The Effect of Sound on Attention Restoration

A Senior Honors Thesis

Submitted in Partial Fulfillment of the Requirements
for Graduation in the Honors College

By

Matthew W. Bennett

Psychology Major

The College at Brockport
May 17, 2019

Thesis Director: Dr. Melissa M. Brown, Assistant Professor, Psychology

Abstract

Prolonged use of directed attention leads to cognitive fatigue, characterized in part by a reduced capacity to maintain focus. Studies have shown that recorded nature sounds can have a stimulating effect on the process of recovering from attentional fatigue; other studies also suggest that the emotional response to sound may affect the restoration of directed attention. To see if a preference for natural sounds might explain their restorative quality, we compared the effects of bird calls and music on the rate of attention restoration. We induced cognitive depletion in 116 college undergraduates and then tested their attentional capacity using a digit span backwards task before and after exposure to a sound condition. We also used the Geneva Emotional Music Scale to evaluate participants' emotional response to the sound. Hypothesis tests revealed no statistically significant effect of sound type on restoration, nor any significant correlations between measures of emotional response to music and attention restoration. Findings do offer some weak support for our hypothesis; we discuss this alongside problems with the study design and suggestions for future studies.

The Effect of Sound on Attention Restoration

Of all the cognitive processes humans are generally aware of, none may be as frequently used as attention. We use attention to maintain our cognitive focus on a variety of objects and thoughts: things interesting and uninteresting, transcendent and mundane, stimulating and sleep-inducing. Modern psychologists' interest in attention stretches back at least to William James, who identified two major systems of attention: *involuntary attention*, which is responsible for focusing on naturally captivating and exciting objects, and *voluntary attention* (commonly called *directed attention*) which keeps our focus on less-interesting phenomena and requires willful effort to maintain (James 1892, quoted in Kaplan and Berman 2010). Kaplan (1995) describes the mechanism of directed attention as a combination of the concepts of the *will* and voluntary attention and attributes certain executive abilities (such as the ability to focus and inhibition) to this mechanism (Kaplan 1995; Kaplan and Berman 2010).

One theoretical and research paradigm that relies on directed and involuntary attention is Attention Restoration Theory (ART), proposed by Kaplan and Kaplan (1989). ART states that the prolonged use of the directed attention system leads to a state of fatigue, where a person's mind begins to wander and their task performance begins to suffer. ART also states that interactions with nature are conducive to the restoration of the directed attention system and attempts to explain why exposure to nature may reverse this kind of cognitive depletion (Kaplan 1995). The central premise of ART is that the directed attention system goes through a restorative period when it is able to rest and the involuntary attention system takes over; furthermore, because natural settings are full of the kind of naturally captivating stimuli that engage the involuntary attention system, natural environments tend to be more restorative than anthropogenic environments (Kaplan and Kaplan 1989, Kaplan 1995).

According to ART, there are four required characteristics for an environment to be restorative: being away, fascination, extent, and compatibility (Kaplan 1995). *Being away* refers to a sense of being removed from everyday worries and stressors. *Fascination* refers to an environment being innately interesting, to encourage the activation of the voluntary attention system instead of the involuntary system. *Extent* refers to a feeling of the environment feeling expansive, and therefore being able to fully encompass an individual, and *Compatibility* refers to a sense of being compatible with the environment—a “special resonance between the natural setting and human inclinations” (Kaplan 1995, p. 174). When an environment has all of these characteristics, it has the necessary qualities to induce attentional (or, cognitive) restoration.

This construct of directed attention is similar, conceptually, to Working Memory (WM), a model of short-term memory first proposed by Baddeley and Hitch (Baddeley and Hitch 1974) and, specifically, to WM capacity. WM is the part of short-term memory that interacts most directly with cognitive representations of physical stimuli and part of the construct—the central executive—is closely related to directed attention (Baddeley 1996; 2001). The two constructs are so interrelated that ART relies on tests of WM capacity as measures of directed attention.

Engle (2002) suggests that WM capacity—conceptualized as the ability to maintain cognitive focus on important information despite distractors—is synonymous with executive attention. There are multiple ways to measure WM capacity, including digit span tasks (forward and backward; see Woods, et al. 2011 for a discussion), word span tasks (Mattys, Baddeley, and Trenkic 2018), and the Stroop task (Kane and Engle 2003). Traditionally, ART research uses digit span tasks alongside other tasks to measure directed attention (see Ohly et al. 2016 for more detail concerning the different research methodologies used in ART research).

Recent studies using the ART framework have included investigations of how natural sounds and soundscapes affect cognitive restoration, and have found a positive effect (Payne 2013; Ratcliffe, Gatersleben, and Sowden 2013, 2016; Abbott et al. 2016). Much of this work has evaluated how participants interpret the sounds they hear. For instance, Ratcliffe, Gatersleben, and Sowden (2016) found that the calls of bird species that were associated with more relaxing environments (e.g. songbirds, which were associated with calming, green environments) were rated by participants as being more restorative than birds evocative of more tumultuous environments.

This characteristic of stimuli—source attribution—has been shown to be able to affect the perceived restorative potential of a stimuli for other sounds as well. Participants who listened to pink noise (static-like, synthetic sound) and were told they were hearing a waterfall found it more relaxing and perceived it as more restorative than those who were told it was the result of industrial equipment. Moreover, of those who were not told the sound's possible source (and therefore made their own attributions), those who thought it was a natural sound rated it as more restorative than those who thought it was anthropogenic in nature (Haga, Halin, Holmgren, and Sörqvist 2016). In both of these cases, stimuli that evoked more calming environments (which may be higher in the "compatibility" quality of ART) were perceived as more relaxing and restorative. However, in the second study, there was no significant difference in the improvement of performance on the Attention Network Test (a test designed to measure specific executive functions; Fan et al. 2002) between the groups. That does not mean that there was no restoration of directed attention, however—the Attention Network Test has failed, in other studies, to find evidence in support of ART whereas studies that use measures that focus more on general executive functioning (such as DSF and DSB) have (see Ohly et al. 2016).

