Testing the Biophilia theory: Automatic approach tendencies towards nature

Tanja Schiebel[2], Jürgen Gallinat[2], & Simone Kühn[1,2]

[1] Max Planck Institute for Human Development,
Lise Meiter Group for Environmental Neuroscience,
Lentzeallee 94, 14195 Berlin, Germany

[2] University Clinic Hamburg-Eppendorf,
Clinic and Policlinic for Psychiatry and Psychotherapy,
Martinistraße 52, 20246 Hamburg, Germany

Corresponding author: Simone Kühn

E-mail: kuehn@mpib-berlin.mpg.de

keywords: urbanicity, implicit association test, approach avoidance task, dot probe task, urban, rural, schizophrenia

Abstract

It seems as if modern urban lifestyle disconnects people from nature, this may be associated with adverse health effects. In line with this notion it has been consistently shown that psychiatric diagnoses are more frequent in urban compared to rural regions. Most of the studies addressing potential causal mechanisms of this urban-rural difference focus on detrimental aspects of city living. In contrast, biophilia theory has posited an automatic, potentially deep-rooted need for contact with nature. Acting against this proposed tendency to seek contact to natural environments may affect mental health. As scientific evidence for this psycho-evolutionary biophilia theory is lacking by now, we utilized implicit test strategies developed to assess automatic associations between mental representations and action tendencies to put this theory to test. In an online study ($N = 109$), we administered three reaction time paradigms: the dot probe task (DPT), the implicit association test (IAT) and the approach avoidance task (AAT). All tasks reveal a tendency to approach nature and avoid cities (DPT: $F(1,105)=11.15$, $p=.001$, $\eta^2=.096$; IAT: $F(1,107)=17.10$, $p=7.068E-5$, $\eta^2=.138$; AAT: $F(1,103)=4.36$, $p=.039$, $\eta^2=.041$). Interestingly, the results of the AAT, the only test that allows this differentiation, suggest that the tendency to approach nature seems to play a more important role than the avoidance of built environments. The present findings provide clear evidence in support of biophilia theory and can therefore inspire and foster further studies investigating whether acting against an automatic and potentially deep-rooted need for contact with nature by living in cities e.g. may contribute more prominently to the emergence of mental health problems than (or at least in addition to) environmental or societal stressors individuals are exposed to in cities.

275 words

1. *Introduction*

Urbanization is steadily increasing, with more than half of the world's population living in urban areas today and prospectively 68% in the year 2050. Since urban settings are a relatively new phenomenon in phylogenesis, their long-term impact on human well-being and mental health cannot be fully estimated yet. Generally, it seems as if urban inhabitants enjoy better health than their rural counterparts (Dye, 2008), in particular considering physical ailments such as obesity, diabetes and premature morbidity (Eberhardt & Pamuk, 2004; Wagner & Brath, 2012). However, mental health seems to be a striking exception. It has been consistently shown that psychiatric diagnoses such as mood and anxiety disorders as well as schizophrenia are more frequent in urban compared to rural areas (Jaap Peen & Dekker, 2004; J. Peen, Schoevers, Beekman, & Dekker, 2010). Most of the time this urban-rural difference has been explained by a higher prevalence of stress in the city (Abbott, 2012; Kennedy & Adolphs, 2011). Albeit, the specific factors causing stress and therewith the urban increase in psychiatric diseases are still unknown. Most of the present literature focuses on social stressors such as decreases in social support, increases in social isolation (Holz, Tost, & Meyer-Lindenberg, 2020; Tost, Champagne, & Meyer-Lindenberg, 2015) or environmental stressors such as air pollution (Khan, et al., 2019; Newbury, et al., 2019) which are more prominent in cities.

Although the focus on detrimental aspects of city living is predominant in the search for factors causing higher prevalence of psychiatric disease in urban contexts, a potential role of the absence of nature also has been discussed. First evidence revealed that exposure to green spaces during childhood may reduce the risk of later psychiatric disorder (Engemann, et al., 2019). And presence of green and blue spaces (namely water), around the home address of individuals has been shown to be negatively associated with the occurrence of mental disorders (de Vries, et al., 2016). In accordance to this, nature interventions have been shown to elicit positive effects on mental health (Hubbard, et al., 2020; Tillmann, Tobin, Avison, & Gilliland, 2018; Trostrup, Christiansen, Stolen, Nielsen, & Stelter, 2019). Moreover, longitudinal data from Britain revealed that individuals who moved to greener areas showed better mental health three years post movement (Alcock, White, Wheeler, Fleming, & Depledge, 2014). In line with this, East Asian countries have a long tradition in research on and exposure to nature as a facilitator of health. In particular the Japanese practice of "Shinrin-yoku", which translates to "forest bathing", is considered as a remedy for urban stress (Park, Tsunetsugu, Kasetani, Kagawa, & Miyazaki, 2010).

In terms of theoretical background, most of the studies on the positive effects of nature exposure draw onto psycho-evolutionary theories such as the Biophilia theory (Wilson, 1984) (or likewise the Attention Restoration (Berman, Jonides, & Kaplan, 2008; R. Kaplan & Kaplan, 1989) or Stress Reduction Theory (Ulrich, et al., 1991)), all positing that humans have an innate tendency to seek connection with nature, which is seen as the product of biological evolution. A common criticism of evolutionary theories is that they make predictions that are difficult to falsify. However, for the Biophilia theory, a study design and hypothesis that pushes itself to the fore is, to explore whether humans have an implicit tendency to approach nature (or to avoid cities). Social psychology has developed and applied implicit test strategies to assess automatic associations between mental representations and action tendencies. These paradigms comprise reaction time tasks where participants respond to stimuli presented on a computer screen by means of button presses or similar reactions, e.g. movement of a joystick or a computer mouse. Tests such as the dot probe task (DPT, *Figure 1 a*) (MacLeod, Mathews, & Tata, 1986), the implicit association test (IAT, *Figure 1 b*) (Anthony G. Greenwald, McGhee, & Schwartz, 1998), and the approach avoidance task (AAT, *Figure 1 c*) (Rinck & Becker, 2007), have been used to assess stereotypes, attitudes and perceptions, but are also utilized in clinical contexts to test for cognitive biases in individuals suffering from addiction disorders, phobias or suicidality (Nock, et al., 2010; C. E. Wiers, et al., 2013). We set out to utilize these implicit tests to investigate whether humans have an automatic tendency to approach nature (and/or to avoid cities respectively) as posited by the Biophilia theory.

In case this biophilic tendency proves true this may in the future add to the understanding of the preponderance of psychiatric diseases in urban contexts. Acting against the assumed automatic and potentially deep-rooted need for contact with nature may cause stress and contribute to the emergence of mental health problems, rather than or in addition to specific environmental or societal stressors of the city. This notion has already been introduced by the microbiologist Dubos who argued that access to and contact with natural environments was essential to the mental health of populations (Logan, Katzman, & Balanza-Martinez, 2015).

2. Methods

2.1. Participants

109 healthy individuals (see *Table 1* for sample characteristics) took part in the study (sample size was guided by Joye 2013, Study 2), matching the following eligibility criteria: age 18- 75 years, sufficient German language skills, no lifetime diagnosis of a neurological or a severe

psychiatric disease, no acute suicidal thoughts or tendencies, informed consent for participation, owning and using a computer mouse and consent to be reimbursed via money transfer. Participants were recruited via online posts and flyers. The local psychological ethics committee of the University Medical Center Hamburg-Eppendorf, Germany, approved of the study (LPEK-0019).

