with an average of roughly 51 ratings per image.

5.2.3. Analysis

To test whether participants' ratings of aesthetics (beauty) in Study 5 were similar to affinity ratings from Study 1, correlations between images' average beauty ratings (Study 5) and affinity ratings (Study 1) were conducted using the 'cor' function in R. The analyses specified in the pre-registration (testing for differences in beauty in image sets that were matched on preference) were also conducted initially and are reported in the supplementary materials, but ultimately the correlation analyses in the full image dataset were more informative, so they are the primary results reported.

5.3. Results

When examined across all 375 images, the correlation between beauty ratings and affinity ratings was r = 0.97 [Fig. 8]. Additionally, when examined within environment type, the correlations were similar in magnitude. For nature images this correlation was r = 0.96 and for urban images the correlation was r = 0.95.

5.4. Discussion

The results of Study 5 failed to find support for the hypothesis that there is something special, unique, or "rewarding" (Valtchanov & Ellard, 2015) about the nature scenes in our study which make them preferred (liked) above and beyond aesthetics. These results demonstrated a near perfect correlation between affinity and aesthetics for the full sample of images, as well as when broken up by environment type. Therefore, our data do not suggest that there is some missing factor that is specific to natural environments which makes them preferred (liked) in our study.

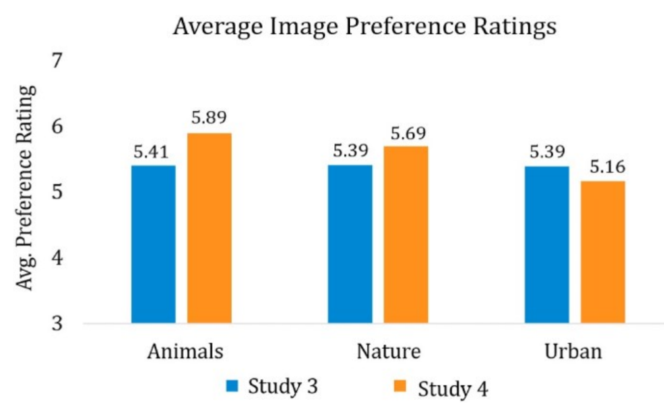

Fig. 6. Average preference ratings by category for Study 3 and 4.

6. Study 6: Replication of study 2 with naturalness ratings

6.1. Study Intro

Study 2 tested the roles of preference and environment type in predicting affect change and found greater evidence for a role of preference than environment. However, in this study (and Study 4) participants were asked to make preference ratings as they were rating the images, potentially priming the importance of preference. Therefore, Study 6 was conducted using the identical study procedures to Study 2, with the exception of the image rating task, where participants were asked to rate the images on naturalness. If cuing participants to the importance of preference is what drove the effects found in Study 2, then priming participants on naturalness should 1) remove or greatly reduce any preference effects and/or 2) lead to an effect of environment type. In contrast, if the question used in the image rating task is not important, this should replicate the results of Study 2.

6.2. Method

The data for Study 6 are publicly available at: https://osf.io/ehtk9/.

6.2.1. Participants

607 US-based adults (376 male, 228 female, 3 other) were recruited from Amazon Mechanical Turk. Age of the participants ranged from 19 to 74 years (M = 37.6, SD = 11.2). The full study procedures were expected to take approximately 15 min and participants were compensated $1.50 for participating. Informed consent was administered by the University of Chicago Institutional Review Board. Sample size was chosen to match that of Study 2.

6.2.2. Conditions & procedure

Participants were randomly assigned to one of the 6 conditions: VHA-Nat (n = 102), HA-Nat (n = 102), LA-Nat (n = 99), HA-Urb (n = 101), LA-Urb (n = 99), VLA-Urb (n = 104). The experimental procedure, measurements, and analyses were identical to those of Study 2, with the exception of the image rating task, where participants rated the images on naturalness (1 = "very man-made", 7 = "very natural"). Table 7 shows the mean and standard deviation of the naturalness ratings in Study 6. Unfortunately, the average naturalness ratings here (when rating only a single environment type) were substantially changed from those collected in Study 1 (where participants rated both nature and urban scenes). The overall range of responses for participants in the nature conditions was reasonable (98% reported average naturalness ratings over 4) though the average naturalness (5.94) was much lower than in Study 1 (6.67). However, the range for urban participants was much more variable, with only 73% providing values less than 4, meaning over ¼ of participants perceived the city images as being more natural than man-made. Due to this and a much higher average