The reason for the difference in perceived restoration potential may be connected to the emotional response evoked by the image of the stimulus' source. According to basic emotion theory, emotions can be categorized according to two major dimensions: valence (whether the emotion is pleasurable or not) and arousal (a sense of energy and excitement accompanying the emotion, ranging from low to high; Russell 1980; Kuhbander and Zehletleitner 2011; Zentner, Grandjean, and Scherer 2008). Baldwin and Lewis (2017) sorted a selection of popular music according to these two dimensions, ranking them according to the mean score given each to each song by a group of participants ("speed", e.g. how slow or fast participants felt the song tempo was, was substituted for "arousal"). Songs were also evaluated using a version of the Geneva Emotional Music Scale-9 (GEMS-9) scale, which evaluates the emotions induced by music (Zentner, Grandjean, and Scherer 2008; the GEMS scales will be explained in more detail below). Baldwin and Lewis (2017) then tested a second group to see whether a song's quality affected performance on a Sustained Attention to Response Task (a Go-NoGo task used to assess response inhibition, an executive function task). They found that listening to positively-valenced music increased performance, and the effect was enhanced if the participant liked the song (as measured by a 5-point Likert scale)—that is, participants had the most pronounced response if they had a positive emotional response to a positively-valenced song.

Somewhat paradoxically, Jefferies, Smilek, Eich, and Enns (2008) found that, after inducing various affective states in participants via music, that those who had sadness induced (marked by low valence and low arousal) had the highest increase in performance on an attentional blink task, while those who had calmness and happiness (positive valence-low arousal and positive valence-high arousal, respectively) performed only intermediately. The final group, an anxiety (negative valence-high arousal) group, performed the most poorly. In the interpretation of these results, Jefferies and colleagues note that this performance increase was only for part of the test, and the test only evaluated visual attention. Therefore, sadness may have an effect in how visual attention functions or items are prioritized, but this may not generalize to the executive functions of attention. One possible confound for this experiment, which Baldwin and Lewis controlled for, was how familiar the music was to the participant. Pereira and colleagues found that familiarity with pieces of music was a key factor in triggering the blood oxygen level dependence response in various emotion-related regions in the brain (2011). It is possible that individual differences in the degree of familiarity with the music used to induce the desired mood confounded the data.

Another possibility to explain the different findings is that Baldwin and Lewis (2017) induced emotion by music alone, whereas Jefferies et al. (2008) focused on inducing emotion through a combination of music and remembering past experiences where they felt that emotion. According to emotion refinement theory, emotions can be broadly divided into two categories: ‘coarse’ and ‘refined’ (Frijda and Sundararajan 2007; Zentner and Eerola 2011). Coarse emotions are those which are strongly connected to action tendencies and may have played a strong adaptive role in human survival (e.g., the physiological activity accompanying the positive emotions associated with being with friends may increase social behavior that historically increased survival rates; see Winkielman & Cacioppo 2001). Refined emotions, however, have an attenuated connection to their related action tendency and are, therefore, more often only felt and not acted upon (Frijda and Sundararajan 2007). Zentner, Grandjean, and Scherer (2008) suggest the existence of an emotion subcategory: ‘aesthetic emotions’, a subset of refined emotions induced by aesthetic experiences, like listening to music. It is possible that aesthetic and coarse emotions have different effects on body physiology (their different relationships to action tendencies suggests this) and, if so, they may have different effects on cognitive processes such as cognitive restoration. Jefferies et al. (2008) induced emotion by music *and* the recall of a past emotional experience—it is possible that this method may induce a significantly different affective state than if emotion were induced through music alone, as Baldwin and Lewis (2017) did.

The different effects of refined or aesthetic emotions vs. coarse emotions on physiology and cognition may also help explain the restorative effect of nature on the mind.

Previous studies have concluded there is a link between how restorative an environment is and whether a threat was present (i.e. the presence of a threat was connected to less restoration induced by an environment; Herzog and Rector 2009, Andrews and Gatersleben 2010). Ratcliffe, Gatlersleben, and Sowden (2016) cite this as an explanatory factor in why, in their study, certain bird species were rated as being more restorative than others—each bird's call was evocative of the environment in which it lived. Certain environments were interpreted as safer—especially lush, forested, or green areas, which are often associated with safety. It is possible that the link between the type of bird call and their natural environment could explain the potential for restoration inherent in listening to bird calls.

One further implication of this difference in aesthetic and coarse emotions is that a person could confound coarse and aesthetic emotions and use the same word to describe different emotion states. While discussing this problem, Zentner, Grandjean, and Scherer (2008) use the example of sadness. Generally, sadness (and closely related emotions, like melancholy) is an unpleasant, negatively-valenced emotion. However, musically-induced 'sadness' was often unconnected to an aversive response to it in their studies. They quote Levinson (1990) in their interpretation that the various sad-like "coloring[s] of consciousness" experienced while listening to music are devoid of accompanying, unpleasant thoughts feelings (such as a sense of loss accompanying the death of a loved one). Consequently, these sad-like feelings can be experienced solely as aesthetic emotions, allowing a person to "savor" the emotion "like the bitter taste of whiskey" (Zentner, Grandjean, and Scherer 2008). When discussing emotion, therefore, we should be cautious not to confound the coarse emotions experienced in everyday life with the aesthetic emotions felt while listening to music. Connections do exist between music and emotions, but the link may be more complex than expected.

Different theories on cognitive restoration and WM capacity posit different causes and mechanisms. There is a theme of positive affect aiding cognitive restoration, but the direct mechanisms remain in question. Joye and van den Berg (2011) argue that perceptual fluency plays a role, and they suggest that part of natural-environment-induced restoration stems from a positive affective state resulting from the ease of perceiving certain visual qualities (i.e. perceptual fluency; see also Reber, Schwarz, and Winkielman 2004 and Joye, Ünal, and Pals 2016). This is harmonious with the assumption made by ART that restoration in natural environments stems from a person's comfort while in them. In ART, evolutionary accounts are often employed to explain the restorative mechanisms (Joye and van den Berg 2011); however, the fact that the similar cognitive restoration can be achieved by listening to music (Kuhbandner and Zehetleitner 2011; Baldwin and Lewis 2017; Jefferies et al. 2008) suggests that either the restorative processes that result from nature exposure are not domain-specific (i.e., that the reasons nature-bound restoration occurs is due to mechanisms that are also activated by music) or that there are multiple pathways to restoration, and that nature exposure and music exposure activate different pathways.