2.2. Procedure

The experiment was implemented online using Inquisit 5 (www.millisecond.com). Participants were sent the link to the study as well as a participation number. Participants were presented the study information, asked to give their informed consent to participate and to confirm matching the eligibility criteria. In case of participation, they were instructed to answer sociodemographic questions. Afterwards, they completed two questions assessing information on residence: (1) current place of residence (choices: city with over 100.000 inhabitants, town with over 10.000 inhabitants, rural area) and (2) how many years they grew up living in a city, a town and in a rural area until the age of 15 years (Pedersen & Mortensen, 2001). Next, they were asked to rate 40 pictures using a 100-point visual analogue scale ranging from "not at all" to "very much" (German translation: „überhaupt nicht" to „sehr stark") answering the following two questions: (1) "How much does the place in the picture appeal to you?" (German translation: „Wie gut gefällt Ihnen der Ort auf dem Bild?") and (2) "Please rate the aesthetics of the place in the picture." (German translation: „Bitte schätzen Sie die Schönheit/Ästhetik des Ortes auf dem Bild ein."). After the survey part of the experiment, participants completed six experimental paradigms (DPT, IAT, AAT; each twice).

Participants received 12€ for study participation. In total, the experiment lasted for about 75-90 minutes.

2.3. Stimulus material and randomization of the tasks

40 different photographs were used as stimulus material during all experimental paradigms. The pictures were selected from the website "Scenic or not" (http://scenicornot.datasciencelab.co.uk/) where pictures all over Great Britain (originating from http://www.geograph.org.uk/) can be rated with regard to their aesthetics ("Scenic or not?" on a 10 point Likert scale from 0 = "not scenic" to 10 = "very scenic"), while the collected data can be openly accessed. We chose ten pictures each of the following categories and ratings: (1) B1: built environment, low scenic rating ~1; (2) B5: built environment, medium scenic rating ~5; (3) N5: natural environment, medium scenic rating ~5 (4) N9: natural environment, high scenic rating ~9 (descriptive statistics of the online ratings in *Table 2*) (*Figure 1d*). All

photographs depicted unthreatening scenes (Joye, Pals, Steg, & Evans, 2013). To obtain a stimulus set which differed in terms of picture content and aesthetics, we combined B1 and N9, whereas B5 and N5 were merged as a second stimulus set which only differed concerning the respective picture content but not in terms of aesthetics. To make sure the chosen B5 and N5 pictures did not differ in terms of rated aesthetics on the data collected by the website, a paired t test was performed, $t(13.91)=1.16$, $p=.265$. Twenty additional pictures served as practice stimuli. The online links to the pictures used are provided in the *Supplementary Material (S6).* Participants performed each of three paradigms twice: once with each picture set (B1N9 or B5N5). Half of the participants started with the B1N9 pictures, the other half started with the B5N5 pictures. Tasks were presented in the same order for both picture sets within each participant, while task order was counterbalanced across participants (three tasks: A, B, C; two picture sets: 1, 2; e.g. participant X: B1,C1,A1,B2,C2,A2; not: B1,B2,C1,C2,A1,A2; participant Y: C2,B2,A2,C1,B1,A1).

2.4.Dot Probe Task (DPT)

After participants viewed a fixation cross for 500ms in the center of the screen, two pictures (one built and one natural) were displayed on the left and right side of the screen for 500ms. One of the pictures was followed by the presentation of an “X” (= probe). Participants were instructed to indicate the position of the probe (left or right) as quickly as possible by pressing “E” or “I” with their index fingers (*Figure 1a*). The probe was presented for a maximum of 1000ms. As soon as a valid key was pressed, the probe vanished. If the response was incorrect or no response was made during 1000ms, a red error sign (“Fehler”) was displayed for 400ms. Twenty pictures (10x built and 10x natural of the picture set B1N9 or B5N5) were sorted into 10 fixed pairs, which were used as stimuli. Each pair was presented 16 times, resulting in 160 trials. The probe appeared equally often in the position of the built and the natural picture as well as on the left and on the right side. The order of trials was fully randomized. At the start participants practiced the task during 10 trials.

2.5.Implicit Association Test (IAT)

Participants were instructed to press “E” or “I” on the keyboard as quickly as possible with their index fingers to assign stimuli to categories displayed on the left and right upper corner of the screen (*Figure 1b*). Twenty pictures (10x built and 10x natural of the corresponding picture set B1N9 or B5N5) and 10 words (see *Table 3*) served as stimuli. While the pictures had to be assigned to the categories “city” (“Stadt”) or “landscape” (“Land”), words were to be categorized as “approach” (“Annäherung”) or avoidance (“Vermeidung”).

The IAT consisted of 220 trials presented in 7 test blocks. The order of the stimuli within each block was fully randomized. After participants practiced the categorization of the pictures in block 1 (20 trials), they had to categorize the words in block 2 (20 trials). During the next two blocks 3+4 (40 trials each), pictures and words were presented alternately while each key “E” or “I” was associated with two categories (e.g. “E” = “city” and “approach”; “I” = “landscape” and “avoidance”; incompatible condition). Next, only pictures had to be assigned, but the corresponding categories had switched sides on the screen (block 5, 20 trials). The last two blocks 6+7 resembled blocks 3+4 with the difference that the categories belonging to one key were paired vice versa (e.g. “E” = “landscape” and “approach”; “I” = “city” and “avoidance”; compatible condition). The order of the conditions (compatible and incompatible) was counterbalanced across participants.

If participants pressed the wrong key, a red error sign (“Fehler”) was presented for 200ms and the answer had to be corrected. As reaction time (RT), the time between stimulus onset and correct keypress was recorded (built-in error penalty, (A. G. Greenwald, Nosek, & Banaji, 2003)). The inter trial interval was 250ms. The categories were constantly displayed during each test block in the upper corners of the screen.

2.6. Approach Avoidance Task (AAT)

For the AAT, participants were instructed to respond to pictures by pulling the computer mouse towards themselves (approach) or pushing it away from themselves (avoidance) as quickly as possible. The type of reaction (pull/push) was determined by the thickness (thin/thick) of a black frame around the picture (Lawrence, et al., 2015) (*Figure 1c)*. Which frame type required which reaction type was counterbalanced across participants. We used an irrelevant feature version of the AAT (reaction type depends on frame type, not on picture content) to facilitate the measurement of “automatic” tendencies as Wiers and colleagues propose (C. E. Wiers, et al., 2013). The approach and avoidance reactions were visually elucidated: While pulling the mouse towards oneself, the picture size increased, whereas it decreased while pushing the mouse away (zooming effect).

Each of the 20 stimuli (10x built, 10x natural content) was presented four times with each frame type in a fully randomized order, resulting in a total of 160 trials. Consequently, both pictures types (built and natural) had to be pulled and pushed equally often. Participants practiced the task in a block of 20 trials.

At the beginning of each trial, participants had to click on a red “X”, to make sure the cursor was in the center of the screen. Afterwards, the picture was presented. As soon as the mouse cursor reached the lower or upper rim of the screen the picture vanished. The inter-trial interval was 300 ms long. If the mouse was not moved in the right direction, an error sign (“Fehler”, in red color) was displayed for 400 ms. As long as the cursor had not yet reached the wrong rim of the screen, participants were able to correct their movement.

Two different RTs were recorded (Solarz, 1960): The time to initiating the response (initial RT: stimulus onset until start of mouse movement) and the time of response execution (movement RT: start of mouse movement until the cursor reaches the upper or lower rim of the screen).

