**Characteristics of fascination: Using eye-tracking to explore the impact of spatial frequency on the allocation of attention to nature and urban scenes.**

*Catherine Thompson1\*, Tobiasz Trawiński1, David Beevers2,*

*Neil Harrison1, and Nick Donnelly1*

1 Liverpool Hope University

2 University of Salford

\* Corresponding author

HCA EW113

Hope Park

Liverpool Hope University

Liverpool

L16 9JD

[thompsc1@hope.ac.uk](mailto:thompsc1@hope.ac.uk)

Word count: 9035

**Abstract**

Evidence suggests that nature environments capture attention effortlessly and/or are easy to process, but the mechanisms responsible for attention restoration are not fully understood. This study manipulated low-level properties of nature and urban images to measure their impact on gaze behaviour. Stimuli comprised 20 grayscale images (10 nature, 10 urban) shown in their original form, and with low or mid-to-high spatial frequencies filtered out. Eye movements were recorded whilst participants viewed the scenes, rating each one for pleasantness and ease of identifying objects. Participants made fewer, longer fixations to nature scenes, unless mid-to-high spatial frequencies were removed, and explored urban scenes more, unless low spatial frequencies were removed. Nature scenes were rated as more pleasant, providing mid-to-high spatial frequencies were present, and identification of objects was easier in urban scenes, particularly when mid-to-high spatial frequencies were removed. Further analysis revealed important findings to guide future research; gaze behaviour was highly consistent within participants, and differences in eye movements to urban and nature scenes were most prominent in the first 5-6 seconds of viewing. This shows the potential for using eye movements, especially early in viewing, to study individual differences in the perception of, and attention to, nature and urban environments.

**Keywords**

Attention restoration, perceptual fluency, eye-tracking, spatial frequency, gaze behaviour

**Introduction**

Attention restoration is the recovery of limited-capacity cognitive resources that have been depleted during the course of a demanding task. It falls under the wider concept of restoration, which encompasses the recovery of mental, physical, psychological, and even social resources (e.g., Hartig, 2021). A large body of evidence shows that attention restoration can be achieved via exposure to certain environmental contexts, specifically contexts that contain natural stimuli. Most of this research measures cognitive performance before and after engagement with nature compared to urban settings, and findings consistently show better cognitive performance following exposure to nature relative to urban environments. Attention restoration from nature has been shown across varying methods of engagement (e.g., viewing scenes and videos, walking in real environments, being in a virtual environment), for different durations of exposure (from minutes to hours), and for a range of different cognitive tasks. For instance, engagement with nature has been found to improve concentration (Tennessen & Cimprich, 1995), sustained attention (Berto, 2005; Lee et al., 2015), working memory (Berman et al., 2012), verbal working memory (Bratman et al., 2015), executive control (Berman et al., 2008; Taylor et al., 2002), and executive function (Bourrier et al., 2018).

The most prominent theory to account for such findings is the Attention Restoration Theory (ART; Kaplan & Kaplan, 1989; Kaplan, 1995). The theory outlines that focusing on a task utilises limited-capacity directed attention resources, resulting in attentional fatigue. Attention restoration occurs when directed attention resources can rest and recover, and this can be achieved in settings that occupy the use of indirected attention resources (Kaplan & Berman, 2010). The ART outlines four different features of restorative environments; being away (restorative settings are those which allow for a shift of focus from everyday thoughts), extent (restorative settings are those that occupy the mind sufficiently to allow resources to restore), compatibility (a restorative environment is one that aligns with an individual’s aims and intentions), and fascination (a restorative environment captures and holds attention in an effortless manner). Research has mostly focussed on the concept of fascination. Kaplan (1995) proposed that restorative settings feature stimuli that are “softly fascinating” in that they are interesting enough to elicit attention, but they are not demanding of attention so do not require the use of directed attention resources. In contrast, settings that are associated with “hard fascination” contain stimuli that capture attention in a demanding, effortful way, require the use of directed attention, and provide limited opportunity for restoration. The restorative benefits of nature on cognition have, therefore, been explained with the proposal that stimuli in nature settings capture attention “modestly”, whilst stimuli in urban settings capture attention “dramatically”.

The concept of soft fascination has been criticised, with researchers arguing that it is unclear what features make a setting softly fascinating or how much soft fascination is required to enable restoration (e.g., Joye & Dewitte, 2018; Schertz & Berman, 2019). An alternative explanation is the Perceptual Fluency Account (PFA; Joye & van den Berg, 2012; Joye et al., 2016). The PFA suggests that nature settings are perceptually fluent and are therefore easier to visually process than urban settings. This theory attributes the restoration of cognitive resources following engagement with nature stimuli to the ease with which nature environments can be processed. However, as with soft fascination, it is unclear what makes an environment perceptually fluent. One proposal is that whilst nature settings may contain the same amount of information and be just as visually complex as urban settings, they are more repetitive and predictable (Menzel & Reese, 2021). As such, nature scenes may have more “self-similarity” or “visual redundancy” than urban scenes (Bourrier et al., 2018; Joye et al., 2016).

The importance of self-similarity was investigated by Joye et al. (2016) using fractal patterns. In two experiments they asked participants to solve visual puzzles that were surrounded by (Experiment 1) or preceded by (Experiment 2) block patterns that were high or low in fractalness (therefore high or low in self-similarity). Performance on the puzzle task was better when paired with the high fractal patterns. Joye and colleagues argued that the high fractal patterns were more perceptually fluent, leaving resources to be focussed on the pattern task.

Although Joye et al. (2016) did not directly assess the impact of nature and urban stimuli on the visual puzzle task performance, their study offers a different method to those used previously for investigating attention restoration. The focus of past research has predominantly been on comparing cognition before and after exposure to/engagement with natural versus urban environments, rather than behaviour and processing within the environments themselves (i.e., *during* the recovery period). Fleming et al. (2024) claim that this style of pre-post design does not allow for the investigation of the underlying mechanisms responsible for attention restoration. Given that it is currently unclear what contextual features lead to restoration, measuring the allocation of attention whilst observers engage with nature and urban stimuli offers an alternative approach to this field of research.

Eich and Beck (2023) investigated attention whilst participants were viewing nature and urban scenes. They used the rapid serial visual presentation paradigm (RSVP) to measure the Attentional Blink (AB) whilst participants searched for two targets (coloured nature or urban scenes) amongst non-targets (grayscale nature or urban scenes). The images were shown for 140ms each, within each trial the images were either all nature or all urban, and the second target (T2) was presented at differing temporal lags following the first target (T1). The magnitude of the AB effect is thought to reflect the investment of resources to the first target. For instance, Chun and Potter (1995) proposed that T2 performance suffers at early lags because resources are devoted to processing T1, leaving insufficient resources to identify T2. Recovery from the AB occurs at later lags because T1 has been processed, allowing resources to be used for monitoring the RSVP for T2. Consequently, the AB may be a useful method for measuring differences in the allocation of attention in nature and urban settings. Eich and Beck found that the AB was longer, and the magnitude of the AB was greater during presentation of nature scenes. This finding initially appears to contradict earlier findings of attention restoration. However, they explained this result by suggesting that it is easier to disengage attention from urban scenes. They concluded that resources are allocated more effectively in an urban than nature RSVP stream.