One way to test this would be to test whether there are differential restorative effects resulting from exposure to sounds already shown to encourage restoration, to see if a nature-based stimulus had a greater effect than an anthropogenic one. If, as ART posits, there is something unique about natural environments' restorative potential, one would expect the natural stimulus to have a greater effect than the anthropogenic. If, however, sound-based restoration arising from natural and man-made sources show the same restorative power, that lends credence to the domain-free restorative mechanism hypothesis. Additionally, if one's affective response to sound affected restoration more than the sound type, that would show a great deal of support to the domain-free restoration theory and be further evidence for the general link between affect and restoration.

This research has bearing on a wide range of human activity, including city planning. ART strongly implies that we should include natural settings and qualities in places where humans live and work, but the more we learn about *why* nature helps restore cognition and reduce stress, the more power we will have to be able to create healthy urban landscapes. If it is the case that the affective response to stimuli is the main component in cognitive restoration, or if it perceptual fluency encourages positive affective responses, then efforts to ensure future cityscapes and buildings incorporate easy-to-perceive qualities will pay off greatly in terms of the encouragement of mental health in urban populations.

There are two central questions for this project. The first is whether one category of stimuli (natural or anthropogenic) has a greater potential for inducing cognitive restoration than another. The second questions is whether the affective response to stimuli is a mediating factor in cognitive restoration. To test these questions, we designed a study that would measure the rate of cognitive restoration, using DSB scores as measures. We formulated two hypotheses at the outset: The first is that natural sounds will have a greater positive effect on attention restoration than music, and the second is that a positive affective response will enhance experienced restoration.

Method

Participants

116 undergraduate university students (89 women, M_{age} = 19.53 years, age range: 18-51) were recruited from introductory psychology classes. Participants were recruited via the SONA web-based participant recruitment system: information about the experiment was posted and students volunteered in exchange for class credit. The study was approved by the Institutional Review Board at The College at Brockport (10/19/2018).

Materials

Affective measures: PANAS and GEMS-45. The Positive and Negative Affect Scale (PANAS; Appendix A) assesses how strongly participants are feeling different emotions in the present moment using a 5-point Likert scale (Watson, Clark, and Tellegen 1988). It has 20 items divided into two subscales (one for positive [PosAff] and one for negative emotions [NegAff]). Each item is an adjective describing an emotion; participants are asked to rate how strongly they are feeling that emotion at the present moment. Participants were given the PANAS twice: once at the beginning of the experiment and once at the very end, after the GEMS-45.

The Geneva Emotional Music Scale-45 (GEMS-45; Appendix B) is a 45-item questionnaire that uses a 5-point Likert scale to measure the emotion that participants experienced while listening to their assigned sound. Participants were asked how much the sound induced certain emotions within themselves, using common descriptive words or phrases (such as ‘fascinated’ or ‘filled with wonder’). The GEMS-45 contains nine subscales, each describing a musical emotion (Wonder, Transcendence, Power, Tenderness, Nostalgia, Peacefulness, Joyful Activation, Sadness, and Tension) and three supraordinate scales representing more broad emotion categories (Sublimity, Vitality, and Unease). The GEMS-45 was given immediately after the second administration of the DSB test and asked participants about the emotions induced by their assigned sound clip. Permission to use the GEMS-45 was obtained by Dr. Marcel Zentner.

Cognitive depletion task. The cognitive depletion task (Appendix C) consisted of two parts: either a nine-digit or a six-digit number which participants memorized (485126953 or 485126; all participants were given the nine-digit number first and given the six-digit only if they failed to memorize the first within 2 minutes, and none needed a shorter number), and a list of 75 increasingly difficult math problems designed to tax the working memory of the average undergraduate student (complex math problems have been shown to engage the central executive; see Otsuka and Osaka 2014 and Hubber, Gilmore, and Cragg 2014). After memorizing the number, participants were given 15 minutes to complete as many math problems as possible, with the option to use a calculator if they reached a more difficult section including algebra problems. After the time was over, participants were asked to repeat the memorized number to the researcher.

Digit span backwards. We used a DSB test based on the work of Woods and Colleagues (2011). The DSB test was administered electronically via the Inquisit computerized platform, a subscription-based computerized test library. The test requires participants to listen to a string of digits spoken by the program and then use a mouse to click on-screen numbers to enter the correct (backwards) sequence. The test uses a 1:2 staircase method—for every span a participant gets correct, the next span's length increases by one digit; if a participant fails twice in a row, the span length is reduced by one. The program starts with on-screen instructions and two trial runs to ensure that participants understand the process before starting the real test.

The test measured participant performance via four metrics: two-error maximum length (the longest span length before a participant made two successive errors; TE_ML), two-error total trial (the total number of trails a participant made before two successive errors; TE_TT), maximum length (the longest span length completed correctly; ML), and mean span (a computation measuring the list length where half of trials would be correctly completed; MS). We categorized TE_ML and TE_TT as "two-error metrics" and MS and ML as "overall metrics". Two-error metrics reflect participant performance in the first part of the test and have poor test-retest reliability: TE_ML, $r = .67$; TE_TT, $r = .53$. Overall metrics measure performance through the whole test and have much better test-retest reliability: ML, $r = .81$; MS, $r = .84$ (Woods et al. 2011, Table 3).

Each participant took the DSB test twice: once right after the Cognitive Depletion Task, and another right after listening to their assigned sound. The test was administered via a Dell Latitude E 5540 15" laptop connected to AUVIO Concert Class Stereo Headphones.