2.7.Data Analysis

Manipulation check. To make sure the groups of pictorial stimuli (B1, B5, N5, N9) were appropriately selected for our sample, we performed a manipulation check on the picture ratings and checked (1) descriptive statistics and (2) via paired *t* tests, if the ratings between picture groups differed significantly (B1vs.N9, B1vs.B5, N5vs.N9) or were the same (B5 and N5) according to expectations.

DPT. Only correct trials were regarded as valid for the analysis of the DPT (Waechter, Nelson, Wright, Hyatt, & Oakman, 2013). Furthermore, trials with extremely short RTs (< 200 ms) were deleted (van Ens, Schmidt, Campbell, Roefs, & Werthmann, 2019). 95.9 % of the original data remained. As participants were only given the possibility to respond during a time span of 1000 ms, there were no outliers with extreme long RTs. Two participants had less than 65% valid trials in one of the DPTs and had to be excluded from the analysis ($N = 107$) (R. W. Wiers, Eberl, Rinck, Becker, & Lindenmeyer, 2011). We calculated medians for each combination of the factors “congruency” (incongruent vs. congruent) and “picture set” (B1N9 vs. B5N5) (Schoenmakers, Wiers, & Field, 2008) and conducted an ANCOVA considering the covariate “age”.

IAT. The data of the IAT was prepared based on an improved scoring algorithm (D_2) proposed by Greenwald and colleagues (A. G. Greenwald, et al., 2003) with slight changes. Trials with RTs above 10.000 ms and below 400 ms were deleted. As our version of the IAT contained a built-in error penalty, error trials were not excluded. 99.09 % of the data remained valid. No participants had to be excluded from analyses (N = 109). As we wanted to take into account the factor "congruency" in our analysis for a more detailed understanding of the IAT effect, we decided to deviate from the original D_2 procedure. Instead of subtracting the means (incongruent – congruent) and standardizing the differences, we calculated the mean per condition (congruent: mean of block 3 and 5; incongruent: mean of block 7 and 9). This procedure enabled us to perform an ANCOVA with the factors "congruency" (congruent vs. incongruent) and "picture set" (B1N9 vs. B5N5) while considering "age" as a covariate.

AAT. Only trials with correct responses of the AAT were used for further analyses. A correct response was defined as a mouse movement, which started into the right direction and reached the correct rim of the screen without any changes of direction. Furthermore, trials with extremely long RTs were deleted based on visual screenings of the distributions. Cut-Offs were specified liberally (initial RT: > 5000 ms; movement RT: > 2000 ms). 89.1% of the data (for both initial and movement RT) remained in the analyses. In the last step, participants with less than 65% valid trials in one of the AATs were removed from the dataset (R. W. Wiers, et al., 2011). As the data of four participants had to be deleted, AAT analyses were performed with a sample of $N = 105$. To aggregate the single RTs, we calculated medians instead of means, as common in the field, because of their lower sensitivity for outliers (Rinck & Becker, 2007). The medians for all possible combinations of the factors "direction" (push vs. pull), "picture content" (built vs. natural environment) and "picture set" (B1N9 vs. B5N5) served as basis for the calculation of an ANCOVA, which considered "age" as a covariate (Paslakis, Kühn, Grunert, & Erim, 2017). To further examine significant interaction effects, *t* tests were conducted. Two analyses were separately conducted for both types of RTs (initial and movement).

Reliability of reaction time tasks. In order to check if the tasks served as reliable measurement techniques for the bias towards natural/ against built environments, we performed reliability calculations. For each task, the different stimuli (DPT: 10 picture pairs; IAT: 20 pictures and 10 words; AAT: 20 pictures) were regarded as "items" which were used to calculate Cronbach's α. As each stimulus was presented various times during each task, we calculated the average reaction time for each stimulus to get one value per "item". Cronbach's α was calculated separately for each group of stimuli that we expected to produce similar reaction times (DPT and IAT: separately for the four combinations of the factors "picture set" and "congruency"; AAT: separately for all possible combinations of the factors "movement direction", "picture content", and "picture set". Since we compare the stimulus groups separately in our analyses, we chose this procedure to calculate the reliability. However, many studies use bias scores (difference score: incongruent-congruent condition (DPT, IAT) or push-pull reaction times (AAT)) in there analyses, here those scores should be used to assess reliability (Anthony G. Greenwald, et al., 1998; MacLeod, et al., 1986; Rinck & Becker, 2007). In order to report reliability measures comparable to the literature, we also calculated split-half reliabilities using the difference scores as measures. To do so, we further summarized the data calculating a difference score per "item". We randomly assigned the items to two test halves (using the online random generator from "matheretter.de") with the constraint of a balanced design (e.g. equal number of pictures and words in both halves). As trials were randomly presented and reaction times of various presentations of each picture/word were averaged, we considered possible confounding effects addressed (Pronk, Molenaar, Wiers, & Murre, 2021).

Relationship between implicit biases and explicit picture ratings. As former studies have been using explicit measures to assess participants' connection to nature (Whitburn, Linklater, & Abrahamse, 2020), we calculated Pearson correlations to explore the relationship of the implicit biases and explicit measures for the concept of liking with respect to natural environments, indicating approach motivation. We therefore used the picture ratings of the first question "How much does the place in the picture appeal to you?" (from "not at all" to "very much"). We calculated an average rating per picture category B1, B5, N5, N9 per participant and further summarized the ratings by calculating the difference "natural" - "built" which should result in positive values given higher ratings for nature pictures as Biophilia theory posits. For the reaction time tasks we calculated biases (DPT and IAT: incongruent condition – congruent condition; AAT: first step: bias= push-pull RTs, and second step: bias for natural – bias for built pictures) which should also produce positive values while higher values mean stronger biases towards nature. Correlations were calculated separately for both picture sets B1N9 and B5N5 as well as for all 40 pictures, independent from ratings of aesthetic pleasantness.

Data was prepared using R (R) and analyzed using SPSS 24. We decided to restrict our reports to main effects as well as interaction effects which are relevant for our research question. All analyses were based on a significance level of α=.05. In case of multiple testing, Bonferroni correction was used. Apart from "age" we also took "sex" into account as a covariate, but as the results did not show any differences, we refrained from reporting them for the sake of clarity.

3. Results

3.1. Manipulation Check

An overview of the descriptive statistics of the aesthetic ratings for each picture group can be found in *Table 2.* The ratings produced by our sample resemble those of the online sample. While the B1, B5 and N5 pictures were rated higher than expected, absolute ratings for the N9 pictures were slightly lower. However, paired *t* tests (see *Table 4*) show that the expected pattern of differences and parity between picture groups prevails: While the difference in the ratings of the B1-N9, B1-B5 and N5-N9 pictures reached statistical significance, the B5-N5 pictures were rated as equally aesthetic.

3.2. DPT

In a 2x2 ANCOVA with the two factors "congruency" (congruent vs. incongruent) and "picture set" (B1N9 vs. B5N5) we found a significant main effect of "congruency", $F(1,105)=11.15$, $p=.001$, 95% CI [8.05, 11.42], $\eta^2=.096$. This effect reveals an attention bias towards nature as RTs are faster for congruent (probe at the position of previous natural picture) than for incongruent trials (probe at the position of previous built picture) (*Figure 2*). The two-way interaction of "congruency" and "picture set" did not reach significance, $F(1,105)=0.11$, $p=.746$, $\eta^2=.006$. Thus, there is no evidence that the aesthetics of the pictures influences the attention bias. The main effect of "picture set" was not significant, $F(1,105)=0.793$, $p=.375$, 95% CI [-6.04, 3.49], $\eta^2=.007$.