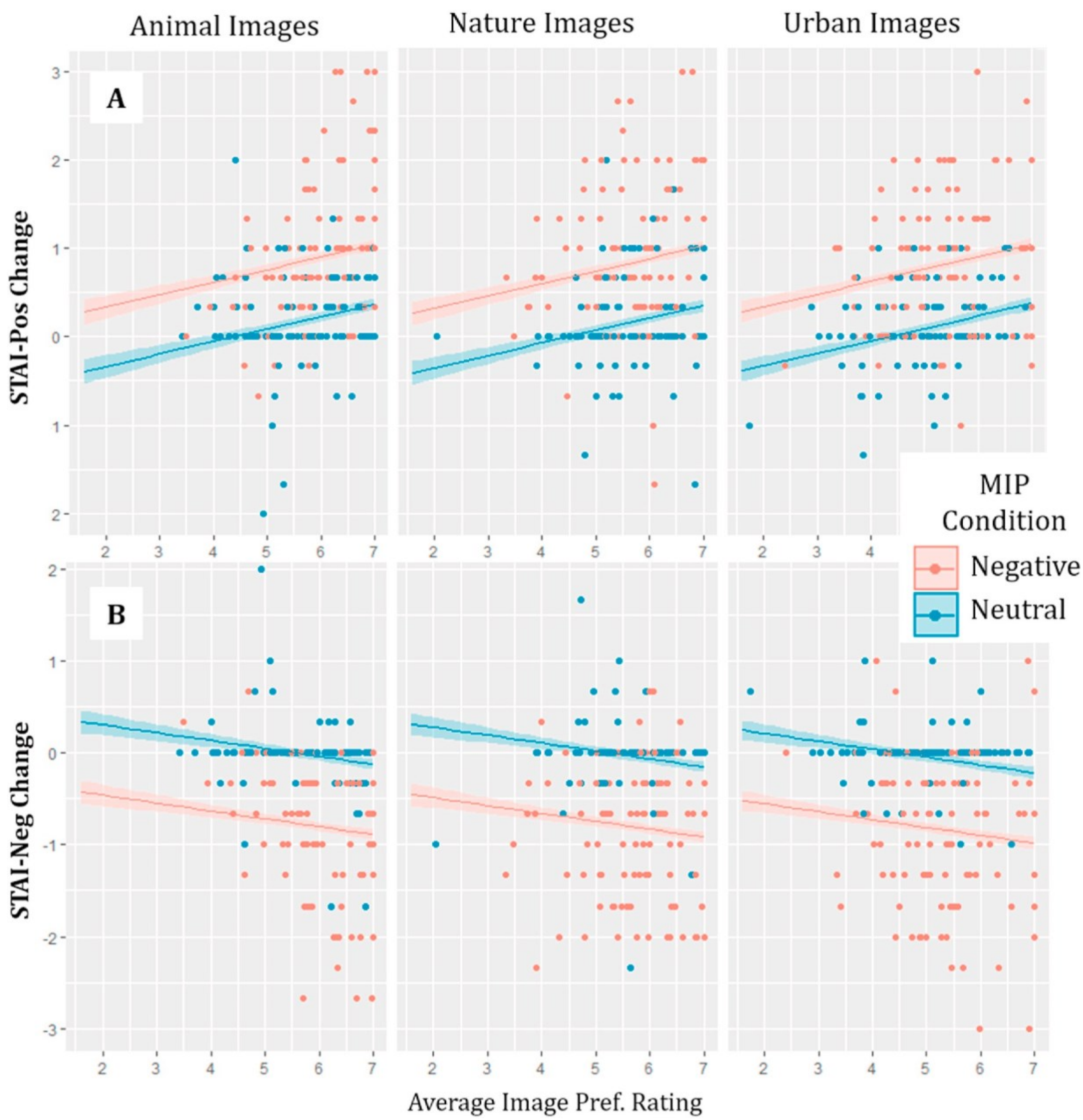

Fig. 7. Study 4 Multiple regression plots for affect change.
Regression plots predicting change in STAI-Pos (A) and STAI-Neg (B) by VHA Image Category + Participants' Average Image Rating + Mood Induction (MIP) Condition.

naturalness rating in this study (2.75) than in Study 1 (1.42), it seems likely that many participants either 1) didn't understand the question in the context of the stimuli they were viewing (i.e., one participant commented that they were evaluating the extent to which the image looked edited/doctored), or 2) used a different criterion to evaluate naturalness than what would be used when presented with both environments. As its therefore unclear what these naturalness ratings reflect, the analyses reported here will focus on the categorical factors (environment type and aesthetic value level) rather than incorporating individuals' average naturalness ratings.

6.3. Results

6.3.1. Analysis 1: preliminary tests for significant affect change by condition

The results of these preliminary tests can be found in Table 8. Overall, all conditions except for VLA-Urb elicited a significant improvement in STAI-Pos. Results for positive affect mostly replicated those in Study 2, though most of the effect sizes were somewhat larger. Additionally, while in Study 2, the LA-Urb did not lead to a significant increase in positive affect, it did in Study 6. The results for STAI-neg diverged from Study 2, however. In this case, no condition significantly improved negative affect (i.e., lower STAI-neg after picture viewing relative to baseline). However, the low and very low aesthetic value conditions all lead to an increase in negative affect. This discrepancy is likely due to higher levels of baseline negative affect when conducting this study, the data for which were gathered on June 1, 2020 amid protests against police brutality towards Black Americans and the global Covid-19 pandemic.

6.3.2. Analysis 2: affect change in completely crossed data subset

This analysis was conducted using only those conditions which were completely crossed (HA-Nat, HA-Urb, LA-Nat, LA-Urb), excluding the VHA-Nat and VLA-Urb conditions. To test the effect of environment and

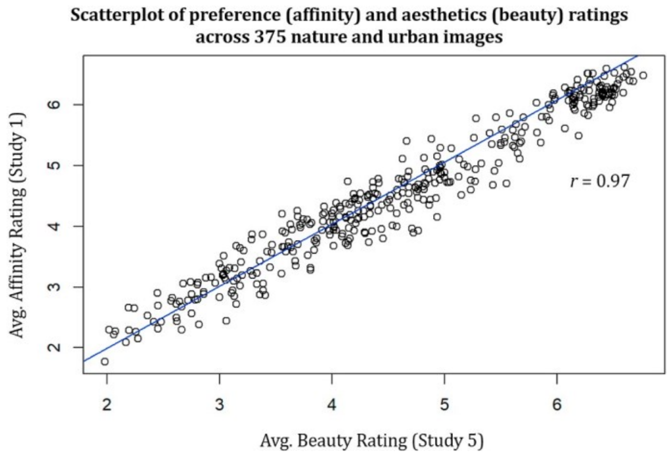

Fig. 8. Correlation between preference and aesthetics.