Sound clips. Three different sound clips were purchased on Amazon.com and modified using Audacity (an open-source audio editor) to ensure all were of the same length of four minutes. The sound clips consisted of bird songs (a Wood Thrush call), the song "Somewhere Over the Rainbow" by Israel Kamakawiwo'ole, and a clip of pink noise. A silent 4-minute audio file was created for use with the passive control group, and a short period of silence (15 seconds) was added before and after “Somewhere Over the Rainbow” to bring the total run time to four minutes. All sound files were played via Groove Media Player software.

Design and Procedure

Participants were tested individually, and each participant was tested only once. Participants were asked to silence and put away their phone, and then researchers collected demographic data via a demographic questionnaire (Appendix D) and administered the first PANAS. Afterwards, researchers explained to participants how the cognitive depletion task would work and were told that they would receive reward pieces (individually-wrapped candy pieces or small raisin boxes) at the end of the study. Participants received one reward piece for every 10 math problems correctly completed, and an additional piece if they were able to repeat the memorized number after the math portion. This incentive structure helped focus participants’ attention on both retaining the memorized number in memory while undergoing the more complex mental task of transforming quantities and problem solving. The intent of this task was to induce cognitive fatigue and decrease working memory capacity.

After the cognitive depletion task, participants sat at a desk with a laptop with connected headphones, put the headphones on, and began the first DSB set. Upon completion, participants listened to their assigned sound clip. Participants were instructed to close their eyes and focus on the sound they heard (or the silence, for the passive control group). After the clip was over, participants took the DSB test one more time.

After the second DSB set was complete, participants completed the GEMS-45 and the PANAS again and were given the correct amount of candy pieces or raisin boxes, according to how well they performed on the cognitive depletion task.

Results

Although Woods and colleagues (2011) found that ML and MS were better predictors of many aspects of working memory, we used and tested all four DSB metrics to see if any differences in patterns emerged in the data. We also tested both Affect metrics to further elucidate findings. All tests were performed on SPSS, version 25. For metric names and summary, see Table 1.

Table 1

Descriptive Statistics for DSB and Affect metrics

Metric Name	**Before: *M*(*SD*)**	**After: *M(SD)***	***N***
TE_ML	5.25 (1.24)	5.59 (1.16)	116
TE_TT	4.78 (1.62)	4.94 (1.34)	116
ML	5.9 (1.06)	6.13 (1.12)	116
MS	5.40 (0.95)	5.73 (1.14)	116
PosAff	2.74 (0.76)	2.21 (0.81)	115
NegAff	1.39 (0.36)	1.38 (0.47)	114

Hypothesis 1

To test our first hypothesis—that bird calls will have a greater positive effect on restoration than music—we ran mixed ANOVAs for every variable of interest (see Table 2). None of the metrics showed a statistically significant change; in fact, some F values were far below 1.00, ML: $F(3, 112) = .7$; MS: $F(3, 112) = .32$; NegAff: $F(3, 110) = .34$. This occurs when the estimation of data variance based on the within-group variance is much larger than the estimation of variance based on the between-group variance, and suggests that some unknown effect was interfering with either the data collection or participant performance.

Table 2

Mixed ANOVA table for DSB and Affect metrics

Source	***SS***	***df***	***MS***	***F***	***p***
TE_ML					
Pre-Post	6.90	1	6.90	9.96	*
Group	6.98	3	2.33	1.07	.37
Error (Group)	245	112	2.18		
TE_TT					
Pre-Post	1.56	1	1.56	1.14	.29
Group	12.7	3	4.22	1.38	.25
Error (Group)	342	112	3.06		
ML					
Pre-Post	3.14	1	3.14	9.74	*
Group	4.36	3	1.45	.70	.55
Error (Group)	231	112	2.06		
MS					
Pre-Post	6.64	1	6.64	23.1	**
Group	1.85	3	.62	.32	.81
Error (Group)	218	112	1.95		
PosAff					
Pre-Post	15.9	1	15.9	77.3	**
Group	3.12	3	1.04	1.02	.39
Error (Group)	113	111	1.02		
NegAff					
Pre-Post	< .01	1	< .01	.01	.92
Group	.26	3	.090	.34	.80
Error (Group)	28.0	110	.26		

* $p < .01$, ** $p < .001$

As part of an exploratory post hoc analysis, we looked at effect size and observed statistical power. To test whether our study was underpowered, we ran effect size and power analyses (see Table 3). We found that η_p^2 for the DSB metrics were all small to medium (M = 0.02, range 0.008 – 0.036; see Table 3). According to Cohen (1973)**,** a small effect size is 0.01 and a medium is 0.06.

Table 3

Effect Size and Observed Power for between-group DSB and Affect metrics

Source	**η_p^2**	**Observed Power**
TE_ML	.028	.28
TE_TT	.036	.36
ML	.019	.20
MS	.008	.11
PosAff	.027	.27
NegAff	.009	.11

Observed statistical power for the DSB metrics was also low (M = 0.24, range 0.11-0.36). Because of the very low statistical power, we examined the means of the difference scores for all DSB and both Affect metrics to see whether the data offered any possible support for our hypothesis.

An examination of the means for the change of DSB performance showed that they were sometimes in the hypothesized direction (see Figure 1 and Table 4). Specifically, participants in the Bird Calls condition showed the greatest improvement in the Two-Error metrics, TE_ML = +0.48; TE_TT = +.38. However, the reverse was true for the Overall metric category, where the Song condition was associated with equal (ML = +.24) or superior improvement (MS = +0.39) to the Bird Calls. Bird Calls were also associated with the greatest drop in PosAff (PosAff = -0.68) and NegAff (NegAff = -0.031), whereas the Song condition was associated with the smallest drop in PosAff (PosAff = -0.35) and the second smallest in NegAff (NegAff = -0.003; see Table 4).