3.3. IAT

A 2x2 ANCOVA with the factors "congruency" (congruent vs. incongruent) and "picture set" (B1N9 vs. B5N5) as well as "age" as covariate revealed a highly significant main effect of congruency, $F(1,107)=17.10$, $p=7.068E\text{-}5$, 95% CI [-129.01, -60.80], $\eta^2=.138$. RTs were faster during the congruent test blocks than during the incongruent test blocks, suggesting an approach bias towards natural and avoidance bias towards built environments. This main effect was extended by the significant two-way interaction of "congruency"x"picture set", $F(1,107)=5.80$, $p=.018$, $\eta^2=.051$, shown in *Figure 3*. This reflects that the IAT effect (RT difference between congruent and incongruent blocks) is higher for B1N9 than for B5N5 pictures. However, an ANCOVA conducted separately for the B5N5-IAT with "age" as covariate likewise shows a significant main effect of congruency, $F(1,107)=5.24$, $p=.024$, 95% CI [-120.29, -45.57]

η^2=.047. This result indicates that the approach bias towards natural and avoidance bias towards built environments measured by the IAT is present in both picture sets. However, it is not purely driven by picture content, and further influenced by the aesthetics of the stimulus material. The main effect of "picture set" was not significant, $F(1,107)=1.81$, $p=.181$, 95% CI [-38.62, 29.25], η^2=.017.

3.4.AAT

Initial RT

A 2x2x2 ANCOVA considering the factors "picture content" (built vs. natural), "movement direction" (pull vs. push) and "picture set" (B1N9 vs. B5N5) while controlling for "age" revealed no statistically significant main effects (picture content: $F(1,103)=1.90$, $p=.171$, 95% CI [3.43, 8.60], η^2=.018; movement direction: $F(1,103)=3.40$, $p=.068$, 95% CI [1.51, 13.25], η^2=.032; picture set: $F(1,103)=1.262E-4$, $p=.991$, 95% CI [-15.13, 4.36], η^2=1.226E-6). The two-way interaction "movement direction" x "picture content" indicating the prevalence of an AAT effect reached significance, $F(1,103)=4.36$, $p=.039$, η^2=.041. Follow- up *t* tests show that the effect is driven by an approach bias for natural environments: RTs for pulling (=approach) are significantly faster for natural compared to built picture content, while the other categories did not differ significantly (see *Table 5*).

The three-way interaction "movement direction" x "picture content" x "picture set" was not significant, $F(1,103)=0.85$, $p=.360$, η^2=.008. Thus, the observed AAT effect did not depend on the perceived aesthetics, but only on picture content.

Movement RT

We found a significant main effect of picture content, $F(1,103)=4.23$, $p=.042$, 95% CI [0.55, 2.47], η^2=.039. Reactions were quicker with respect to pictures showing natural than built environments. No other main effect reached significance (movement direction: $F(1,103)=0.24$, $p=.627$, 95% CI [-5.35, 4.10], η^2=.002; picture set: $F(1,103)=0.18$, $p=.672$, 95% CI [-8.17, 5.50], η^2=.002). The two-way interaction "movement direction" x "picture content" was also significant, $F(1,103)=10.42$, $p=.002$, η^2=.092. None of the follow up *t* tests reached significance (see *Table 6*). Plotting the interaction (see *Figure S1, Supplementary material*) shows a pattern which seems to support the existence of an approach bias towards nature (higher slope for pulling reactions, faster for pictures of natural than built content).

The three-way interaction "movement direction"x"picture content"x"picture set" did not reveal a significant influence of picture content on the AAT effect, $F(1,103)=2.19$, $p=.142$, η^2=.025.

Consequently, there is no evidence that the AAT bias is based on differences in aesthetics, but only depends on picture content.

Reliability of Reaction Time Tasks

The results of the reliability calculations for the reaction time tasks are presented in *Tables 7a-c* (Cronbach's α separately for each item group) and *Tables 8a-c* (Split-Half Reliability of difference scores). Due to missing values, some reliabilities had to be calculated based on a reduced sample size.

High reliabilities (all Cronbach's $\alpha > .9$) were reached for all tasks when reaction times for stimulus groups were regarded separately. Considering the difference scores, reliability turned out weak (low to moderate size) for the DPT and AAT (ranging from .01 to .77) as previously reported in the literature. However, we observed relatively high reliability for the IAT (~ .89), which may be due to the higher item number in this task (George & Mallery, 2003).

Relationship between implicit biases and explicit picture ratings

No significant correlations between the biases of the DPT and the liking ratings emerged, B1N9 (n = 107): r = .03, p = .784; B5N5: (n = 109): r = .09, p = .353; total (n = 107): r = .09, p = .386. By contrast, the biases of the IAT were significantly correlated to the explicit ratings, B1N9 (n = 109): r = .24, p = .014; B5N5: (n = 109): r = .33, p = 4.482E-4; total (n = 109): r = .34, p =2.701E-4, which can be interpreted as small to moderate effect size (Cohen, 1988). Participants with higher biases towards nature rated pictures of natural environments as more likeable than pictures of built environments regardless of their aesthetic beauty (the effect also emerged for the B5N5 picture set). However, regarding the AAT biases again no significant correlations to the picture ratings were found: initial RT: B1N9 (n = 105): r = -.11, p = .276; B5N5: (n = 105): r = .08, p = .429; total (n = 105): r = -.04, p = .663; move RT: B1N9 (n = 105): r = -.13, p = .183; B5N5: (n = 105): r = -.02, p = .856; total (n = 107): r = -.05, p = .646.

4. Discussion

In line with the biophilia hypothesis positing an innate tendency of humans to seek connection with nature we found evidence for a tendency to approach nature stimuli in all three implicit tests in the present study. In the DPT participants were shown pairs of pictures (one built one natural) on the screen and were asked to respond to the spatial side on which a visual probe was shown afterwards. What we observed is a tendency for faster responses when the probe appeared behind the natural picture. This phenomenon is typically explained as the result of an attentional bias for the respective picture category. Originally, the DPT paradigm has been developed using threatening vs. neutral stimuli and applied in individuals diagnosed with anxiety disorders (MacLeod, et al., 1986). Within the context of the present study we interpret the result as revealing that participants' attention seems to be more strongly drawn to pictures of the natural in comparison to built environments. A similar finding has previously been shown by Joye and colleagues (Joye, et al., 2013). Unfortunately, the task design does not enable us to determine whether the attention of participants is actually driven *towards* the natural pictures or actually *away* from the built pictures. In depressed patients the phenomenon that they are faster to respond to probes appearing after the presentation of negative information has recently been re-interpreted, as an attention bias away from positive content (Winer & Salem, 2016). However, this was only possible because it is quite obvious what a neutral condition in terms of affect is and against which positive and negative content could be compared. This is more complex when comparing natural and built environments where the neutral category is unclear and almost no research is available as of now.

In order to further explore our hypothesis, we conducted an IAT in which participants needed to classify the content of pictures into belonging to "city" or "landscape" and words (e.g. "to dodge") belonging to the category "approach" or "avoid". In line with the predictions of the biophilia hypothesis participants were indeed faster to classify pictures and words when "approach" and "landscape" as well as "avoid" and "city" were mapped onto the same buttons as compared to the opposite mapping. This implies that our participants automatically associate the concept "nature" with "approach" and "city" with "avoidance". The more congruent the key mapping and therewith tighter the link between the concepts in the mental representation of the participants is, the faster they can respond. However, we still cannot say whether the effect is driven by human beings' automatic tendency to approach nature or respectively the avoidance of cities.