Another approach to investigating attention whilst viewing nature and urban scenes is to use eye-tracking. In the context of studies contrasting nature and urban settings, the spatial and temporal distribution of eye movements provides a direct and on-line measure of attention (Schertz & Berman, 2019). An early study by Berto et al. (2008) compared eye movements to photographs of nature and urban scenes that had been rated as high and low on fascination. They found significantly more fixations for the low fascination (urban) scenes and suggested this reflected the effortful capture of attention by stimuli in urban settings. Franěk et al. (2019) compared eye movements to urban and nature scenes with and without foliage (therefore manipulating the fractal complexity of the nature scenes). They showed that scenes with foliage (and therefore a higher fractal dimension) were associated with fewer and longer fixations. More recently, Batool et al. (2022) found that relative to nature scenes, viewing urban scenes led to more fixations, shorter fixations, and more saccades.

Fractal complexity is a low-level scene characteristic and the theories to account for attention restoration place significant importance on such properties. For instance, the ART predicts that restorative environments contain features that capture attention in a bottom-up manner, allowing top-down resources to restore. It has been proposed that low-level visual differences between nature and urban settings play a role in attention restoration and this has been supported by research. For example. Berman et al. (2014) found that scenes rated as more natural (compared to those rated as more man-made) were characterised as having increased edges, fewer straight lines, and less hue diversity. This was supported by Kardan et al. (2015) who asked participants to rate over 300 scenes for naturalness and aesthetic preference and found that edge density was important in naturalness ratings, and that reduced hue was associated with higher preference ratings. They suggested that preference for nature scenes is driven by the initial bottom-up processing of basic scene properties before semantic knowledge and experience have an influence.

Gaze behaviour to nature and urban scenes has also been explored using scenes that have been filtered to modify their low-level features. Low-level features (e.g., spatial frequencies, edges, colour) are processed early in the visual system and support the processing of scene gist (e.g., Torralba & Oliva, 2003). Valtchanov and Ellard (2015) measured the impact of low-level scene properties by measuring eye movements whilst participants viewed images that had been filtered to modify their spatial frequency (important for scene recognition (e.g., Baddeley, 1997; Berman et al., 2017; Oliva et al., 1999)). They presented observers with four nature and four urban images in their original grayscale form and filtered such that only the mid-to-high or low spatial frequencies were present in the images. Each image was presented for 15-seconds and participants were asked to rate how pleasant each image was and how easy it was to identify items within the images. Consistent with Berto et al. (2008), they found that participants made more frequent but shorter fixations when viewing urban scenes than when viewing nature scenes. Participants also made more, and shorter, fixations in mid-high spatial frequency images compared to low spatial frequency images. Analysis of the image ratings showed that nature scenes were rated as more pleasant than urban scenes, but only in original images and images with mid-to-high spatial frequencies intact. From this they suggested that the presence of mid-to-high spatial frequencies are an important component in the affective benefits gained from nature scenes (e.g., Ulrich et al., 1991)[[1]](#footnote-1).

There are now increasing efforts to investigate, understand, and categorise the underlying mechanisms responsible for attention restoration, however this work is still in its infancy and warrants further exploration. For example, Valtchanov and Ellard (2015) used a very small set of images, and it is important to discover whether their effects hold across a wider variety of experimental settings. Many studies that compared eye movements in nature and urban settings used a viewing time of 15-seconds (e.g., Berto et al., 2008; Franěk et al., 2018; Valtchanov & Ellard, 2015; Wu et al., 2021) but it has been argued that presenting scenes for this long may lead to irrelevant fixations which then impact on attention and restoration (Wu et al., 2021). Exploring variations in gaze behaviour over time may provide insight about the allocation of attention in nature and urban settings, and the present study addresses this by examining how differences in eye movements to nature and urban scenes change over the time-course of viewing.

Additionally, although comparing the average of a group of eye movements to nature and urban images is important, it is possible that reliable individual differences exist in gaze behaviour across different types of scenes. Research has shown that visual attention and search is impacted by individual differences. For instance, in a free-viewing task, Risko et al. (2012) found that higher levels of perceptual curiosity correlated with increased visual exploration. Studies using the change blindness paradigm have also found effects of neuroticism and extraversion on visual attention (e.g., Bendall et al., 2021; Hahn et al., 2015). More recently, Grady et al. (2022) found that higher levels of conscientiousness were associated with improved visual search performance in an online baggage search task, and that performance for high conscientious individuals was not impacted by increased levels of mental fatigue. In addition, extent of engagement with a scene also impacts attention. Duvall (2011) found evidence for improved attention when participants were more engaged when walking in an outdoor setting. A study by Harrison (2023) also showed that connectedness to nature impacted the amount of attention given to nature scenes and also affected aesthetic responses to those scenes. If there are differences in how individuals view and engage with nature and urban settings, this may have implications for attention restoration.

The significance of individual variation in eye movement behaviour, which would be an important precursor to studying individual differences in attentional restoration, is suggested in a study reported by De Haas et al. (2019). They analysed eye movements from over 100 observers viewing 700 images and found a high level of consistency in fixation behaviour within observers across images. Measuring consistency and reliability of visual exploration within and across participants might reveal the potential for studying individual differences in eye movement behaviour to nature and urban scenes. Moreover, comparing individual differences in eye movements across scenes that have been filtered to modify their low-level scene properties may provide an important index of the relative contribution of low-level versus high-level processing in the attention restoration effect (see Batool et al., 2022). The current study examines both the reliability of individual differences in eye movements to scenes, and how these are impacted by image filtering.

In sum, the current study aimed to further explore the underlying mechanisms responsible for attention restoration by measuring gaze behaviour to nature and urban scenes. Participants were asked to view nature and urban images in their original format, and with mid-high and low spatial frequencies removed, and rate each image for pleasantness and ease of object identification. The first objective was to assess the consistency and reliability of individual differences in eye movements to nature and urban scenes. This would show the extent to which eye movement patterns are similar within participants across different scenes. A second objective was to measure differences in eye movement behaviours to urban and nature scenes when in their original form or spatial frequency filtered, extending previous findings. It was predicted that gaze behaviour would be significantly different between nature and urban scenes, with more fixations and shorter fixations to urban scenes. It was also predicted that there would be greater exploration on urban scenes compared to nature scenes. Following Valtchanov and Ellard (2015), manipulations to spatial frequency were expected to change this pattern. It was also predicted that participants would rate nature images as more pleasant than urban images and easier to identify objects, but that again, manipulations to spatial frequency would impact the ratings. A third objective was to explore the time course of the eye movement differences found when viewing nature and urban scenes. It was predicted that eye movements would change over the course of scene viewing, but with limited past research to indicate a specific trend, this analysis was largely exploratory.