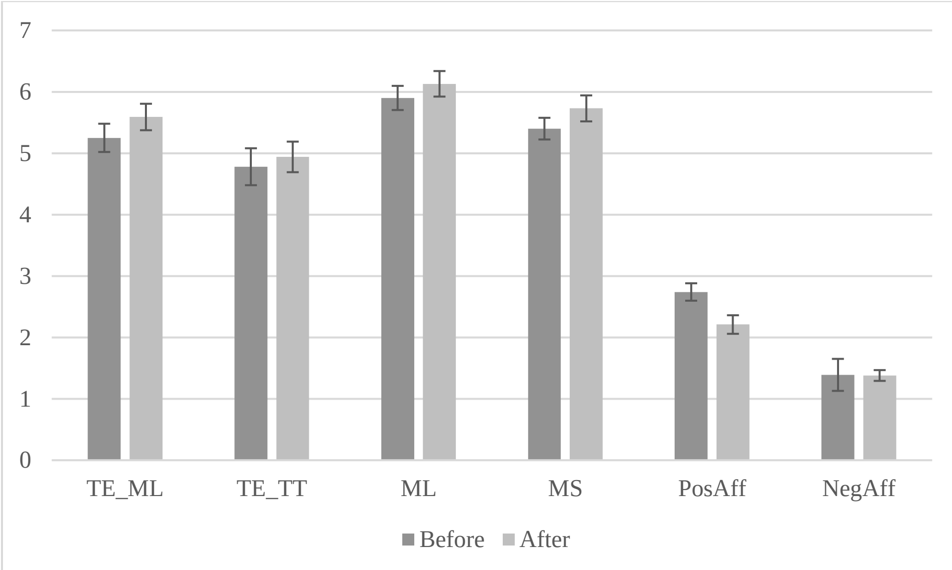

Figure 1. Mean scores on DSB and Affect measures before and after sound intervention.

Error bars represent +/- 2 standard errors of the mean.

Table 4

Mean differences: DSB and Affect Metrics vs. Average

Group	Two-Error		Overall		Affect	
	TE_ML	TE_TT	ML	MS	PosAff	NegAff
Bird Call	0.483	0.379	0.241	0.327	-0.676	-0.031
Song	0.379	0.276	0.241	0.39	-0.345	-0.003
Pink Noise	0.207	-0.035	0.138	0.352	-0.554	0.004
Silence	0.31	0.035	0.31	0.284	-0.528	0.014
Average	0.345	0.164	0.233	0.338	-0.525	-0.004

Note: Cells are shaded to reflect average size difference. White cells are the highest scores; dark grey are the lowest. The range for shading is specific to each metric.

Hypothesis 2

To test our second hypothesis—that a positive affective response to sound, as measured by the GEMS-45, mediates attentional restoration—we ran serial mediation analyses using ordinary least squares path analysis as recommended by Hayes (2013). Each of the three composite GEMs subscales (Sublimity, Vitality, and Unease) were tested as mediators for each of the four restoration difference scores (TE_ML, TE_TT, ML, and MS).

There was no evidence of mediation, regardless of the valence of the scale (Positive valence: Sublimity and Vitality; Negative valence: Unease).

To further explore our second hypothesis, we ran bivariate correlation analyses for the difference scores of all metrics and GEMS subscale scores. Of the DSB metrics, only one—TE_ML—showed a significant correlation, with the Power subscale, $r = -.19$, $p = .047$ (see Table 5). No other DSB metrics showed any statistically significant correlation with any GEMS scales, and when we divide the data according to group, the correlation disappeared for all but the Pink Noise group, $r = -.39$, $p = .044$. Therefore, it is likely that this result is due to chance.

Of the PANAS subscales, there were weak significant correlations observed between both PosAff and NegAff and the Peace subscale (PosAff, $r = .225$, $p = .017$; NegAff, $r = -.258$, $p = .006$), the Tension subscale (PosAff, $r = -.432$, $p < .001$; NegAff, $r = .385$, $p < .001$), and the Unease subscale (PosAff, $r = -.387$, $p < .001$; NegAff, $r = .414$, $p < .001$; see Table 5).

Table 5

Relevant Correlations between metrics and GEMS-45 subscales

	Power	Peacefulness	Tension	Unease
TE_ML	-0.188*	--	--	--
PosAff	--	0.225*	-0.432**	-0.387**
NegAff	--	-0.258**	0.385**	0.414**

* $p < .05$, ** $p < .01$

Discussion

Previous research has suggested that listening to calming sounds—both natural and man-made—can have a positive restorative effect on direct attentional control (Abbott et al. 2016; Baldwin and Lewis 2017; Haga, Holmgren, and Sörqvist 2016; Ratcliffe, Gatersleben, and Sowden 2013, 2016), but there are different theories as to the exact mechanisms this restoration operates by (cf. Kaplan 1995, Kaplan and Berman 2010 with Joye and van den berg 2011). This study found mixed support for the hypothesis that natural sounds have a more salubrious effect than anthropogenic sounds, and it replicated previous findings where both kinds of sounds can have some sort of beneficial effect.

The chief obstacle to interpreting our results is the lack of statistical significance between participant performance in the treatment and control groups. The low effect size and observed statistical power suggests that we should look to the direction of the means and see whether there is support for the hypotheses there; if so, further experimentation with a more refined study design would be warranted.

Bird Calls produced better participant performance for the Two-Error metrics, and contributed significantly to loss of participant affect (both positive and negative). Conversely, the Song seemed to mitigate loss of positive affect and was correlated with an equivalent or larger increase in performance for the overall metrics. Additionally, the Bird Call group was connected to the largest decrease of PosAff and the Song group had the smallest decrease. One possible explanation is that the song may have led to more relaxation among participants (evidenced by the smaller drop in PosAff), leading to greater performance on Overall metrics, but the Bird Calls put participants “on edge” (more on this below), leading to greater initial performance that quickly drops off. However, a study with higher power would need to be run before stronger conclusions can be drawn about what the data truly suggests. Based on the pattern of the differences of the means, we believe another study would be warranted.