Table 7
Study 6 Naturalness Ratings of Images in Each Condition Average ratings of all images in each condition from participants in Study 1 and Study 6. Naturalness ratings are on a 1–7 scale (1 = very man-made, 7 = very-natural).

Aesthetic Value Condition	Nature Conditions	Urban Conditions
	M (SD)	M (SD)
Very High Aesthetic Value	5.96 (0.96)	
High Aesthetic Value	6.05 (0.73)	2.79 (1.81)
Low Aesthetic Value	5.8 (0.86)	2.70 (1.77)
Very Low Aesthetic Value		2.76 (1.95)

aesthetic value on affect change, 2 (Nat vs. Urb) x 2 (High vs. Low Aesthetic Value) factorial ANOVAs were conducted on change scores (T2 minus T1) for STAI positive and STAI negative scores. Results of these analyses did not yield significant effects of environment, aesthetic value condition, or the interaction for STAI-Pos (all $p > 0.2$). For STAI-Neg, a marginal effect of aesthetic value was found, $F(1,397) = 2.99$, p $= 0.08$, $\eta_p = 0.007$, 95% CI [0.00, 0.033], where Low Aesthetic Value images lead to a larger increase in negative affect ($M = 0.14$, $SD = 0.58$) than did High Aesthetic Value images ($M = 0.03$, $SD = 0.58$).

6.3.3. Analysis 3: affect change between aesthetic value conditions within an environment

To handle the conditions which were not completely crossed in our design (VHA-Nat and VLA-Urb), separate factorial ANOVAs were conducted on each of the three aesthetic value levels for nature and urban images.

6.3.3.1. Nature conditions. A one-way ANOVA with the 3 aesthetic value levels (Very High, High, Low) in the nature condition was performed for change in STAI-Pos and STAI-Neg. Results of this ANOVA for STAI-Pos yielded a significant effect of aesthetic value level, $F(2, 300) = 3.41$, $p = 0.034$, $\eta_p^2 = 0.022$, 95% CI [0.0, 0.061]. The partial eta-squared indicates a small effect size. Post hoc comparisons were conducted and family-wise error corrected using Tukey's HSD, which showed a significant difference between the VHA-Nat ($M = 0.42$, $SD = 0.71$) and LA-Nat conditions ($M = 0.17$, $SD = 0.63$, $p = 0.029$, $d = 0.37$), indicating a greater increase in STAI-Pos for those in the VHA-Nat condition relative to those in the LA-Nat condition. There were no differences between VHA-Nat and HA-Nat ($M = 0.25$, $SD = 0.70$, $ps = 0.20$) or between HA-Nat and LA-Nat ($ps = 0.66$).

Results of this analysis for STAI-Neg showed a trending effect of aesthetic value, $F(2,300) = 3.27$, $p = 0.061$, $\eta_p^2 = 0.019$, 95% CI [0.0, 0.055]. The partial eta-squared indicates a small effect size. Post hoc comparisons showed a trending difference between VHA-Nat ($M = -0.06$, $SD = 0.49$) and LA-Nat ($M = 0.13$, $SD = 0.59$, $p = 0.053$, $d = 0.34$). This difference indicates a greater reduction in STAI-Neg for participants in the VHA-Nat condition relative to LA-Nat (which led to an increase in negative affect). No significant difference was found between VHA-Nat and HA-Nat ($M = 0.07$, $SD = 0.49$, $ps = 0.27$) or between HA-Nat and LA-Nat ($ps = 0.71$).

6.3.3.2. Urban conditions. For the urban images, a one-way ANOVA with the 3 aesthetic values (High, Low, Very Low) was also performed for change in STAI-Pos and STAI-Neg. Results of the first ANOVA did not yield a significant effect of aesthetic value level on STAI-Pos. Results of this analysis for STAI-Neg did show a significant effect of aesthetic value, $F(2,301) = 3.06$, $p = 0.04$, $\eta_p^2 = 0.02$, 95% CI [0.0, 0.057]. The partial eta-squared indicates a small effect size. Post hoc comparisons yielded a significant difference between VLA-Urb ($M = 0.18$, $SD = 0.59$) and HA-Urb ($M = -0.01$, $SD = 0.58$, $p = 0.046$, $d = 0.34$), but not between VLA-Urb and LA-Urb ($M = 0.13$, $SD = 0.53$, $ps = 0.81$) or between LA-Urb and HA-Urb ($ps = 0.19$). This indicates a larger increase in negative affect for those in the VLA-Urb condition relative to the HA-Urb condition.