**Method**

***Design***

A 2x3 within-participants design was used with two independent variables, *image type* (nature and urban) and *image manipulation* (original, mid-high spatial frequency, and low spatial frequency). The dependent variables consisted of three eye-movement metrics including number of fixations, mean fixation duration, and saccade length, and two image ratings (pleasantness and ease of object identification). A further independent variable of *eye movement epoch* was included following the main investigation, to allow comparison of viewing behaviour over time. Eye movements across the 15-seconds of viewing time on each image were broken down into epochs on the basis of temporal location (e.g., Epoch 1 consisted of fixations 1-10 on an image, Epoch 2 consisted of fixations 11-20, etc.). Full ethical approval was obtained from the School of Health Sciences Research Ethics Panel at the University of Salford.

***Measures***

The primary measure of eye movement behaviour was fixations (x and y coordinates and fixation duration). Three separate fixation measures were calculated: the number of fixations made to each image, the mean fixation duration on each image (in milliseconds), and the distance between fixations (saccade length). Following the method of Valtchanov and Ellard (2015), the total (summed) saccade length was also calculated for the full 15-seconds of scene viewing by taking the sum of the Euclidean distances between each fixation in an image and this was measured in pixels. This provided a measure of exploration (greater distance travelled indicating greater exploration). However, when measuring eye movement patterns over time with fixations separated into epochs, mean saccade length was used, rather than total saccade length, to control for the number of fixations and the duration of fixations within each epoch.

Two further measures of how pleasant participants found each image and how easy it was to identify objects within each image were also calculated. These were operationalised as “pleasantness” and “ease of object identification” and the two rating types were compared across the conditions of image type and image manipulation and correlated with the eye movement metrics.

***Participants***

The participants were an opportunity sample of 40 staff and students at the University of Salford who responded to invitations to take part. All participants had normal or corrected-to-normal vision. All were compensated for their participation (£5). Age ranged from 18 to 64 with a mean of 28.87 years (SD = 12.54). Fifteen participants identified as male, 24 as female, and one as non-binary.

***Stimuli and Apparatus***

Twenty images of outdoor scenes (10 nature and 10 urban) were chosen, some taken by the authors, and some obtained from royalty-free image directories. The nature images included nine woodland scenes and one lakeside mountain image. The urban scenes were all city or town centre images. The 20 images were converted to grayscale and resized to 900 x 900 pixels using Microsoft Photo Editor. The lack of colour elements in the images ensured that both nature and urban scenes were neutral and not influenced by colour preferences. Each image was manipulated to create a further 40 pictures, 20 mid-high spatial frequency images (with low spatial frequencies removed) and 20 low-spatial frequency images (with mid-high spatial frequencies removed). The manipulated images (mid-high and low spatial frequency images) were created using the method of Valtchanov and Ellard (2015)[[2]](#footnote-2). See Figure 1 for examples of the images used. Visual contrast and complexity were compared for the original grayscale urban and nature scenes using the imagefluency package (Mayer & Landwehr, 2018). Urban images were more complex (*t* (18) = -3.81, *p =* .001, *Cohen’s d =* -1.71) and had a higher amount of contrast than nature images (*t* (18) = -3.44, *p =* .003, *Cohen’s d =* -1.54).

The experiment was designed and conducted on a Tobii Pro Spectrum with a screen size of 531mm x 298mm (aspect ratio 16:9), a screen resolution of 1920 x 1080 pixels, and a sampling rate of 120Hz. Tobii Pro Lab gathered raw fixation data using a velocity-based algorithm. The minimum fixation duration was 100ms and the minimum fixation dispersion threshold was 100 pixels. A chin rest was used to minimise head movements and maintain a set distance from the screen.

***Procedure***

Participants were given full instructions about the experiment and were asked to sit with their chin in the chin rest 55cm from the screen. Eye movements were calibrated using nine targets, and then validated using four targets. Participants were then presented with the 60 images, one after another, and were asked to rate each image. In each trial a fixation cross was presented for one second, an image was then presented for 15 seconds, and participants were instructed to view the image as they would normally view a photograph. Images were presented to the centre of the screen, subtending 25.5° of visual angle. Following this, a screen appeared asking participants “How pleasant did you find this image?” and they responded by pressing keys on the keyboard from 1 (extremely unpleasant) to 5 (extremely pleasant). A second screen was then presented asking participants “How well could you identify types of objects in the image?” with the instruction to respond using the keyboard on a scale of 1 (not at all) to 5 (very well). All images were presented in a random order, and the experiment took approximately 20 minutes to complete.

***Data Analysis***

The data collected consisted of eye movements and image ratings (pleasantness and ease of identification) for all 60 photographs. Forty participants completed the experiment; however, analysis was conducted on 38 participants. One participant was discarded due to a recording error, and data from a second participant was removed as they were classed as an outlier, with most of their ratings and eye movements falling more than 2.5 standard deviations from the mean.

The reliability of individual differences in gaze behaviour was tested across the three conditions of image manipulation. Using 10,000 random splits, the Spearman-Brown corrected reliability estimates (*r*-splits) for the number of fixations on original, mid-high, and low spatial frequency images were calculated with the associated 95% highest density intervals (Parsons, 2021). Higher estimated *r*-splits scores indicate greater consistency in the number of fixations made to the images. Using the traditional approach, all estimates above 0.80 were interpreted as highly reliable (Hedge et al., 2018). This analysis was conducted separately for nature and urban images.

The eye movements selected for the primary analysis were number of fixations, mean fixation duration, and total saccade length. Fixations shorter than 100ms and longer than 1000ms were removed (6.16% of all fixations). To measure the impact of the independent variables on gaze behaviour, a 2 (image type) x 3 (image manipulation) repeated-measures ANOVA was conducted for each of the three measures (number of fixations, mean fixation duration, and saccade length). For the variable of image manipulation, where sphericity was violated the Greenhouse-Geisser values are reported, unless this did not change the level of significance. Partial eta squared (*η2p*) was used as a measure of effect size in the ANOVA analyses. Any significant effects of image manipulation were explored using post-hoc comparisons with a Bonferroni correction. Where relevant, interactions between image type and image manipulation were analysed using post-hoc paired samples t-tests comparing nature and urban images within each condition of image manipulation and corrected for multiple comparisons. Effect sizes were calculated using Cohen’s d (*Cohen’s d*).

Following this, to explore gaze behaviour across time, each fixation made on each of the images was categorized into one of five fixation epochs: Epoch 1 (fixations 1–10), Epoch 2 (fixations 11–20), Epoch 3 (fixations 21–30), Epoch 4 (fixations 31–40), and Epoch 5 (fixation 40 and beyond). As the average fixation duration was 332ms, each epoch was approximately 3-seconds in length. A series of paired samples t-tests were conducted to compare gaze behaviour (mean fixation duration and mean saccade length) between nature and urban images. These tests were performed separately for each of the five epochs, for each condition, and *p*-values were adjusted for number of comparisons.

Two final 2 x 3 repeated measures ANOVAs were used to investigate the effect of image type and image manipulation on ratings of pleasantness and ease of identifying objects. Pearson’s Product Moment Correlations were used to assess the correlation between ratings of pleasantness and the eye movements metrics, but this analysis was only conducted for original images, not for images that had been filtered to manipulate spatial frequency. Effect sizes were calculated using R2.