Our second hypothesis stated that there would be a positive correlation between a participant’s positively-valenced emotional response to music and the degree of restoration experienced. We ran mediation analyses to test for relationships between GEMS scores and changes in DSB performance, but we found no evidence of mediation. Our post hoc exploratory analyses, consisting of running bivariate correlations between all six metrics and every GEMS subscale and second-order scale, found little evidence that there were any correlations between the metrics and GEMS scores (Appendix E). It is possible that the one statistically significant correlation we found between a DSB measure and GEMS subscale (TE_ML and Power; r = -0.19) was due to chance. We did find a possible explanation for the Bird Call group's large increase in Two-Error Metrics: the strong negative correlation between PosAff and the Tension (r = -0.43, $p < .01$) and Unease (r = -0.39, $p < .01$) subscales on the GEMS-45 suggests that it is possible that the Bird Call intervention put participants on edge, and this state of irritated alertness contributed to better initial performance on the second DSB and a larger drop in PosAff.

However, the low statistical power of the study makes the correlations somewhat difficult to interpret. A few correlations were marginally significant; if we increase alpha to .10, we begin to find more correlations; to .2, and we find that five correlations between a DSB measure and GEMS subscale emerge (all listed alpha values are for 2-tailed tests). Increasing statistical power may result in these and other correlations emerging as statistically significant.

Limitations

As mentioned before, the primary limitations of this study were low statistical power; the low effect size contributed significantly to the low power, as did the relatively small group size (n = 29). Because different statistics had different effect sizes, the a posteriori power estimate was different for each. Because a good study would include high power for each metric of interest, and the lowest power level was for MS (power = .11), future studies should calibrate their designs so as to be high-powered (with a power level of .80 or more; Cohen 1992) in reference to MS. Future studies exploring this topic should aim to increase effect sizes by inducing greater cognitive fatigue before the first DSB administration, increasing the length (and therefore efficacy) of the sound intervention, or both.

Beyond that: the significant reduction in positive affect associated with the Bird Call condition suggests that something about the sound clip used may be off-putting to participants. (Support for this theory comes from the fact that, of all the sound conditions, the correlation between the Bird Call condition and the Tension and Unease subscales was the strongest; see A-2). It is possible this was due in part to the fact we used isolated bird call sounds—and not a full soundscape—for the sound clip. Future studies should see whether there is a difference in restorative power between isolated nature sounds (such as bird calls) and full soundscapes (e.g. “sounds of the forest”).

References

Abbott, L. C., Taff, D., Newman, P., Benfield, J. A., & Mowen, A. J. (2016). The Influence of Natural Sounds on Attention Restoration. *Journal of Park & Recreation Administration*, *34*(3), 5–15.

Andrews, M., & Gatersleben, B. (2010). Variations in perceptions of danger, fear and preference in a simulated natural environment. *Journal of Environmental Psychology*, *30*(4), 473–481. https://doi.org/10.1016/j.jenvp.2010.04.001

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *49A*(1), 5–28. https://doi.org/10.1080/027249896392784

Baddeley, A. D. (2001). Is working memory still working? *American Psychologist*, *56*(11), 851–864. https://doi.org/10.1037/0003-066X.56.11.851

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *Recent Advances in Learning and Motivation* (Vol. 8, pp. 47-89). New York: Academic Press.

Baldwin, C. L., & Lewis, B. A. (2017). Positive valence music restores executive control over sustained attention. *PLoS ONE*, *12*(11), 1–11. https://doi.org/10.1371/journal.pone.0186231

Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and Psychological Measurement*, 33(1), 107–112.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159. http://dx.doi.org/10.1037/0033-2909.112.1.155

Engle, R. W. (2002). Working Memory Capacity as Executive Attention. *Current Directions in Psychological Science*, *11*(1), 19–23.

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience*, *14*(3), 340–347. https://doi.org/10.1162/089892902317361886

Frijda, N. H., & Sundararajan, L. (2007). Emotion refinement: A theory inspired by Chinese poetics. *Perspectives on Psychological Science*, *2*(3), 227–241. https://doi.org/10.1111/j.1745-6916.2007.00042.x

Haga, A., Halin, N., Holmgren, M., & Sörqvist, P. (2016). Psychological restoration can depend on stimulus-source attribution: A challenge for the evolutionary account? *Frontiers in Psychology*, *7*.

Herzog, T. R., & Rector, A. E. (2009). Perceived Danger and Judged Likelihood of Restoration. *Environment and Behavior*, *41*(3), 387–401. https://doi.org/10.1177/0013916508315351

Hubber, P. J., Gilmore, C., & Cragg, L. (2014). The Roles of the Central Executive and Visuospatial Storage in Mental Arithmetic: A Comparison across Strategies. *Quarterly Journal of Experimental Psychology*, *67*(5), 936–954. https://doi.org/10.1080/17470218.2013.838590

James, W. (1892). Psychology: The briefer course. New York: Holt.

Jefferies, L. N., Smilek, D., Eich, E., & Enns, J. T. (2008). Emotional valence and arousal interact in attentional control. *Psychological Science*, *19*(3), 290–295. https://doi.org/10.1111/j.1467-9280.2008.02082.x

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. https://doi.org/10.1037/xhp0000105

Joye, Y., & van den Berg, A. (2011). Is love for green in our genes? A critical analysis of evolutionary assumptions in restorative environments research. *Urban Forestry & Urban Greening*, *10*(4), 261–268. https://doi.org/10.1016/j.ufug.2011.07.004

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, *132*(1), 47–70. https://doi.org/10.1037/0096-3445.132.1.47

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

Kaplan, S., & Berman, M. G. (2010). Directed attention as a common resource for executive functioning and self-regulation. *Perspectives on Psychological Science*, *5*(1), 43–57. https://doi.org/10.1177/1745691609356784

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

Kuhbandner, C., & Zehetleitner, M. (2011). Dissociable effects of valence and arousal in adaptive executive control. *PLoS ONE*, *6*(12). https://doi.org/10.1371/journal.pone.0029287

Levinson, J. (1990). *Music, art, and metaphysics*. Cornell University Press.