6.4. Discussion

The results of Study 6 generally replicated those found in Study 2, where environment type did not have an influence on affect change in the completely crossed conditions, and aesthetic value was predictive of affect change within an environment. Additionally, in Study 6 a trending effect of aesthetic value was found in the completely crossed conditions, which was not evident in Study 2. As such, it does not appear that priming participants to think about how much they like or dislike the images accounts for the previously reported results.

7. General discussion

Nature interactions reliably elicit positive changes in affect (MacKerron & Mourato, 2013; McMahan & Estes, 2015), however, the underlying mechanism remains unknown. Though many researchers have demonstrated the robust impact of nature on emotions, much of this previous work has not controlled for preference when examining

Table 8
Study 6 STAI Results: Change in affective state relative to baseline.

		STAI-Positive			STAI-Negative		
Condition	n	t-statistic	p-value	Cohen's d	t-statistic	p-value	Cohen's d
VHA-Nat	102	5.9**	<0.001	0.58	−1.16	0.25	0.11
HA-Nat	102	3.62**	<0.001	0.36	1.13	0.26	0.11
HA-Urb	101	2.50*	0.014	0.25	−0.17	0.86	0.02
LA-Nat	99	2.65*	0.009	0.27	2.14*	0.035	0.22
LA-Urb	99	3.28**	0.001	0.33	2.26*	0.026	0.23
VLA-Urb	104	0.65	0.52	0.06	3.37**	0.001	0.33

Results of 1-sample t-tests comparing STAI positive and negative change to zero in each of the 6 conditions. **Significant p-value with Bonferroni family-wise multiple comparisons correction (α = 0.008) *Significant p-value uncorrected (α = 0.05).

mood effects. The goal of the present project was to clarify whether there is something unique about the affective benefits of nature stimuli over and above individuals' preference for these stimuli. Across several studies, consistent evidence for a preference-based account of affect change was found. That is, in this work, nature seems to have a positive effect on emotional state because it is highly preferred. If nature was sufficiently low on aesthetics or compared to an equally preferred urban image, we failed to find evidence for an additional benefit of nature on affective state. Additionally, even nature very high on aesthetics did not elicit larger emotional responses than other equally preferred stimuli, suggesting there was not an additional benefit to affective state of viewing natural scenery per se.

Study 2 demonstrated that, once equated on preference, there were no significant differences in affect change between nature and urban environments. However, aesthetic value, as measured by participants' own ratings or by pre-established conditions within an environment type, did predict the extent to which participants' affective state improved post-picture viewing. In Study 2, the largest condition-level effects were found for images in the most extreme aesthetic value conditions (VHA-Nat and VLA-Urb), which were not completely crossed with environment type.

To overcome this, Study 4 used very high aesthetic value images in categories other than natural environments, and test whether the improved effect on affect after VHA-Nat in Study 2 was due to the high aesthetic value or to the environment category nature itself. For these purposes, urban scenes and animal images with very high preference ratings were utilized. When comparing change in affect before and after image viewing, very high aesthetic nature did not have a larger effect than the animal or urban images. This finding was unaffected by whether participants had been induced to a negative mood state at baseline. Further, the results of multiple regression analyses, which examined both participants' average preference ratings and image category, showed that while rating the images as more highly preferred was significantly predictive of affect change, image category did not have a significant effect.

Study 5 was conducted to address the possibility that the preference measure used in Study 1 and 2 (affinity) captured something unique about nature above and beyond aesthetic preferences in our stimuli. This idea was not supported by the data, as explicit ratings of beauty were almost perfectly correlated with affinity ratings across all images. Lastly, in Study 6 participants were asked to rate the images on naturalness instead of preference to examine whether affect change was affected by the type of judgment being made during the image rating task. These results generally replicated the effects of Study 2, suggesting that priming preference could not account for the previous effects.

The focus of the present research was to address whether there is something unique about natural environments that can lead to changes in affect even when preference is taken into account. Interestingly, the results of this research do not support that viewing nature scenes has an acute effect on affect that can be attributable to something beyond preference. However, it is important to note that the current study does not shed light on why, in general, natural environments are so highly preferred to begin with.

Decades of research have spawned theoretical accounts of the origins of nature preferences. These include evolutionary theories such as Biophilia (Kellert & Wilson, 1995), and Stress Reduction Theory (Ulrich et al., 1991), which propose that because our evolutionary history took place in predominantly natural environments, humans therefore feel an innate affinity towards nature. Other theories propose that the ease of processing visual features often found in natural environments (e.g., fractalness) causes nature to be preferred and causes a positive affective response (Perceptual Fluency Account; Joye et al., 2016). Still others propose that we prefer natural environments due to their potential restorative value (Hartig & Staats, 2006).

Though the current research emphasizes preference as the key ingredient of nature-related affect change, it does not suggest that the nature category is meaningless. Indeed, the robust preference for natural environments might be a vital part of why nature is viewed as unique. From the image preference ratings obtained in Study 1, it was challenging to find urban environments that were as preferred as the high aesthetic nature images (HA-Nat) to be used in Study 2, and was not possible to find urban images to match the very high aesthetic (VHA) nature stimuli. Similarly, in Study 3, only two (animals, urban scenes) of four other categories examined yielded sufficiently overlapping preference distributions to use in Study 4. Additionally, it is noteworthy that although the content of the animal images are quite different from nature scenery, and do not comprise natural spaces, this category is indeed composed of natural stimuli. It is likely that, because of this lack of preference overlap, many studies examining the effects of nature interventions have used images or videos that were not similarly preferred. For example, in the 2008 study conducted by Berman and colleagues (Berman et al., 2008), the average preference rating for the nature images used (on a 1–7 Likert scale) was approximately 5.5, whereas for urban images it was approximately 2.8. Therefore, it is worth noting that though the current data suggest that the affective benefits are only due to preference and not due to anything unique to nature scenes, in some sense, nature is a "special" kind of stimulus due to it being so overwhelmingly preferred (at least among adults, please see Meidenbauer et al., 2019 for research examining environmental preferences in children).