**Results**

***Consistency and reliability of eye movements***

When analysing the reliability of fixation behaviour in the nature scenes, the Spearman-Brown corrected reliability estimate for the number of fixations to the original, mid-high, and low spatial frequency images was 0.93, 95% CI [0.89, 0.96], 0.92, 95% CI [0.88, 0.95], and 0.93, 95% CI [0.89, 0.96] respectively. For the urban scenes, the Spearman-Brown corrected reliability estimate for the number of fixations to the original, mid-high, and low spatial frequency images was 0.95, 95% CI [0.93, 0.97], 0.91, 95% CI [0.86, 0.95], and 0.94, 95% CI [0.90, 0.96] respectively. The data suggest a high level of consistency in the number of fixations made by participants to exemplars of both nature and urban scenes across each of the three conditions of image manipulation (see Figure 2).

***Effects of image type and image manipulation on eye movements***

*Number of fixations:*

Across the full 15-seconds of image viewing, participants made significantly more fixations to urban scenes (*M =* 32.89) than to nature scenes (*M =* 29.91), *F* (1,37) = 29.117, *MSE =* 17.495, *p <* .001, *η2p* = 0.440. A significant effect of image manipulation, *F* (2,74) = 45.603, *MSE =* 27.590, *p <* .001, *η2p* = 0.552, showed that more fixations were made to original images (*M =* 34.58) than to mid-high (*p =* .005) and low (*p <* .001) spatial frequency images, and more fixations were made to mid-high spatial frequency images (*M =* 32.80) than to low spatial frequency images (*M =* 26.82, *p* < .001).

There was also a significant interaction between image type and image manipulation, *F* (2,74) = 15.250, *MSE =* 6.012, *p <* .001, *η2p* = 0.292, and this was a large effect. Paired samples t-tests showed that whilst more fixations were made to urban than to nature scenes in the original, *t* (37) = -6.119, *p* < .001, *Cohen’s d =* 0.99, and the mid-high spatial frequency images, *t* (37) = -3.474, *p* = .001, *Cohen’s d =* 0.56, when correcting for multiple comparisons there was no significant difference in number of fixations made to nature and urban scenes when the images only included low spatial frequencies, *t* (37) = -2.124, *p* = .040, *Cohen’s d =* 0.35. See Figure 3a.

*Mean fixation duration:*

Overall, across the full viewing time of each trial, participants made significantly longer fixations to nature scenes (*M =* 352.59ms) than to urban scenes (*M =* 334.39ms), *F* (1,37) = 39.325, *MSE =* 480.159, *p <* .001, *η2p* = 0.515. For mean fixation duration there was also a significant effect of image manipulation, *F* (2,74) = 47.956, *MSE =* 1276.450, *p <* .001, *η2p* = 0.564. The effects of image type and image manipulation on mean fixation duration were large. Post-hoc comparisons showed significantly shorter mean fixation durations to original images (*M =* 315.43ms) than to mid-high (*M =* 342.86ms, *p <* .001) and to low spatial frequency images (*M =* 372.18ms, *p <* .001), and significantly shorter mean fixation durations to mid-high than to low spatial frequency images (*p* < .001).

The interaction between image type and image manipulation was also significant, *F* (1.740,64.388) = 6.536, *MSE =* 512.979, *p =* .004, *η2p* = 0.150. Paired samples t-tests showed that mean fixation durations were significantly longer to nature than urban scenes in the original images, *t* (37) = 8.500, *p* < .001, *Cohen’s d =* 1.38, and in the mid-high spatial frequency images, *t* (37) = -2.609, *p* = .013, *Cohen’s d =* 0.42, but not when mid-high spatial frequencies were removed from the nature and urban scenes, *t* (37) = 1.793, *p* = .081, *Cohen’s d =* 0.29. See Figure 3b.

*Mean fixation duration across viewing time:*

When mean fixation duration was analysed across the viewing time, by breaking fixations into epochs, mean fixation durations were significantly longer to nature and urban scenes in the original images in the first *t* (993) = 2.160, *p* = .031, *Cohen’s d =* 1.53, and second epochs, *t* (993) = 2.483, *p* = .013, *Cohen’s d =* 1.76. In contrast, in the last two epochs of viewing, this pattern changed. Specifically, participants made longer fixations to urban than nature scenes in the low spatial frequency images in the fourth epoch, *t* (993) = -2.400, *p* = .017, *Cohen’s d =*-1.70, and similarly, participants made longer fixations to urban than nature scenes in the original images in the fifth (last) epoch, *t* (993) = -2.218, *p* = .027, *Cohen’s d =* -1.57. Again, these effect sizes were large, indicating a clear influence of the variables on mean fixation duration across the viewing time. No other comparisons were significant (all *p*'s > 0.06). These results demonstrate that longer fixations to nature scenes were only apparent within the first two epochs. In addition, this effect of image type was only observed in the original images (see Figure 4).

*Saccade length:*

The spatial extent to which participants explored each image was measured using saccade length. When considering the full viewing time, analysis showed a significant effect of image type with longer saccade length in urban scenes (*M =* 6223.99 pixels) compared to nature scenes (*M =* 5770.17 pixels), *F* (1,37) = 12.105, *MSE =* 969821.647, *p =* .001, *η2p* = 0.247. There was also a significant effect of image manipulation, *F* (2,74) = 37.695, *MSE =* 1588684.12, *p <* .001, *η2p* = 0.505. Post-hoc comparisons showed longer saccade length in original images (*M =* 6840.13 pixels) relative to mid-high spatial frequency (*M =* 6080.45 pixels) and low spatial frequency images (*M =* 5070.67 pixels), and longer saccade length in mid-high compared to low spatial frequency images (all *p’s* < .001).

The interaction between image type and image manipulation was also significant, *F* (2,74) = 10.264, *MSE =* 523121.488, *p <* .001, *η2p* = 0.217. Post-hoc t-tests showed significantly longer saccade length in urban than nature scenes when the images were in their original format, *t* (37) = -4.389, *p* < .001, *Cohen’s d =* -0.71. The same pattern was found in low spatial frequency images, *t* (37) = -3.439, *p* = .001, *Cohen’s d =* -0.56. Effect sizes across the measure of saccade length were moderate to large. There was no significant difference in saccade length between nature and urban scenes for mid-high spatial frequency images, *t* (37) = 0.689, *p* = .495, *Cohen’s d =* 0.11 (see Figure 5).

*Saccade length across viewing time:*

Analysing the mean saccade length across the five epochs showed no differences (all *p*'s > 0.09) other than longer saccades in nature than urban scenes when the images were in mid-high spatial frequency format in the fourth epoch, *t* (993) = 2.126, *p* = .034, *Cohen’s d =* 1.50. Together, the analysis of mean saccade length across viewing time appears contradictory to the findings of total saccade length over the full viewing time, however the longer *total* saccade length to urban relative to nature images (across the full, 15-seconds of viewing) can be attributed to the increased number of fixations to urban scenes (participants are exploring the images more widely). In conjunction with this, the analysis of *mean* saccade length (within each epoch) shows that, apart from one exception, these saccades were of similar length.