Mattys, S. L., Baddeley, A., & Trenkic, D. (2018). Is the superior verbal memory span of Mandarin speakers due to faster rehearsal? *Memory & Cognition*, *46*(3), 361–369. https://doi.org/10.3758/s13421-017-0770-8

Ohly, H., White, M. P., Wheeler, B. W., Bethel, A., Ukoumunne, O. C., Nikolaou, V., & Garside, R. (2016). Attention Restoration Theory: A systematic review of the attention restoration potential of exposure to natural environments. *Journal of Toxicology & Environmental Health: Part B*, *19*(7), 305–343. https://doi.org/10.1080/10937404.2016.1196155

Otsuka, Y., & Osaka, N. (2015). High-performers use the phonological loop less to process mental arithmetic during working memory tasks. *Quarterly Journal of Experimental Psychology*, *68*(5), 878–886. https://doi.org/10.1080/17470218.2014.966728

Payne, S. R. (2013). The production of a Perceived Restorativeness Soundscape Scale. *Applied Acoustics*, *74*(2), 255–263. https://doi.org/10.1016/j.apacoust.2011.11.005

Pereira, C. S., Teixeira, J., Figueiredo, P., Xavier, J., Castro, S. L., & Brattico, E. (2011). Music and Emotions in the Brain: Familiarity Matters. *PLOS ONE*, *6*(11), e27241. https://doi.org/10.1371/journal.pone.0027241

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

Ratcliffe, E., Gatersleben, B., & Sowden, P. T. (2016). Associations with bird sounds: How do they relate to perceived restorative potential? *Journal of Environmental Psychology*, *47*, 136–144. https://doi.org/10.1016/j.jenvp.2016.05.009

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: is beauty in the perceiver's processing experience? *Personality And Social Psychology Review: An Official Journal Of The Society For Personality And Social Psychology, Inc*, *8*(4), 364–382.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology*, *81*(6), 989–1000. https://doi.org/10.1037/0022-3514.81.6.989

Woods, D. L., Kishiyama, M. M., Yund, E. W., Herron, T. J., Edwards, B., Poliva, O., … & Reed, B. (2011). Improving digit span assessment of short-term verbal memory. *Journal* *of Clinical and Experimental Neuropsychology*, *33*(1), 101–111. https://doi.org/10.1080/13803395.2010.493149

Zentner, M., & Eerola, T. (2011). Continuous self-report methods. In P. N. Juslin & J. A. Sloboda (Eds.), Handbook of Music and Emotion: Theory, Research, Applications (1st ed., p. xiv, 975 p. : ill.). Oxford: Oxford University Press.

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, *8*(4), 494–521. https://doi.org/10.1037/1528-3542.8.4.494

Appendix A

Participant ID: ____________________

Group Number: ____________________

52

Worksheet 3.1 The Positive and Negative Affect Schedule (PANAS; Watson et al., 1988)

PANAS Questionnaire

This scale consists of a number of words that describe different feelings and emotions. Read each item and then list the number from the scale below next to each word. **Indicate to what extent you feel this way right now, that is, at the present moment**

1	2	3	4	5
Very Slightly or Not at All	A Little	Moderately	Quite a Bit	Extremely

________ 1. Interested	________ 11. Irritable
________ 2. Distressed	________ 12. Alert
________ 3. Excited	________ 13. Ashamed
________ 4. Upset	________ 14. Inspired
________ 5. Strong	________ 15. Nervous
________ 6. Guilty	________ 16. Determined
________ 7. Scared	________ 17. Attentive
________ 8. Hostile	________ 18. Jittery
________ 9. Enthusiastic	________ 19. Active
________ 10. Proud	________ 20. Afraid

Appendix B

Participant ID:_________________
Group Number: ________________

GEMS-45 Scale & Instructions

Instructions

When providing your ratings, please describe how the music you listen to makes you *feel* (e.g., this music makes me *feel* sad). Do not describe the music (e.g., this music is sad) or what the music may be expressive of (e.g. this music expresses sadness). Keep in mind that sometimes a piece of music can be sad or can sound sad without making you feel sad. Please rate the intensity with which you *felt* each of the following feelings on a scale ranging from 1 (*not at all*) to 5 (*very much*). If you should find that one or more important labels are missing from the list to describe what you feel, please add those at the end and give it a rating (NN).

1	**2**	**3**	**4**	**5**
Not at all	Somewhat	Moderately	Quite a lot	Very Much

1. ______ Moved
2. ______ Fascinated
3. ______ Strong
4. ______ Tender
5. ______ Nostalgic
6. ______ Serene
7. ______ Animated
8. ______ Sad
9. ______ Tense
10. ______ Bouncy
11. ______ Filled with Wonder
12. ______ Sentimental
13. ______ Affectionate
14. ______ Overwhelmed
15. ______ Agitated
16. ______ Feeling of Transcendence
17. ______ Calm
18. ______ Joyful
19. ______ Tearful
20. ______ Soothed
21. ______ Energetic
22. ______ Dreamy
23. ______ Mellowed (Softened-up)
24. ______ Allured
25. ______ Triumphant
26. ______ Happy
27. ______ Inspired
28. ______ In love
29. ______ Melancholic
30. ______ Fiery
31. ______ Meditative
32. ______ Feel like dancing
33. ______ Dazzled
34. ______ Sorrowful
35. ______ Chills (shivers & goose bumps)
36. ______ Irritated
37. ______ Amused
38. ______ Heroic
39. ______ Relaxed
40. ______ Feeling of Spirituality
41. ______ Sensual
42. ______ Stimulated
43. ______ Nervous
44. ______ Admiring
45. ______ Impatient

(X.)_______Blue
(Y.)_______Amazed
(Z.)_______In awe
NN ______

Appendix C

For the following problems, solve to the best of your ability and write the answer next to the equation. Do not use a calculator.