The primacy of aesthetics in nature-elicited affect change has a number of notable implications. Perhaps one of the most important relates to the ongoing debate of the role of nature preferences in cognitive restoration. In particular, Stress Reduction Theory (Ulrich et al., 1991) posits that the cognitive benefits of nature interactions occur due to changes in affective state and reductions in stress. The Stress Reduction Theory framework would therefore predict that if one experienced natural and urban environments which were equally preferred, superior changes in cognitive performance would not be expected after nature exposure. In comparison, Attention Restoration Theory (Kaplan, 1995; Kaplan & Berman, 2010), does not assume that restoration relates to affect in any way. Attention restoration theorists instead focus on features of natural environments which restore directed attention resources while softly capturing involuntary attention, and do not propose that the extent to which a natural environment is preferred matters. Support for this comes from recent evidence that affective and cognitive benefits of nature are dissociable (Stenfors et al., 2019). Though this study demonstrates the difficulty in finding preference-equated environments that match nature preferences, a strong test of whether preference plays a role in the cognitive benefits of nature would be to compare objective performance on cognitive tasks before and after exposure to preference-equated nature and urban images, videos, or walks.

Another implication relates to the use of biophilic design in architecture and urban planning (Joye, 2007). Though the idea of designing buildings to contain nature-like features is not new (Alexander, 2002; Kellert, 2012; Salingaros, 1998), recent research has generated compelling evidence for the overlap between architectural aesthetics and naturalness (Coburn et al., 2019). Broadly speaking, there are many visual features common in natural environments which are also highly aesthetically preferred, such as fractalness or recursive complexity (Van den Berg et al., 2016), density of curved edges (Berman et al., 2014), or color-related properties such as blue-green hue and high saturation diversity (Kardan et al., 2015). The results of this study would suggest that, if one goal of biophilic architecture is to promote positive affective responses, design ought to prioritize inclusion of natural features which provide the most aesthetic value over those which may appear natural but not highly predictive of beauty, such as visual disorder (Kotabe et al., 2017). Furthermore, from an urban planning perspective, this research suggests one clear way to improve city residents' affective well-being is through the incorporation of aesthetically pleasing urban green infrastructure (UGI). Implementing beautiful UGI would be a more feasible way to improve the aesthetic value of currently developed spaces. In addition, nature exposure is associated with improved attentional resources (Berman et al., 2008; Schertz & Berman, 2019), improved mental health (Bratman et al., 2019), positive thinking (Schertz et al., 2018, 2020, Schwartz et al., 2019), reduced crime ((Kuo & Sullivan, 2001a, 2001b; Schertz et al., 2019) and greater neighborhood social cohesion (de Vries et al., 2013; Kuo et al., 1998). Therefore, it is likely that the benefits obtained from this urban green infrastructure would not be limited to residents' emotional functioning.

This study contains a few notable limitations. Though changes in affect have been documented across both real and simulated nature interventions (McMahan & Estes, 2015), these data do not directly speak to whether the results would be different after real life environmental experiences. Given the difficulty in finding urban images that were sufficiently preferred to be able to perform this research, conducting a similar study in preference-equated real environments would likely be very challenging, if not altogether impossible. However, we cannot rule out the possibility that there are qualities of natural environments which contribute to affect change that cannot be captured in images (including additional sensory stimuli), or that the results would not be different with longer term exposures. Additionally, as these studies were conducted via Mturk, a within-subjects design with many sessions was not logistically feasible but might have provided some insight into the variation in preference ratings across studies. Relatedly, though we measured state affect in a multiple ways, there are certainly other elements of the affective benefits of nature that are not examined in the current study which may or may not show different effects, such as awe and compassion (Joye & Bolderdijk, 2014; Swami et al., 2019). It is an empirical question that is certainly worth pursuing.

Another limitation is that we have focused this work primarily on nature's benefits rather than examining the detriments related to urban environments. There is some evidence for a preference effect here as well. In Study 2 the very low aesthetic value urban images were the only category to induce negative affective responses, and in Study 4 very highly preferred urban images elicited a positive affective response. However, Study 2 did not have an equally 'un-preferred' nature condition, so this is still an open question. Lastly, though Study 5 demonstrated near perfect correlations of preference (affinity) and aesthetics (beauty), it is possible that ratings of other scene elements not tested here may differentially predict preference ratings between scene types. Though there is evidence that the visual features associated with preference for urban and nature scenes are very similar (Coburn et al., 2019; Kardan et al., 2015), there may be other ways of evaluating or judging these images (e.g., restorativeness, pleasantness) that may or may not show divergent results. This is a question that deserves future investigation.