The third implicit task that we administered, the AAT, lends itself to compare actual approach and avoidance movements that participants make in response to "natural" or "built" pictures. We observed that participants were significantly faster in pulling (approaching) natural pictures towards themselves rather than built pictures. In contrast there was no difference in pushing (avoiding) the two different picture types. This strongly suggests that the automatic tendencies that we have been observing across the different tasks are driven by a tendency to approach nature and not to avoid built environments / cities.

The previous environmental psychological literature oftentimes employed stimuli that did not only differ in terms of the displayed content (natural vs. built) but also in terms of aesthetic pleasantness. Typically, nature is much preferred compared to built environments (S. Kaplan, Kaplan, & Wendt, 1972). Even to the extent that unspectacular or mediocre natural views consistently elicit higher aesthetic preference than do all except a very small percentage of urban scenes (Ulrich, 1986). This calls many of the previous findings comparing natural vs. built environments (Joye, et al., 2013) into question since it is unclear whether the observed effects are due to differences in liking of the places or actually due to the place characteristics. To address these confounds formally, we performed each implicit task twice, once in a picture set which showed high discrepancies in scenic ratings between natural and built environments (B1N9) and one picture set where the aesthetic pleasantness ratings of individuals were not different from one another (B5N5). Across all tasks we did not observe any evidence for the observed effects to be limited to the picture sets with strong disparities in aesthetic pleasantness. Therefore, we feel confident to dismiss any explanation based on differences in aesthetic pleasantness.

The present study goes way beyond previous studies focussing on differences between natural and built environments in terms of aesthetic pleasantness ratings, since these previous explicit and conscious assessments may simply be based on common beliefs such as "nature does you good" instead of accurately reflecting the individuals' experiences, biases and motivations. Instead we employed six implicit tests that objectively verified that individuals possess an attentional bias towards and an automatic tendency to approach nature and therefore confirmed the biophilia hypothesis. This is in line with first results showing an association between "me" and "nature" in an IAT setting, that was related to environmental concern and connectedness (Bruni & Schultz, 2010). Similar methodology has previously been used to show that the concept of nature is implicitly associated with women (aka "mother nature"), by both sexes (Liu, Geng, Ye, & Zhou, 2019).

To obtain a better understanding how our measures of implicit biases towards nature relate to explicit measures, namely picture ratings regarding the likeability of the depicted places we looked into their associations. Only for the IAT, significant correlations emerged which shows a congruency of implicit and explicit measures of connection to nature. However, it may be possible that during the IAT – as opposed to the DPT and AAT – not only automatic, but also conscious components of processing are involved, as pictures have to be categorized by content, and categories (e.g. "landscape") have to be mentally paired to successfully achieve the task, whereas the picture content is actually irrelevant from the perspective of the participant in DPT and AAT. Apparently, this preliminary finding has to be extended by further investigations concerning the validity of the reaction time tasks and their relationship to explicit measures.

Surprisingly, although individuals do commonly rate natural environments as more pleasant than built environments, they systematically underestimate the hedonic benefit that spending time in nature gives them ("affective forecasting error") (Nisbet & Zelenski, 2011). Soga and Gaston describe the phenomenon that people's direct interaction with nature diminishes over generations which leads to a loss of nature's positive influence on health and well-being (Soga & Gaston, 2016). This demonstrates that individuals fail to maximise their time spent in nature and therefore miss opportunities to increase their happiness by going out into nature. It seems as if modern lifestyle erodes people's connection with nature.

Dual-process models (Evans & Frankish, 2009; Strack & Deutsch, 2004), which are often referred to in order to explain the working mechanisms behind implicit tests, posit that behavior is determined by two different information processing systems: automatic/impulsive vs. controlled/reflexive processing. The automatic system is captured by means of implicit tests, and assesses fast, implicit, effortless, affective and motivational responses to stimuli. In contrast the controlled processing is slow, effortful and explicit and encompasses conscious decision-making, choices based on personal goals and standards. Within the former, processes are assumed to be innate and to use heuristics that evolved to solve specific adaptive problems. In the latter, processes are taken to be learned, flexible, and responsive to rational norms (Evans & Frankish, 2009). Dual-process models assume that the two systems are in conflict and decisions are determined by the relative strength of both processes. Note that the two systems must not be regarded as distinct and isolated, but rather as interdependent capacities of mental processing as Keren and Schul criticize the common understanding of two-system theories (Keren & Schul, 2009).

It could be that the act of forecasting the effects of nature draws mostly on the controlled, reflexive system and therefore undermines the automatic tendency to seek nature out. The focus on and praise of the controlled processing and willpower that is characteristic of our present society may therewith contribute to a growing estrangement from our innate knowledge that we thrive in nature.

As an implication of our results, it seems necessary to facilitate people's contact to nature in order to foster mental health and prevent the emergence of psychological disorders. Strategies might include city planning (creating parks and green neighborhoods as opportunities to engage with nature) or education (programs at school/for parents to inform about the importance of direct contact to nature)(Soga & Gaston, 2016).

However, the results of this study should be interpreted with caution considering some limitations. The study sample was relatively small, consisting of rather young people living mainly in cities. Future research should address this problem and replicate our findings in a larger sample with a higher variance regarding sample characteristics, e.g. age and current residence. Additionally, implicit methods smilar to the tasks used in the present study have come under criticism lately (Gawronski, 2019) regarding their reliability and validity. The reliability measures for difference scores observed in the present study definitely support the aforementioned deficiency. However, quite contrary to this notion, we found high reliabilities considering stimulus groups separately and since those were used in the main analyses, we consider them most relevant. This higher reliability argues against the use of difference scores in the respective paradigms. Nevertheless, it seems indispensable to further scrutinize and advance implicit paradigms, especially when it comes to validity as it was already mentioned before. We did not address the question of validity in our study, in the context of Biophilia the reference criterion to internally validate the proven biases remains unclear. In future studies one may consider to use the *Inclusion of Nature in Self Scale* for validation (Martin & Czellar, 2016; Schultz, 2002). Most importantly, the relationship between the implicit biases and mental health (problems) should be investigated in future research, to put our hypothesis of a link between mental health problems and biophilia to test.

Taken together the biophilic tendency revealed by the presented implicit test results may provide a first step to understanding the preponderance of psychiatric diseases in urban contexts. Living at greater distance to and at places with lower availability of green spaces seems to act against an automatic and potentially deep-rooted need for contact with nature and may contribute to stress and in turn to the emergence of mental health problems than (or at least in addition to) environmental or societal stressors individuals are exposed to in cities.

5. References

Abbott, A. (2012). Stress and the city: Urban decay. *Nature, 490*, 162-164.

Alcock, I., White, M. P., Wheeler, B. W., Fleming, L. E., & Depledge, M. H. (2014). Longitudinal effects on mental health of moving to greener and less green urban areas. *Environ Sci Technol, 48*, 1247-1255.

Berman, M. G., Jonides, J., & Kaplan, S. (2008). The cognitive benefits of interacting with nature. *Psychol Sci, 19*, 1207-1212.

Bruni, C. M., & Schultz, P. W. (2010). Implicit beliefs about self and nature: Evidence from an IAT game. *Journal of Environmental Psychology, 30*, 95-102.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum Associates.

de Vries, S., Ten Have, M., van Dorsselaer, S., van Wezep, M., Hermans, T., & de Graaf, R. (2016). Local availability of green and blue space and prevalence of common mental disorders in the Netherlands. *BJPsych Open, 2*, 366-372.