***Image ratings and the relationship between ratings and eye movements***

*Pleasantness ratings:*

For pleasantness ratings there was a significant effect of image type, *F* (1,37) = 26.862, *MSE =* 0.660, *p <* .001, *η2p* = 0.421. Participants rated the nature scenes as significantly more pleasant than the urban scenes (means of 3.12 and 2.56 respectively). There was also a significant effect of image manipulation, *F* (2,74) = 107.656, *MSE =* 0.461, *p <* .001, *η2p* = 0.744. The original images were rated as significantly more pleasant (*M* = 3.68) than the mid-high (*M* = 2.78) and the low spatial frequency images (*M* = 2.06), and the mid-high spatial frequency images were rated as significantly more pleasant than the low spatial frequency images (all *p’s* < .001). Large effect sizes showed clear differences between the conditions.

There was also a significant interaction between image type and image manipulation, *F* (2,74) = 25.471, *MSE =* 0.157, *p <* .001, *η2p* = 0.408. Paired samples t-tests showed that whilst pleasantness ratings for nature and urban scenes differed for the original images, *t* (37) = 6.431, *p* < .001, *Cohen’s d =* 1.04, and the mid-high spatial frequency images, *t* (37) = 3.828, *p* < .001, *Cohen’s d =* 0.62 (with nature scenes rated as more pleasant than urban), there was no significant difference between pleasantness of nature and urban low spatial frequency images, *t* (37) = 1.475, *p* = .149, *Cohen’s d* = 0.24 (see Figure 6a).

Pearson’s correlations showed that pleasantness ratings were significantly negatively correlated with number of fixations, *r* (20) = -0.597, *p <* .01, *R2* = 0.36, and significantly positively correlated with mean fixation duration, *r* (20) = 0.669, *p <* .01, *R2* = 0.45. Both relationships were moderate. Images rated as more pleasant were associated with fewer fixations but longer fixations. There was no relationship between ratings of pleasantness and total saccade length, *r* (20) = -0.357, *p =* .122, *R2* = 0.13.

*Ease of identification ratings:*

For the ease with which objects could be identified across the different images, objects in urban scenes were rated as significantly easier to identify (*M* = 3.28) than objects in nature scenes (*M* = 3.05), *F* (1,37) = 30.223, *MSE =* 0.106, *p <* .001, *η2p* = 0.451. There was also a significant effect of image manipulation, *F* (2,74) = 472.634, *MSE =* 0.333, *p <* .001, *η2p* = 0.927. The post-hoc comparisons showed that objects were rated as easier to identify in original images (*M* = 4.71) compared to both mid-high (*M* = 2.92) and low spatial frequency images (*M* = 1.86), and that identification was easier in mid-high compared to low spatial frequency images (all *p’s* < .001). The large effect sizes show the influence of image type and manipulation on the ability to identify objects within a scene.

The interaction between image type and image manipulation was also significant, *F* (1.703, 63.012) = 5.097, *MSE =* 0.068, *p =* .012, *η2p* = 0.121. Paired samples t-tests showed that for all image manipulations participants consistently rated identification as easier in the urban scenes than the nature scenes for original images, *t* (37) = -3.186, *p* = .003, *Cohen’s* *d =* -0.52, the mid-high spatial frequency images, *t* (37) = -272, *p* = .010, *Cohen’s d =* -0.44, and low spatial frequency images, *t* (37) = -6.535, *p* < .001, *Cohen’s d =* -1.06. The relative difference between ratings was however smallest for the original images and greatest for the low spatial frequency images, see Figure 6b.

When analysing the relationship between eye movements and the ease of identification ratings for original images there was no correlation between the ratings and number of fixations, *r* (20) = 0.405, *p =* .077, *R2* = 0.16, mean fixation duration, *r* (20) = -0.278, *p =* .235, *R2* = 0.07, or total saccade length, *r* (20) = 0.290, *p =* .215, *R2* = 0.08.

**Discussion**

To further investigate the underlying mechanisms responsible for attention restoration, the present study measured gaze behaviour whilst participants viewed nature and urban scenes and rated them for pleasantness and ease of identifying objects. The intention was to compare eye movements made to nature and urban images, both spatially (the extent of exploration across a scene) and temporally (whether exploration changes over the course of viewing a scene). To investigate the importance of low-level scene properties in the allocation of attention in nature and urban scenes, the images were presented in their original (grayscale) format, and with mid-to-high or low spatial frequencies removed. Additionally, consistency in gaze behaviour was explored across the different image types to gain insight about the potential for measuring individual differences in attention restoration.

The first important finding was that eye movements were highly consistent within participants across exemplars of different types of scene. Participants were therefore viewing each image within both urban and nature categories in a highly similar way. This reliability in gaze behaviour within participants has been found for other image types, for example complex everyday scenes (De Haas et al., 2019) and paintings (Trawiński et al., 2023), but this is the first time such analysis has been used to consider the allocation of attention in nature and urban settings. Not only does this finding suggest a role for individual differences in the perception of, and the allocation of attention to nature and urban environments, the high level of consistency in gaze behaviour indicates that these effects would generalise to other scenes within each class of image. Consequently, despite the relatively small number of images used in this study, the reliability analysis provides a high level of confidence in the findings.

In addition to the novel finding of reliability within eye movements to nature and urban scenes, the findings also replicated other studies in this field. Participants made fewer but longer fixations when viewing nature scenes compared to urban scenes. There was also more visual exploration (measured using saccade length) in urban scenes than nature scenes. Across the analysis, effect sizes were moderate to large, indicating a meaningful difference in gaze behaviour between nature and urban settings. These findings support past research (Batool et al., 2022; Berto et al., 2008; Valtchanov & Ellard, 2015), and the patterns of eye movements provide evidence for easier disengagement of attention from stimuli in urban scenes than stimuli in nature scenes (Batool et al., 2022; Eich & Beck, 2023). Scene viewing involves disengagement of attention from one object/location, a shifting of attention, and the engagement of attention to another object/location (Awh et al., 2012; Posner 1980). There is now a body of evidence showing more engagement in nature scenes (longer fixations, fewer fixations, concentrated fixations) and more disengagement in urban scenes (shorter fixations, more fixations, greater exploration).

An important aspect of disengagement is the motivation to shift attention from a current object/location to a new object/location, and it would be expected that environments containing more salient information (i.e., urban environments) would involve more motivation to shift attention. Additionally, when stimuli have a high level of self-similarity (Bourrier et al., 2018; Joye et al., 2016), there is less competition for attention, less need to shift attention, and therefore less motivation to disengage. From this perspective, reduced exploration and increased fixations in nature scenes demonstrate the lack of varied content within an environment; exploration is lower because there is less novelty to draw attention. The concepts of soft fascination and perceptual fluency may therefore be a product of the level of engagement or disengagement of attentional resources.