In summary, the present research suggests that there is nothing unique about nature beyond preference when it comes to improving affective state and that viewing anything that a person prefers will have a positive effect. Yet it remains important to emphasize the difficulty in finding stimuli that were as highly preferred as nature scenes. Thus, while there may not be anything unique about nature for affect change above and beyond aesthetics, the observation that natural environments, as well as scenes containing nature-related stimuli, are preferred remains a significant one. Overall, the results of this research contain not only important implications for the research of other environmental psychologists, but also provide insights which may be useful in domains such as architecture, urban design, and nature-based clinical interventions to improve the well-being of residents.

Author Note

Per OSF guidelines, the authors of this study have reported all measures, conditions, sample size justification, and data exclusions in the main manuscript and supplementary materials.

Data & Code Availability Statement

Data, experiment materials, and pre-registrations are all publicly available at: https://osf.io/ehtk9/

CRediT authorship contribution statement

Kimberly L. Meidenbauer: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing - original draft, Visualization. **Cecilia U.D. Stenfors:** Conceptualization, Methodology, Investigation, Formal analysis, Writing - review & editing. **Gregory N. Bratman:** Conceptualization, Writing - review & editing, Supervision. **James J. Gross:** Conceptualization, Writing - review & editing, Supervision. **Kathryn E. Schertz:** Formal analysis, Writing - review & editing. **Kyoung Whan Choe:** Software, Writing - review & editing. **Marc G. Berman:** Conceptualization, Methodology, Investigation, Formal analysis, Writing - review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported in part by grants from the TKF Foundation, the John Templeton Foundation (University of Chicago Center for Practical Wisdom and the Virtue, Happiness, and Meaning of Life Scholars Group), and NSF Grant BCS-1632445 to MGB. Partial support comes from the international research fellowship grant awarded to CS from the Swedish Research Council (reference no 2015-00190). KES is supported by NSF grant DGE-1746045. GB is grateful for support from the Doug Walker Endowed Faculty Professorship, Craig McKibben and Sarah Merner, John Miller, and Gretchen C. Daily.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jenvp.2020.101498.

References

Alexander, C. (2002). *The nature of order: The phenomenon of life*. Taylor & Francis.
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.
Barrett, L. F. (2006). Solving the Emotion Paradox: Categorization and the Experience of Emotion. *Personality and Social Psychology Review, 10*(1), 20–46. https://doi.org/10.1207/s15327957pspr1001_2
Barrett, L. F., Mesquita, B., Ochsner, K. N., & Gross, J. J. (2007). The Experience of Emotion. *Annual Review of Psychology, 58*, 373–403. https://doi.org/10.1146/annurev.psych.58.110405.085709
Berman, M. G., Hout, M. C., Kardan, O., Hunter, M. R., Yourganov, G., Henderson, J. M., … Jonides, J. (2014). The perception of naturalness correlates with low-level visual features of environmental scenes. *PloS One, 9*(12), e114572.
Berman, M. G., Jonides, J., & Kaplan, S. (2008). The cognitive benefits of interacting with nature. *Psychological Science, 19*.
Beute, F., & de Kort, Y. A. W. (2014). Natural resistance: Exposure to nature and self-regulation, mood, and physiology after ego-depletion. *Journal of Environmental Psychology, 40*, 167–178.
Beute, F., & de Kort, Y. A. W. (2018). The natural context of wellbeing: Ecological momentary assessment of the influence of nature and daylight on affect and stress for individuals with depression levels varying from none to clinical. *Health & Place, 49*, 7–18.
Bowler, D. E., Buyung-Ali, L. M., Knight, T. M., & Pullin, A. S. (2010). A systematic review of evidence for the added benefits to health of exposure to natural environments. *BMC Public Health, 10*, 456.
Bratman, G. N., Anderson, C. B., Berman, M. G., Cochran, B., de Vries, S., Flanders, J., … Daily, G. C. (2019). Nature and mental health: An ecosystem service perspective. *Science Advances, 5*(7), Article eaax0903.
Bratman, G. N., Daily, G. C., Levy, B. J., & Gross, J. J. (2015). The benefits of nature experience: Improved affect and cognition. *Landscape and Urban Planning, 138*, 41–50.

Browning, M. H. E. M., Mimnaugh, K. J., van Riper, C. J., Laurent, H. K., & LaValle, S. M. (2020). Can simulated nature support mental health? *Frontiers in Psychology, 10*, 2667.

Coburn, A., Kardan, O., Kotabe, H., Steinberg, J., Hout, M. C., Robbins, A., … Berman, M. G. (2019). Psychological responses to natural patterns in architecture. *Journal of Environmental Psychology, 62*, 133–145.

Fox, J., & Weisberg, S. (2019). *An {R} companion to applied regression* (3rd ed.). Thousand Oaks CA: Sage. URL: https://socialsciences.mcmaster.ca/jfox/Books/Companion/.

Fuegen, K., & Breitenbecher, K. H. (2018). Walking and being outdoors in nature increase positive affect and energy. *Ecopsychology, 10*(1), 14–25.

Glasgow, T. E., Le, H. T. K., Geller, E. S., Fan, Y., & Hankey, S. (2019). How transport modes, the built and natural environments, and activities influence mood: A GPS smartphone app study. *Journal of Environmental Psychology, 66*, Article 101345.

Hartig, T., Böök, A., Garvill, J., Olsson, T., & Gärling, T. (1996). Environmental influences on psychological restoration. *Scandinavian Journal of Psychology, 37*(4), 378–393.