Dye, C. (2008). Health and urban living. *Science, 319*, 766-769.

Eberhardt, M. S., & Pamuk, E. R. (2004). The importance of place of residence: examining health in rural and nonrural areas. *Am J Public Health, 94*, 1682-1686.

Engemann, K., Pedersen, C. B., Arge, L., Tsirogiannis, C., Mortensen, P. B., & Svenning, J. C. (2019). Residential green space in childhood is associated with lower risk of psychiatric disorders from adolescence into adulthood. *Proc Natl Acad Sci U S A, 116*, 5188-5193.

Evans, J., & Frankish, K. (2009). *In two minds: Dual processes and beyond*.

Gawronski, B. (2019). Six Lessons for a Cogent Science of Implicit Bias and Its Criticism. *Perspect Psychol Sci, 14*, 574-595.

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update (4th ed.)*. Boston: Allyn & Bacon.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*, 1464-1480.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *J Pers Soc Psychol, 85*, 197-216.

Holz, N. E., Tost, H., & Meyer-Lindenberg, A. (2020). Resilience and the brain: a key role for regulatory circuits linked to social stress and support. *Mol Psychiatry, 25*, 379-396.

Hubbard, G., Thompson, C. W., Locke, R., Jenkins, D., Munoz, S. A., Van Woerden, H., Maxwell, M., Yang, Y., & Gorely, T. (2020). Co-production of "nature walks for wellbeing" public health intervention for people with severe mental illness: use of theory and practical know-how. *BMC Public Health, 20*, 428.

Joye, Y., Pals, R., Steg, L., & Evans, B. L. (2013). New methods for assessing the fascinating nature of nature experiences. *PLoS One, 8*, e65332.

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

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

Kennedy, D. P., & Adolphs, R. (2011). Social neuroscience: Stress and the city. *Nature, 474*, 452-453.

Keren, G., & Schul, Y. (2009). Two Is Not Always Better Than One: A Critical Evaluation of Two-System Theories. *Perspect Psychol Sci, 4*, 533-550.

Khan, A., Plana-Ripoll, O., Antonsen, S., Brandt, J., Geels, C., Landecker, H., Sullivan, P. F., Pedersen, C. B., & Rzhetsky, A. (2019). Environmental pollution is associated with increased risk of psychiatric disorders in the US and Denmark. *PLoS Biol, 17*, e3000353.

Lawrence, N. S., O'Sullivan, J., Parslow, D., Javaid, M., Adams, R. C., Chambers, C. D., Kos, K., & Verbruggen, F. (2015). Training response inhibition to food is associated with weight loss and reduced energy intake. *Appetite, 95*, 17-28.

Liu, T., Geng, L., Ye, L., & Zhou, K. (2019). "Mother Nature" enhances connectedness to nature and pro-environmental behavior. *Journal of Environmental Psychology, 61*, 37-45.

Logan, A. C., Katzman, M. A., & Balanza-Martinez, V. (2015). Natural environments, ancestral diets, and microbial ecology: is there a modern "paleo-deficit disorder"? Part I. *J Physiol Anthropol, 34*, 1.

MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology, 95*, 15-20.

Martin, C., & Czellar, S. (2016). The extended Inclusion of Nature in Self scale. *Journal of Environmental Psychology, 47*, 181-194.

Newbury, J. B., Arseneault, L., Beevers, S., Kitwiroon, N., Roberts, S., Pariante, C. M., Kelly, F. J., & Fisher, H. L. (2019). Association of Air Pollution Exposure With Psychotic Experiences During Adolescence. *JAMA Psychiatry, 76*, 614-623.

Nisbet, E. K., & Zelenski, J. M. (2011). Underestimating nearby nature: affective forecasting errors obscure the happy path to sustainability. *Psychol Sci, 22*, 1101-1106.

Nock, M. K., Park, J. M., Finn, C. T., Deliberto, T. L., Dour, H. J., & Banaji, M. R. (2010). Measuring the suicidal mind: implicit cognition predicts suicidal behavior. *Psychol Sci, 21*, 511-517.

Park, B. J., Tsunetsugu, Y., Kasetani, T., Kagawa, T., & Miyazaki, Y. (2010). The physiological effects of Shinrin-yoku (taking in the forest atmosphere or forest bathing): evidence from field experiments in 24 forests across Japan. *Environ Health Prev Med, 15*, 18-26.

Paslakis, G., Kühn, S., Grunert, S., & Erim, Y. (2017). Explicit and Implicit Approach vs. Avoidance Tendencies towards High vs. Low Calorie Food Cues in Patients with Obesity and Active Binge Eating Disorder. *Nutrients, 9*.

Pedersen, C. B., & Mortensen, P. B. (2001). Evidence of a dose-response relationship between urbanicity during upbringing and schizophrenia risk. *Arch Gen Psychiatry, 58*, 1039-1046.

Peen, J., & Dekker, J. (2004). Is urbanicity an environmental risk-factor for psychiatric disorders? *The Lancet, 363*, 2012-2013.

Peen, J., Schoevers, R. A., Beekman, A. T., & Dekker, J. (2010). The current status of urban-rural differences in psychiatric disorders. *Acta Psychiatr Scand, 121*, 84-93.

Pronk, T., Molenaar, D., Wiers, R. W., & Murre, J. (2021). Methods to split cognitive task data for estimating split-half reliability: A comprehensive review and systematic assessment. *Psychon Bull Rev*.

R, R. D. C. T. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria, https://www.R-project.org/*.

Rinck, M., & Becker, E. S. (2007). Approach and avoidance in fear of spiders. *J Behav Ther Exp Psychiatry, 38*, 105-120.

Schoenmakers, T., Wiers, R. W., & Field, M. (2008). Effects of a low dose of alcohol on cognitive biases and craving in heavy drinkers. *Psychopharmacology (Berl), 197*, 169-178.

Schultz, P. (2002). Inclusion with Nature: The Psychology of Human-Nature Relations. In P. Schmuck & W. Schultz (Eds.), *Psychology of Sustainable Development*. Boston, MA: Springer.

Soga, M., & Gaston, K. J. (2016). Extinction of experience: the loss of human-nature interactions. *Frontiers in Ecology and the Environment, 14*, 94-101.

Solarz, A. K. (1960). Latency of instrumental responses as a function of compatibility with the meaning of eliciting verbal signs. *J Exp Psychol, 59*, 239-245.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Pers Soc Psychol Rev, 8*, 220-247.

Tillmann, S., Tobin, D., Avison, W., & Gilliland, J. (2018). Mental health benefits of interactions with nature in children and teenagers: a systematic review. *J Epidemiol Community Health, 72*, 958-966.

Tost, H., Champagne, F. A., & Meyer-Lindenberg, A. (2015). Environmental influence in the brain, human welfare and mental health. *Nat Neurosci, 18*, 1421-1431.

Trostrup, C. H., Christiansen, A. B., Stolen, K. S., Nielsen, P. K., & Stelter, R. (2019). The effect of nature exposure on the mental health of patients: a systematic review. *Qual Life Res, 28*, 1695-1703.

Ulrich, R. S. (1986). Human responses to vegetation and landscapes. *Landscape and Urban Planning, 13*, 29-44.

Ulrich, R. S., Simons, R. F., Losito, B. D., Fiorito, E., Miles, M. A., & Zelson, M. (1991). Stress recovery during exposure to natural and urban environments. *Journal of Environmental Psychology, 11*, 201-230.

van Ens, W., Schmidt, U., Campbell, I. C., Roefs, A., & Werthmann, J. (2019). Test-retest reliability of attention bias for food: Robust eye-tracking and reaction time indices. *Appetite, 136*, 86-92.