Whilst the differences in eye movements to nature and urban scenes found in previous research were replicated here, the present findings showed that these differences (specifically in terms of fixation duration) were most pronounced in the first 5-6 seconds of scene viewing. This effect shows the importance of measuring temporal shifts of attention in addition to spatial shifts of attention. One criticism levelled at studies investigating gaze behaviour in nature and urban scenes has been the length of time that stimuli are presented for. The common approach (which was adopted in the present work) has been to show each image for 15 seconds, however Wu et al. (2021) proposed that this may be too long and will lead to fixations that do not contribute to knowledge of attention restoration. The findings presented here support this and suggest that measuring attention and eye movements across the full viewing time may be diluting any effects of the environmental setting on attention.

Further support for analysing gaze behaviour over the course of a trial comes from Franěk et al. (2018b). They compared eye movements to nature and urban scenes (each presented for 15-seconds) whilst participants listed to fast-tempo music, slow-tempo music, or no music, and they analysed fixation durations over time by separating the 15-seconds of viewing into three time-windows (0-5 seconds, 5-10 seconds, and 10-15 seconds). In line with the present findings, fixations were shorter in urban scenes, however, in direct contrast to the current work, they found these differences to be larger in the later stages of scene viewing. It is difficult to reconcile these findings, although there are clear differences between the experiments, with Franěk and colleagues using colour images with no manipulations to spatial frequency. Both studies do however demonstrate the importance of measuring gaze behaviour over time.

The relatively short-term effects of image type on eye movements leads to the third important finding in the current study; low-level scene properties only appear to have a moderate impact on gaze behaviour. The influence of spatial frequency filtering supports this, but the results still provide valuable insight. Whilst participants made fewer fixations and longer fixations to nature scenes than urban scenes, this effect only occurred when mid-to-high spatial frequencies were present in the images (the difference disappeared when only low spatial frequencies were present). In contrast, the longer saccade lengths in urban compared to nature scenes was only found when low spatial frequencies were present. Mid-to-high spatial frequencies provide information about fine detail and edges, whereas low spatial frequencies provide information about the spatial layout of information (Oliva, 2005). Mid-to-high spatial frequencies may therefore allow for greater focus on objects (longer fixations), and low spatial frequencies may encourage greater exploration (longer saccade lengths).

The impact of spatial frequency was also found in the image ratings, as participants rated nature scenes as more pleasant than urban scenes unless mid-to-high spatial frequencies were removed. Objects in urban images were also rated as easier to identify than in nature scenes, and this difference was greatest when mid-to-high spatial frequencies were removed. The influence of filtering on pleasantness ratings aligns to past research conducted by Berman et al. (2014). They found that natural scenes were characterised by increased edge density. Removing mid-high spatial frequencies from an image reduces the edges, whereas removing low spatial frequencies would magnify the repeating elements within a scene, serving to magnify the edge density. Given that edge density is also associated with ratings of preference (i.e., Kardan et al., 2015) it follows that pleasantness ratings would be contingent on the manipulation of spatial frequency. The current research therefore adds further evidence for the importance of low-level scene properties in the processing of nature and urban settings.

The effects of image manipulation on the image ratings may also relate to the work of Marois et al. (2021) who compared eye movements to nature scenes that had been rated as high or low in mystery. Participants rated high mystery images as more fascinating and more pleasant than low mystery images and eye movements reflected greater engagement with high mystery nature scenes. Filtering images to modify spatial frequency impacts the visibility of a scene. The lack of clarity in low spatial frequency images (mid-high frequencies removed) has a greater impact on visibility, potentially resulting in less interest and engagement in a scene. In contrast, items can still be detected in the mid-high spatial frequency images (low frequencies removed) and the increased visibility of edges may enhance the level of mystery within a scene, making the image more engaging and more pleasant to view. Supporting this, Berman et al. (2017) outlined that high spatial frequencies are important for showing the content within a real-world scene and aiding in the identification of a scene (i.e., as nature or urban). It is therefore posited that removing mid-high spatial frequencies reduces the ability to identify objects in an image and reduces pleasantness of an image. This may have a significant impact on attention restoration due to the influence of pleasantness on attention engagement (Marois et al., 2021).

Valtchanov and Ellard (2015) proposed that mid-to-high spatial frequencies may be important for the affective benefits of restorative environments and low spatial frequencies may be important for the attentional benefits. The current findings add to this, demonstrating that filtering different spatial frequencies has distinct influences on eye movements. Again, these differences may indicate the importance of attentional engagement/disengagement; a lack of fine level detail would make it difficult to focus on stimuli within an image, leading to greater exploration and more frequent shifting between stimuli. Kaplan (1995) proposed that urban scenes contain more distractions and the need to inhibit these distractions using directed attention prevents resources from restoring, however, as an alternative, the limited scope for restoration in urban settings could be a consequence of frequent shifting of attention.

Interestingly, De Cesarei et al. (2023) report that attentional engagement with a scene is impacted by the motivational relevance to the observer. This could relate to another feature of the Attention Restoration theory (ART; Kaplan & Kaplan, 1989; Kaplan, 1995), compatibility (whether an environment is consistent with one’s aims and intentions). Supporting this, when analysing social media (Twitter) comments about urban green spaces in accordance with the ART’s four factors of restorative environments, Wilkie et al. (2020) found that compatibility was the second most important aspect after fascination. Compatibility may therefore warrant further study as the level of compatibility may affect engagement/disengagement.

Motivational relevance and compatibility are top-down, goal-driven influences upon attention and recent research puts more emphasis on the role for such influences on attention restoration. For instance, Batool et al. (2022) found increased visual exploration in nature scenes and reduced exploration in urban scenes as nature relatedness increased. Menzel and Reese (2022) also made the argument for the importance of high-level processing. They presented participants with original, phase scrambled, or line drawings of nature and urban scenes and participants completed a backwards digit span and a questionnaire measuring emotional affect before and after image-viewing. There was no effect of image type on cognition or affect, despite participants rating the nature images as higher in restorativeness, implying that the semantic content in an image is more important than the low-level scene properties. They suggested that low-level features may have only an indirect influence on restoration and visual features are not solely responsible for the positive effects of nature.

Twedt et al. (2019) found that the best predictor of perceived restorativeness of scenes was visual appeal, and the current findings of higher ratings of pleasantness for nature images aligns with the evidence showing that nature settings are more restorative. However, it could be argued that this also suggests a need to explore individual differences, as factors such as past experience may influence perceptions of visual appeal and therefore may influence attentional engagement. Again, the importance of individual differences is being acknowledged in some research, for example studies by Van den Berg et al. (2016) and Wilkie and Clouston (2015) asked participants if they consider themselves to be more of a nature/urban person, or both. If individual differences are important then this could be crucial information. This represents a limitation to the current work because although gaze behaviour was analysed to test for consistency within observers, no measures were included to explore differences between observers. Future work should explore individual differences but should carefully consider the sample size. The participants in this study were from an opportunity-based sample and the sample size is relatively small (although comparable to studies such as Franěk et al., 2018). This limits the generalizability of the results and further research exploring individual differences factors would necessitate a larger sample.