Hartig, T., Evans, G. W., Jamner, L. D., Davis, D. S., & Garling, T. (2003). Tracking restoration in natural and urban field settings. *Journal of Environmental Psychology, 23*, 109–123.

Hartig, T., & Staats, H. (2006). The need for psychological restoration as a determinant of environmental preferences. *Journal of Environmental Psychology, 26*(3), 215–226.

Hawking, S. (1988). *A brief history of time*. New York: Bantam Books.

Johansson, M., Hartig, T., & Staats, H. (2011). Psychological benefits of walking: Moderation by company and outdoor environment. *Applied Psychology. Health and Well-Being, 3*(3), 261–280.

Joye, Y. (2007). Architectural lessons from environmental psychology: The case of biophilic architecture. *Review of General Psychology: Journal of Division 1, of the American Psychological Association, 11*(4), 305–328.

Joye, Y., & Bolderdijk, J. W. (2014). An exploratory study into the effects of extraordinary nature on emotions, mood, and prosociality. *Frontiers in Psychology, 5*, 1577.

Joye, Y., Steg, L., Ünal, A. B., & Pals, R. (2016). When complex is easy on the mind: Internal repetition of visual information in complex objects is a source of perceptual fluency. *Journal of Experimental Psychology. Human Perception and Performance, 42*(1), 103–114.

Kaplan, S. (1995). The restorative benefits of nature: Toward an integrative framework. *Journal of Environmental Psychology, 15*(3), 169–182.

Kaplan, S., & Berman, M. G. (2010). Directed attention as a common resource for executive functioning and self-regulation. Perspectives on psychological science. *A Journal of the Association for Psychological Science, 5*(1), 43–57.

Kaplan, R., & Herbert, E. J. (1987). Cultural and sub-cultural comparisons in preferences for natural settings. *Landscape and Urban Planning, 14*, 281–293.

Kaplan, R., & Kaplan, S. (1989). *The experience of nature: A psychological perspective*. CUP Archive.

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

Kardan, O., Demiralp, E., Hout, M. C., Hunter, M. R., Karimi, H., Hanayik, T., … 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*. https://doi.org/10.3389/fpsyg.2015.00471

Karmanov, D., & Hamel, R. (2008). Assessing the restorative potential of contemporary urban environment(s): Beyond the nature versus urban dichotomy. *Landscape and Urban Planning, 86*(2), 115–125.

Kellert, S. R. (2012). *Building for life: Designing and understanding the human-nature connection*. Island Press.

Kellert, S. R., & Wilson, E. O. (1995). *The Biophilia hypothesis*. Island Press.

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.

Kuo, F. E., & Sullivan, W. C. (2001a). Aggression and violence in the inner city: Effects of environment via mental fatigue. *Environment and Behavior, 33*(4), 543–571.

Kuo, F. E., & Sullivan, W. C. (2001b). Environment and crime in the inner city: Does vegetation reduce crime? *Environment and Behavior, 33*(3), 343–367.

Kuo, F. E., Sullivan, W. C., Coley, R. L., & Brunson, L. (1998). Fertile ground for community: Inner-city neighborhood common spaces. *American Journal of Community Psychology, 26*(6), 823–851.

Kuo, F. E. (2015). How might contact with nature promote human health? Promising mechanisms and a possible central pathway. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2015.01093

Lawrence, M. A. (2016). z: Easy analysis and visualization of factorial experiments. R package version 4.4-0. *e*. https://CRAN.R-project.org/package ez.

Lee, J., Park, B.-J., Tsunetsugu, Y., Kagawa, T., & Miyazaki, Y. (2009). Restorative effects of viewing real forest landscapes, based on a comparison with urban landscapes. *Scandinavian Journal of Forest Research/Issued Bimonthly by the Nordic Forest Research Cooperation Committee, 24*(3), 227–234.

Lindquist, K. A., Satpute, A. B., Wager, T. D., Weber, J., & Barrett, L. F. (2016). The brain basis of positive and negative affect: Evidence from a meta-analysis of the human neuroimaging literature. *Cerebral Cortex, 26*(5), 1910–1922.

Lüdecke, D. (2019). *sjstats: Statistical functions for regression models (version 0.17.6)*. https://doi.org/10.1016/s1473-3099(20)30313-3, 10.5281/zenodo.1284472 (URL: 10.5281/zenodo.1284472), <URL: https://CRAN.R-project.org/package sjstats.

MacKerron, G., & Mourato, S. (2013). Happiness is greater in natural environments. *Global Environmental Change, 23*, 992–1000.

Marteau, T. M., & Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology/the British Psychological Society, 31*(Pt 3), 301–306.

Mayer, F. S., Frantz, C. M., Bruehlman-Senecal, E., & Dolliver, K. (2009). Why is nature beneficial?: The role of connectedness to nature. *Environment and Behavior, 41*(5), 607–643.

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, 10*(6), 507–519.

Meidenbauer, K. L., Stenfors, C. U. D., Young, J., Layden, E. A., Schertz, K. E., Kardan, O., … Berman, M. G. (2019). The gradual development of the preference for natural environments. *Journal of Environmental Psychology, 65*, 101328.

R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Ryan, R. M., Weinstein, N., Bernstein, J., Brown, K. W., Mistretta, L., & Gagné, M. (2010). Vitalizing effects of being outdoors and in nature. *Journal of Environmental Psychology, 30*(2), 159–168.