Waechter, S., Nelson, A. L., Wright, C., Hyatt, A., & Oakman, J. (2013). Measuring Attentional Bias to Threat: Reliability of Dot Probe and Eye Movement Indices. *Cognitive Therapy and Research, 38*, 313-333.

Wagner, K. H., & Brath, H. (2012). A global view on the development of non communicable diseases. *Prev Med, 54 Suppl*, S38-41.

Whitburn, J., Linklater, W., & Abrahamse, W. (2020). Meta-analysis of human connection to nature and proenvironmental behavior. *Conserv Biol, 34*, 180-193.

Wiers, C. E., Kuhn, S., Javadi, A. H., Korucuoglu, O., Wiers, R. W., Walter, H., Gallinat, J., & Bermpohl, F. (2013). Automatic approach bias towards smoking cues is present in smokers but not in ex-smokers. *Psychopharmacology (Berl), 229*, 187-197.

Wiers, R. W., Eberl, C., Rinck, M., Becker, E. S., & Lindenmeyer, J. (2011). Retraining automatic action tendencies changes alcoholic patients' approach bias for alcohol and improves treatment outcome. *Psychol Sci, 22*, 490-497.

Wilson, E. (1984). *Biophilia: The Human Bond with Other Species*. Cambridge: Harvard University Press.

Winer, E. S., & Salem, T. (2016). Reward devaluation: Dot-probe meta-analytic evidence of avoidance of positive information in depressed persons. *Psychol Bull, 142*, 18-78.

Figure captions

Figure 1: Overview over the implicit test paradigms and the picture sets used.

Figure 2: Dot-probe task (DPT): main effect "congruency" (congruent = probe at the position of previous landscape picture, incongruent =probe at the position of previous city picture). The covariate in the model was calculated as follows: age = 28.46. When the *-outlier were removed from the dataset, values changes slightly, but there was no alteration of significances.

Figure 3: Implicit association test (IAT) effect – two way interaction of "congruency" (congruent = approach-landscape; incongruent = approach-city) x "picture set" based on marginal means. The covariate in the model was calculated as follows: age = 28.36

Figure 1

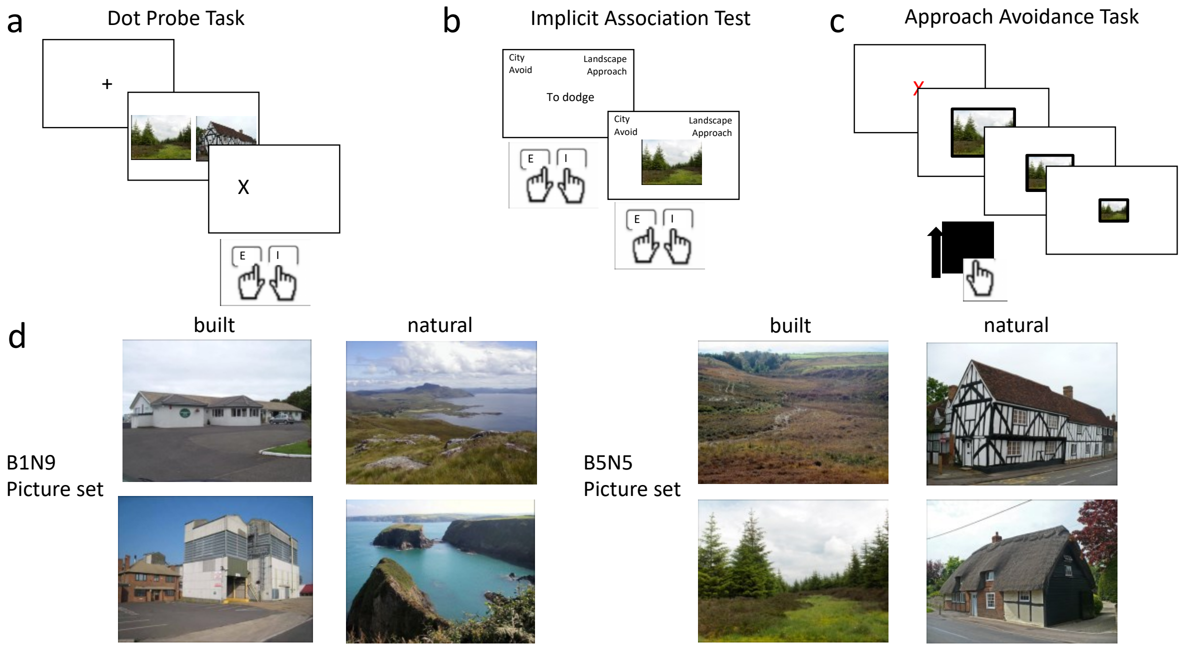


Figure 2

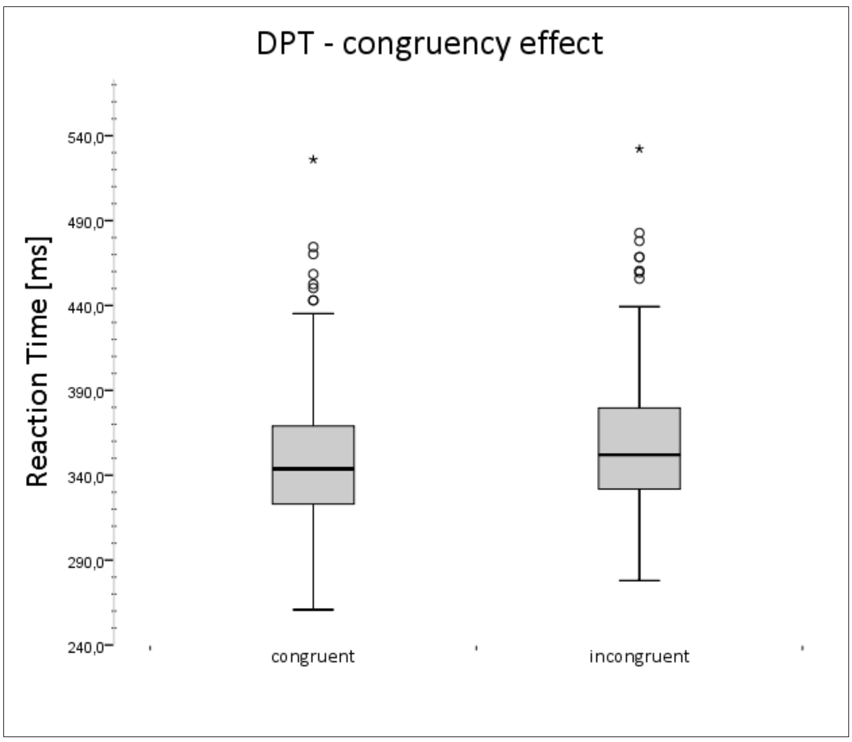
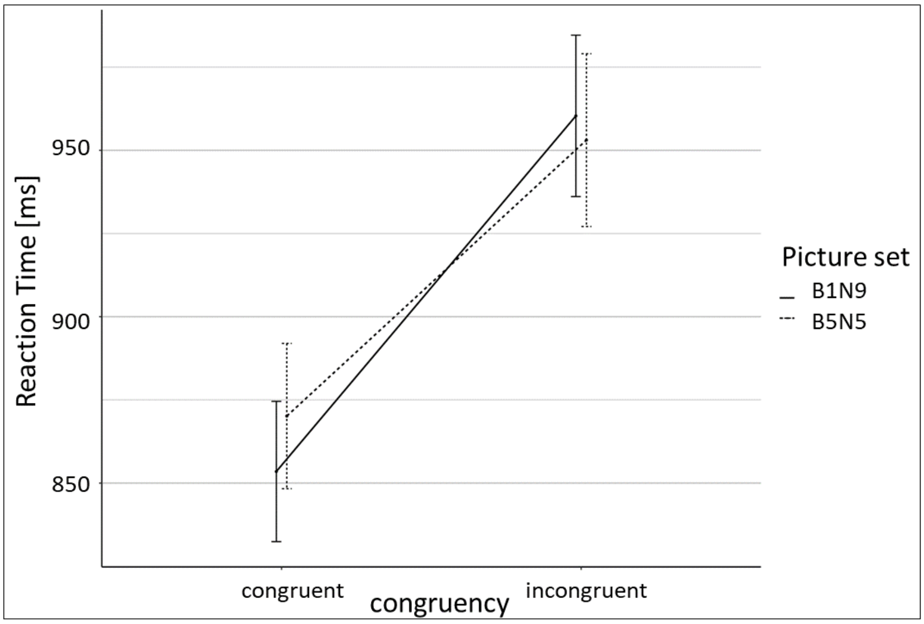