1. 5 + 1 =
2. 8 – 9=
3. 2 + 4 =
4. 0 + 95 – 2 =
5. 3 – 0 =
6. 5 – 3 =
7. 5 / 20 =
8. 82 – 7 =
9. 1 + 9 =
10. 100 / 4 =
11. 11 + 29 =
12. 24 + 39 =
13. 504 / 9 =
14. 63 / 9 + 8 =
15. 36 – 54 =
16. 53 + 48 – 100 =
17. 94 + 63 =
18. 56 – 28.5 =
19. 104 – 83 =
20. 10.25 x 8 =
21. 200 + 401 =
22. 92 – 4 =
23. 12 x 6 =
24. 17 x 6 =
25. 2 x 93 =
26. (45 – 58) / 2 =
27. 56 + 8,529 =
28. 320 – 9,853 =
29. 0 x 95 =
30. 600 / 12 =
31. 649 + 534 – 78 =
32. 13 / 2 =
33. 321 – 956 – 260 + 48 =
34. 613 + 95 =
35. 14 + 18 – 9 =
36. (315 / 15) + 46 =
37. 5,296.7 + 724.8 =
38. (15 / 2) + 3 x 4 =
39. 15 + 4 =
40. 52 + 648 – 995 – 20 + 563 =
41. 235 x 6 =
42. 47 + 6 =
43. 25 x 60 =
44. 1,000 / 25 =
45. (8,765 – 8764) / 0 =
46. 16 – 53 =
47. 48,956 + 569,000 =
48. 98 – 96 / 2 =
49. 5 + 63 =
50. 7 / 7 =
51. 230 / 5 =
52. 54,639 + 978 =
53. 42 + 7 x 16 / 2 =
54. 3,400 x 16 =
55. 45 / 5 x 31 =
56. 87 / 3 =
57. 25 – 20 =
58. 75 – 2 =
59. 58 x (0 + 65) =
60. 32 x 20 =

For the following section, you may use a calculator. Write the answer next to the equation. For all questions, solve for the variable and simplify as much as possible.

61. 90x = 78
    a. x =
62. 20 + 4x = 10
    a. x =
63. 6x / 8 = 7.5
    a. x =
64. 2y + 5x = 68   and   y + x = 16
    a. x =
    b. y =
65. xy = 10   and   x + y = 6.5
    a. x =

    b. y =

66. 98 = 49x
    a. x =
67. $\sqrt{x} + 24y = 57$ and $y - 3 = -1$
    a. x =
    b. y =
68. $z + x + 14 = 37$ and $zx = 132$
    a. x =
    b. z =
69. $18 / x = y - 3$ and $y + 2x = 23$
    a. x =
    b. y =
70. $14 + 2x = 10$
    a. x =
71. $4x + 2y + 3z = 43$ and $3x = 9$ and $z – y = 2$
    a. x =
    b. y =
    c. z =
72. $x^2 + 3 = 84$
    a. x =
73. $502 – (\sqrt{x} + y^2) = 473$
    a. x =
    b. y =
74. $3x – 4y + 20 = 12$ and $4x + 3y = 31$
    a. x =
    b. y =
75. $72x – 4y = 80$ and $[y] = 16$
    a. x =
    b. y =

Appendix D

Demographic Questionnaire

Please answer the following questions.

1. My sex is _______
    1) Male
    2) Female
    3) Transgender, Genderqueer/Gender-fluid, or Questioning/Unsure
    4) Write In: ______________________________
    5) Prefer not to say

2. I am ________ years old.

Appendix E

Table A1

Full Correlation matrix for all metrics and GEMS-45 scales

	TE_ML	TE_TT	ML	MS	PosAff	NegAff	Wonder	Transcendence	Power
TE_ML	1.000								
TE_TT	.873**	1.000							
ML	.401**	.425**	1.000						
MS	.472**	.295**	.683**	1.000					
PosAff	-0.008	0.009	0.036	-0.044	1.000				
NegAff	-0.018	-0.041	-0.047	-0.001	-.217*	1.000			
Wonder	-0.071	-0.008	-0.076	-0.067	0.105	0.069	1.000		
Transcendence	0.012	0.041	-0.043	0.004	-0.058	0.148[A]	.821**	1.000	
Power	-.188*	-0.126[A]	-0.084	-0.086	-0.056	0.126	.693**	.662**	1.000
Tenderness	-0.026	0.013	-0.027	-0.057	0.114	-0.087	.839**	.680**	.483**
Nostalgia	-0.067	-0.003	-0.115	-0.139[A]	0.103	-0.070	.746**	.626**	.380**
Peacefulness	0.067	0.077	-0.101	-0.100	.225*	-.258**	.546**	.324**	0.120
JoyfulActivation	-0.061	0.013	-0.019	-0.028	0.044	0.082	.821**	.687**	.750**
Sadness	-0.084	-0.086	0.043	-0.002	-0.001	0.162[A]	.326**	.426**	.283**
Tension	0.032	0.007	0.038	0.068	-.432**	.385**	-.235*	0.013	-0.051
Sublimity	-0.019	0.026	-0.086	-0.088	0.130[A]	-0.065	.922**	.775**	.532**
Vitality	-0.128[A]	-0.055	-0.049	-0.057	0.001	0.108	.817**	.725**	.915**
Unease	-0.007	-0.030	0.052	0.060	-.387**	.414**	-0.072	.192*	0.075

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

[A]. Correlation is significant at the 0.18 level (2-tailed).

	Tenderness	Nostalgia	Peacefulness	Joyful Activation	Sadness	Tension	Sublimity	Vitality	Unease
TE_ML									
TE_TT									
ML									
MS									
PosAff									
NegAff									
Wonder									
Transcendence									
Power									
Tenderness	1.000								
Nostalgia	.802**	1.000							
Peacefulness	.637**	.648**	1.000						
JoyfulActivation	.685**	.594**	.376**	1.000					
Sadness	.434**	.344**	-0.018	.249**	1.000				
Tension	-.302**	-.248**	-.473**	-0.149	0.024	1.000			
Sublimity	.923**	.885**	.779**	.748**	.327**	-.320**	1.000		
Vitality	.639**	.536**	.283**	.953**	.283**	-0.114	.701**	1.000	
Unease	-0.086	-0.077	-.432**	-0.028	.445**	.906**	-0.148	0.017	1.000

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

A. Correlation is significant at the 0.18 level (2-tailed).