Salingaros, N. A. (1998). A scientific basis for creating architectural forms. *Journal of Architectural and Planning Research, 15*(4), 283–294.

Schertz, K. E., & Berman, M. G. (2019). Understanding nature and its cognitive benefits. *Current Directions in Psychological Science, 28*(5), 496–502.

Schertz, K. E., Kardan, O., & Berman, M. G. (2020). Visual features influence thought content in the absence of overt semantic information. *Attention, Perception & Psychophysics*. https://doi.org/10.3758/s13414-020-02121-z.

Schertz, K. E., Sachdeva, S., Kardan, O., Kotabe, H. P., Wolf, K. L., & Berman, M. G. (2018). A thought in the park: The influence of naturalness and low-level visual features on expressed thoughts. *Cognition, 174*, 82–93.

Schertz, K. E., Saxon, J., Cardenas-Iniguez, C., Bettencourt, L., Ding, Y., Hoffmann, H., & Berman, M. G. (2019). *Neighborhood street activity and greenspace usage uniquely contribute to predicting crime*. https://doi.org/10.31234/osf.io/qzf7a

Schwartz, A. J., Dodds, P. S., O'Neil-Dunne, J. P. M., Danforth, C. M., & Ricketts, T. H. (2019). Visitors to urban greenspace have higher sentiment and lower negativity on Twitter. *People and Nature, 1*(4), 476–485.

Staats, H., Kieviet, A., & Hartig, T. (2003). Where to recover from attentional fatigue: An expectancy-value analysis of environmental preference. *Journal of Environmental Psychology, 23*(2), 147–157.

Stanley, D. (2018). paTables: Create American psychological association (APA) style tables. R package version 2.0.5. *a*. https://CRAN.R-project.org/package apaTables.

Stenfors, C. U. D., Van Hedger, S. C., Schertz, K. E., Meyer, F. A. C., Smith, K. E. L., Norman, G. J., … Berman, M. G. (2019). Positive effects of nature on cognitive performance across multiple experiments: Test order but not affect modulates the cognitive effects. *Frontiers in Psychology, 10*, 1413.

Swami, V., Barron, D., Hari, R., Grover, S., Smith, L., & Furnham, A. (2019). The nature of positive body image: Examining associations between nature exposure, self-compassion, functionality appreciation, and body appreciation. *Ecopsychology, 11*(4), 243–253.

Tsunetsugu, Y., Lee, J., Park, B.-J., Tyrväinen, L., Kagawa, T., & Miyazaki, Y. (2013). Physiological and psychological effects of viewing urban forest landscapes assessed by multiple measurements. *Landscape and Urban Planning, 113*, 90–93.

Ulrich, R. S. (1979). Visual landscapes and psychological well-being. *Landscape Research, 4*(1), 17–23.

Ulrich, R. S. (1983). Aesthetic and affective response to natural environment. In I. Altman, & J. F. Wohlwill (Eds.), *Behavior and the natural environment* (pp. 85–125). Boston, MA: Springer US.

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*(3), 201–230.

Valtchanov, D., Barton, K. R., & Ellard, C. (2010). Restorative effects of virtual nature settings. *Cyberpsychology, Behavior, and Social Networking, 13*(5), 503–512.

Valtchanov, D., & Ellard, C. G. (2015). Cognitive and affective responses to natural scenes: Effects of low level visual properties on preference, cognitive load and eye-movements. *Journal of Environmental Psychology, 43*, 184–195.

Van den Berg, A. E., Joye, Y., & Koole, S. L. (2016). Why viewing nature is more fascinating and restorative than viewing buildings: A closer look at perceived complexity. *Urban Forestry and Urban Greening, 20*, 397–401.

van Hedger, S. C., Nusbaum, H. C., Heald, S. L., Huang, A., Kotabe, H. P., & Berman, M. G. (2019). The Aesthetic Preference for Nature Sounds Depends on Sound Object Recognition. *Cognitive Science*. https://doi.org/10.1111/cogs.12734

Verheyen, C., & Göritz, A. S. (2009). Plain texts as an online mood-induction procedure. *Social Psychology, 40*(1), 6–15.

de Vries, S., van Dillen, S. M. E., Groenewegen, P. P., & Spreeuwenberg, P. (2013). Streetscape greenery and health: Stress, social cohesion and physical activity as mediators. *Social Science & Medicine, 94*, 26–33.

White, M. P., Alcock, I., Wheeler, B. W., & Depledge, M. H. (2013). Would you be happier living in a greener urban area? A fixed-effects analysis of panel data. *Psychological Science, 24*(6), 920–928.

White, M., Smith, A., Humphryes, K., Pahl, S., Snelling, D., & Depledge, M. (2010). Blue space: The importance of water for preference, affect, and restorativeness ratings of natural and built scenes. *Journal of Environmental Psychology, 30*(4), 482–493.

Xiao, J., Hays, J., Ehinger, K. A., Oliva, A., & Torralba, A. (2010). *SUN database: Large-scale scene recognition from abbey to zoo* (pp. 3485–3492). 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*(Issue 2), 151–175. https://doi.org/10.1037/0003-066x.35.2.151

Zeileis, A., & Hothorn, T. (2002). Diagnostic checking in regression relationships. *R News, 2*(3), 7–10. https://CRAN.R-project.org/doc/Rnews/.