Figure 3



Tables

Table 1: Sample Characteristics

Sample characteristics		**Number or Mean**
Sex	male:	*n* = 70
	female:	*n* = 36
Age (*years*)		*M* = 28.36, *SD* = 10.22
Highest level of education	No qualification:	*n* = 0
	Leaving secondary school without graduation:	*n* = 0
	Secondary school (9 years):	*n* = 1
	Secondary school (10 years):	*n* = 8
	High school:	*n* = 100
Current place of residence	City (> 100 000 inhabitants):	*n* = 70
	Town (> 10 000 inhabitants):	*n* = 17
	Rural area:	*n* = 22
Place of growing up for the majority of years until the age of 15[a]	City (> 100 000 inhabitants):	*n* = 21
	Town (> 10 000 inhabitants):	*n* = 29
	Rural area:	*n* = 36
	Information not specified:	*n* = 23

Notes: [a] categorization based on simple majority

Table 2: Mean and Standard Deviation of aesthetic ratings per picture group

	Previous online ratings[a]		***Ratings of study sample***[b]	
Picture group	***M***	***SD***	***M***	***SD***
B1	1.45	0.05	2.34	1.56
B5	4.87	0.61	5.64	1.52
N5	4.61	0.32	5.72	1.49
N9	9.03	0.17	8.74	0.99

Note: [a] *Ratings from 0 to 10, higher scores indicating higher aesthetics, sample sizes range from N = 8 to N = 14 for each picture. Ratings were obtained from the website "Scenic or not" (http://scenicornot.datasciencelab.co.uk/);* [b] *Original ratings from 0 to 100, higher scores indicating higher aesthetics. Original ratings were divided by 10 in order to obtain comparable values. N = 109 for all pictures. Ratings originated from the picture rating conducted in the present study.*

Table 3: German words used as stimuli during the IAT and their English translations

	Approach		**Avoidance**	
	German word	*English translation*	*German word*	*English translation*
1.	nehmen	to take	vermeiden	to avoid
2.	berühren	to touch	ausweichen	to dodge
3.	anfassen	to touch	wegschieben	to push away
4.	ranholen	to fetch	entfernen	to take off
5.	annähern	to approach	verschwinden	to disappear

Table 4: Paired t tests to determine differences in ratings regarding the aesthetics of the selected pictorial stimuli groups

Paired differences	***M***	***SD***	***T***	***df***	***p***	***98.75% CI of difference***
B1 – N9	-64.04	20.18	-33.12	108	2.146E-58*	[-68.95, -59.13]
B5 – N5	-0.86	17.19	-0.52	108	.603	[-5.04, 3.32]
B1 – B5	-32.99	14.61	-23.58	108	2.150E-44*	[-36.55, -29.44]
N5 – N9	-30.18	13.18	-23.92	108	5.889E-45*	[-33.39, -26.98]

Notes. *significant based on a corrected α = .0125

Table 5: Initial reaction times in the AAT - Paired t tests comparing combinations of the factors “movement direction” and “picture content” relevant to clarify their interaction effect

Paired differences	*M*	*SD*	*T*	*df*	*p*	*98.75% CI of difference*
pull, built – pull, natural	6.70	19.05	3.61	104	4.796E-4*	[1.98, 11.43]
push, built – push, natural	5.32	22.01	2.48	104	.015	[-0.14, 10.78]
pull, built – push, built	8.07	34.99	2.36	104	.020	[-0.61, 16.75]
pull natural – push, natural	6.69	33.38	2.05	104	.042	[-1.59, 14.97]

Notes. *significant based on a corrected α = .0125

Table 6: Movement reaction times in the AAT - Paired t tests comparing combinations of the factors “movement direction” and “picture content” relevant to clarify their interaction effect

Paired differences	*M*	*SD*	*T*	*df*	*p*	*98.75% CI of difference*
pull, built – pull, natural	1.93	8.12	2.43	104	.017	[-0.09, 3.94]
push, built – push, natural	1.09	8.34	1.33	104	.185	[-0.98, 3.15]
pull, built – push, built	- 0.20	26.01	- 0.08	104	.936	[-6.66, 6.25]
pull natural – push, natural	- 1.05	24.27	- 0.44	104	.659	[-7.07, 4.97]

Notes. *significant based on a corrected α = .0125

Table 7a: Reliability of the DPT (N = 107), 10 Items (10 picture pairs)

Picture set/ Congruency	*n*	Cronbach’s α
B1N9/congruent	107	.979
B1N9/incongruent	107	.978
B5N5/congruent	107	.978
B5N5/incongruent	107	.978

Table 7b: Reliability of the IAT (N = 109), 30 Items (20 pictures per set B1N9 or B5N5 and 10 words of both categories approach and avoidance)

Picture set/ Congruency	*n*	Cronbach’s α
B1N9/congruent	108	.951
B1N9/incongruent	108	.944
B5N5/congruent	105	.958
B5N5/incongruent	107	.960

Table 7c: Reliability of the AAT (N = 105), 10 Items (10 pictures per category B1, B5, N5, N9)

Movement direction/ picture content/ picture set	***n***	**Cronbach's α Initial RT**	**Cronbach's α Movement RT**
pull/built/B1N9	104	.935	.922
pull/built/B5N5	105	.926	.942
pull/natural/B1N9	105	.949	.949
pull/natural/B5N5	104	.937	.937
push/built/B1N9	105	.948	.944
push/built/B5N5	105	.950	.950
push/natural/B1N9	105	.948	.942
push/natural/B5N5	104	.954	.935

Table 8a: DPT - Split-Half Reliability of the difference score (incongruent-congruent)

Picture set	***n***	**Spearman-Brown Coefficient**
B1N9	107	.100
B5N5	107	.382

Table 8b: IAT - Split-Half Reliability of the difference score (incongruent-congruent)

Picture set	***n***	**Spearman-Brown Coefficient**
B1N9	107	.893
B5N5	104	.890

Table 8c: AAT - Split-Half Reliability of the difference score (push-pull) per picture content and picture set

picture content/ picture set	***n***	**Spearman-Brown Coefficient *Initial RT***	**Spearman-Brown Coefficient *Movement RT***
built/B1N9	104	.219	.609
natural/B1N9	105	.665	.607
built/B5N5	105	.491	.776
natural/B5N5	103	.478	.591