The potential impact of factors such as visual appeal and image complexity raises a further issue with the current work because the images chosen were not controlled for important aspects such as contrast, complexity, and ‘fractalness’, all of which may impact attention, and all of which may vary in different scenes. For instance, Van den Berg et al. (2016) found that magnified portions of nature scenes were rated as more complex, but this was specifically for scenes that had trees, etc., not for scenes with shrubs and grassland. They also found some city scenes were rated as more complex (indicating greater fractalness) when they contained more traditional buildings with increased level of ornamentation. Such aspects were not considered in the present study, and future work should ensure greater control of the stimuli. Additionally, all images were grayscale, which was critical to the modification of spatial frequency, but this level of image-control reduces the ecological validity of the findings.

Finally, another limitation with the present work is the fact that eye movements were used to explore possible mechanisms of the attention restoration effect, but there was no direct measure of restoration (i.e., comparing performance on a cognitive task before and after exposure to the scenes). The goal was to measure attention to nature and urban scenes to explore the mechanisms underpinning restoration, but restoration assumes some level of mental fatigue beforehand, and it is unclear whether participants were fatigued before viewing the images. Stevenson et al. (2018) made the argument that mental fatigue should be a consideration in the research within this field and whilst some studies induce fatigue, others do not. This is important because without comparing cognitive performance before and after scene viewing, and/or without ensuring a level of cognitive fatigue, it is impossible to know whether attention restoration occurred during scene viewing (and therefore whether patterns of eye movements are associated with attention restoration). Without explicitly measuring attention restoration, it is unclear whether the findings reported here reflect the restoration of attentional resources through exposure to different environmental settings.

A possible avenue for future research would be to measure gaze behaviour in real-world nature and urban environments. Investigating eye movements in a controlled setting is a starting point to try and understand the process of restoration, and the large effect sizes reported here indicate important influences of environmental features and properties on viewing, but it would be useful in future to know if visual behaviour in real environments shares the same characteristics. In particular, there are many more factors involved in visual attention and viewing behaviour in the real world, such as navigation, avoiding obstacles, etc., and these would likely have an influence on engagement, shifting, and disengagement of attention. It should also be noted that exposure to real nature environments (relative to virtual nature environments, videos, and images) has a larger effect on the restoration of some cognitive processes (Stevenson et al., 2018). Consequently, to fully understand attention restoration, any effects found in a laboratory setting should be further investigated in the field.

To conclude, this study used eye tracking to measure the allocation of attention in nature and urban scenes that had been manipulated to filter low-level scene characteristics. The intention was to investigate spatial and temporal differences in gaze behaviour and provide evidence that may help to explain the mechanisms associated with attention restoration. In accordance with past findings, gaze behaviour was different in nature and urban images. Participants made fewer fixations and longer fixations when viewing nature scenes, and they explored urban scenes more than nature scenes. Crucially, this effect was most pronounced early in scene viewing, highlighting the importance of studying the time course of attention and eye movements. The pattern of eye movements suggests an important role of attentional disengagement and engagement and these processes may help in understanding the concepts of soft fascination and perceptual fluency. The moderate impact of spatial frequency filtering and the novel finding that eye movements are highly consistent within participants suggests that it is important to study individual differences in the allocation of attention in different environmental settings.

The current work presents key factors that may shed light on the relationship between cognition and the environment. In accordance with Jenkins’ (1979) tetrahedral model, these factors can be organised into materials, participants, and context. First, this experiment manipulated low-level properties of nature and urban stimuli to explore how this may influence attention. The findings showed that spatial frequency impacts the appeal of an image and the guidance of attention within an image, albeit for a short period of time. This highlights the value of utilising different materials within a controlled, laboratory setting. Employing a new approach, the study measured individual consistency and reliability in eye movements to nature and urban scenes. With evidence for highly consistent patterns of visual exploration, and a high reliability, the results suggest the importance of individual differences. This provides a rationale for future research to explore differences between participants that may be driving the attention restoration effect. The work also took the approach of exploring attention and eye movements during the presentation of nature and urban scenes, rather than comparing attention before and after viewing different environments. Eye-tracking allows researchers to study the direct impact of an environment, rather than simply assessing the indirect benefits following engagement with an environment, and provides an objective measure of attentional engagement. Within this context, it is possible to further understand how different environmental settings may be processed, and therefore how and why attention restoration may occur.

**Declaration statement**

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

This work was supported by a New Graduate Research Bursary from the Experimental Psychology Society awarded in April 2021.

**Author Contributions**

CT and DB conceptualised and designed the study and collected the data. CT and TT analysed the data. CT, ND, and TT drafted the manuscript. All authors contributed to the discussion of the findings and approved the final version of the manuscript.

**Data availability statement**

Images, eye tracking data, and rating data are available at <https://osf.io/4wykb/?view_only=f4a54135b01d4447b1172f3b894d4660>. No aspects of the study were pre-registered.

**ORCID**

Catherine Thompson <https://orcid.org/0000-0002-7967-7019>

Tobiasz Trawinski <https://orcid.org/0000-0001-5285-4559>

David Beevers <https://orcid.org/0000-0002-4206-123>

Neil Harrison <https://orcid.org/0000-0002-2181-2046>

Nick Donnelly <https://orcid.org/0000-0002-3577-5434>

**References**

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences, 16,* 437-443. <https://doi.org/10.1016/j.tics.2012.06.010>

Baddeley, R. (1997). The correlational structure of natural images and the calibration of spatial representations. *Cognitive Science, 21*(3), 351-371. <https://doi.org/10.1016/S0364-0213(99)80027-5>

Batool, A., Rutherford, P., McGraw, P., Ledgeway, T., & Altomonte, A. (2022). Gaze correlates of view preference: Comparing natural and urban scenes. *Lighting Research & Technology, 54*(6), 576-594. <https://doi.org/10.1177/14771535211055703>

Bendall, R. C. A., Begley, S., & Thompson, C. (2021). Interactive influences of emotion and extraversion on visual attention. *Brain and Behavior, 11*(11), e2387. <https://doi.org/10.1002/brb3.2387>

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

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

Berman, D., Golomb, J. D., & Walther, D. B. (2017). Scene content is predominantly conveyed by high spatial frequencies in scene-selective visual cortex. *PLoS ONE,* *12*(12), e0189828. <https://doi.org/10.1371/journal.pone.0189828>

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

Berto, R., Massaccesi, S., & Pasini, M. (2008). Do eye movements measured across high and low fascination photographs differ? Addressing Kaplan’s fascination hypothesis. *Journal of Environmental Psychology, 28*(2), 185-191. <https://doi.org/10.1016/j.jenvp.2007.11.004>

Bourrier, S. C., Berman, M. G., Enns, J. T. (2018). Cognitive strategies and natural environments interact in influencing executive function. *Frontiers in Psychology, 9,* 1248. <https://doi.org/10.3389/fpsyg.2018.01248>

Bratman, G. N., Daily, G. C., Levy, B. J., & Gross, J. J. (2015). The benefits of nature experience: Improved affect and cognition. *Landscape and Urban Planning, 138,* 41-50. <https://doi.org/10.1016/j.landurbplan.2015.02.005>

Chun, M. M., & Potter, M. C. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance, 21*(1), 109-27. <https://doi.org/10.1037//0096-1523.21.1.109>

De Cesarei, A., Sambuco, N., D’Ascenzo, S., Nicoletti, R., & Codispoti, M. (2023). Delayed ocular disengagement from arousing scenes. *Frontiers in Psychology, 14,* 1297192. <https://doi.org/10.3389/fpsyg.2023.1297192>

De Haas, B., Iakovidis, A. L., Schwarzkopf, D. S., & Gegenfurtner, K. R. (2019). Individual differences in visual salience vary along semantic dimensions. *Proceedings of the National Academy of Sciences*, *116*(24), 11687-11692. <https://doi.org/10.1073/pnas.1820553116>

Duvall, J. (2011). Enhancing the benefits of outdoor walking with cognitive engagement strategies. *Journal of Environmental Psychology, 31*(1), 27-35. <https://doi.org/10.1016/j.jenvp.2010.09.003>

Eich, B., & Beck, M. R. (2023). Differences in the duration of the attentional blink when viewing nature vs. urban scenes. *Attention, Perception, & Psychophysics, 85,* 1846-1867. [https://doi.org.10.3758/s13414-023-02749-7](about:blank)

Faber, L. G., Maurits, N. M., & Lorist, M. M. (2012). Mental fatigue affects visual selective attention. *PLoS ONE, 7*(10), e48073. <https://doi.org/10.1371/journal.pone.0048073>

Fleming, W., Rizowy, B., & Shwartz, A. (2024). The nature gaze: Eye-tracking experiment reveals well-being benefits derived from directing visual attention towards elements of nature. *People and Nature, 6*(4),1-17. <https://doi.org/10.1002/pan3.10648>

Franěk, M., Petružálek, J., & Šefara, D. (2019). Eye movements in viewing urban images and natural images in diverse vegetation periods. *Urban Forestry & Urban Greening, 46,* 126477. <https://doi.org/10.1016/j.ufug.2018.126477>

Franěk, M., Šefara, D., Petružálek, J., Cabal, J., & Myška, K. (2018a). Differences in eye movements while viewing images with various levels of restorativeness. *Journal of Environmental Psychology*, *57,* 10–16. <https://doi.org/10.1016/j.jenvp.2018.05.001>

Franěk, M., Šefara, D., Petružálek, J., Mlejnek, R., & van Noorden, L. (2018b). Eye movements in scene perception while listening to slow and fast music. *Journal of Eye Movement Research, 11*(2), 8. <https://doi.org/10.16910/jemr.11.2.8>

Grady, J. N., Cox, P. H., Nag, S., & Mitroff, S. R. (2022). Conscientiousness protects visual search performance from the impact of fatigue. *Cognitive Research: Principles and Implications, 7*, 56. <https://doi.org/10.1186/s41235-022-00410-9>

Hahn, S., Buttaccio, D. R., & Lee., T. (2015). Personality and attention: Levels of neuroticism and extraversion can predict attentional performance during a change detection task. *The Quarterly Journal of Experimental Psychology, 68*(6), 1041-1084. <https://doi.org/10.1080/17470218.2015.1032986>

Harrison, N. (2023). Level of attention mediates the association between connectedness to nature and aesthetic evaluations of photographs of nature. *PeerJ,* 11, e14926. <https://doi.org/10.7717/peerj.14926>

Hartig, T. (2021). Restoration in nature: Beyond the conventional narrative. In: A. R. Schutte, J. C. Torquati, & J. R. Stevens, J.R. (Eds.), *Nature and Psychology. Nebraska Symposium on Motivation, 67.* Springer. <https://doi.org/10.1007/978-3-030-69020-5_5>

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods, 50*(3), 1166–1186. <https://doi.org/10.3758/s13428-017-0935-1>

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.

Jenkins, J. J. (1979). Four points to remember: A tetrahedral model of memory experiments. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 429–446). Erlbaum.

Joye, Y., & Dewitte, S. (2018). Nature’s broken path to restoration. A critical look at Attention Restoration Theory. *Journal of Environmental Psychology, 59,* 1-8. <https://doi.org/10.1016/j.jenvp.2018.08.006>

Joye, Y., & van den Berg, A. E. (2012). Restorative environments. In L. Steg, A. E. van den Berg, & J. I. M. de Groot (Eds.), *Environmental Psychology: An Introduction* (pp. 58-66). John Wiley-Blackwell.

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>

Kaplan, S. (1995). The restorative benefits of nature: towards an integrative framework. *Journal of Environmental Psychology, 15,* 169-182. [https://10.1016/0272-4944(95)90001-2](about:blank)

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. and Kaplan, S. (1989). *The experience of nature: A psychological perspective*. Cambridge, UK: Cambridge University Press.

Kardan, O., Demiralp, E., Hout, M. C., Hunter, M. R., Karimi, H., Hanayik, T., Yourganov, G., Jonides, J., & Berman, M. G. (2015). Is the preference of natural versus man-made scenes driven by bottom-up processing of the visual features of nature. *Frontiers in Psychology, 6,* 471. <https://doi.org/10.3389/fpsyg.2015.00471>

Lee, K. L., Williams, K. J. H., Sargent, L. D., Williams, N. S. G., & Johnson, K. A. (2015). 40-second green roof views sustain attention: The role of micro-breaks in attention restoration. *Journal of Environmental Psychology, 42,* 182-189. <https://doi.org/10.1016/j.envp.2015.04.003>

Marois, A., Charbonneau, B., Szolosi, A. M., & Watson, J. M. (2021). The differential impact of mystery in nature on attention: An oculometric study. *Frontiers in Psychology, 12,* 759616. <https://doi.org/10.3389/fpsyg.2021.759616>

Mayer, S. (2024). *Imagefluency: Image statistics based on processing fluency.* R package version 0.2.5. [https://doi.org.10.5281/zenedo.5614665](about:blank)

Menzel, C., & Reese, G. (2021). Implicit associations with nature and urban environments: Effects of lower-level processed image properties. *Frontiers in Psychology, 12,* 591403. <https://doi.org/10.3389/fpsyg.2021.591403>

Menzel, C., & Reese, G. (2022). Seeing nature from low to high levels: Mechanisms underlying the restorative effects of viewing scenes. *Journal of Environmental Psychology, 81,* 101804. <https://doi.org/10.1016/j.jenvp.2022.101804>

Oliva, A. (2005). Gist of the scene. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *Neurobiology of attention* (pp. 251-256). Academic Press.

Oliva, A., Torralba, A. B., Guerin-Dugue, A., & Herault, J. (1999). Global semantic classification of scenes using power spectrum templates. In *Challenge of image retrieval*. BCS Learning & Development.

Parsons, S., (2021). splithalf: robust estimates of split half reliability. *Journal of Open Source Software, 6*(60), 3041. <https://doi.org/10.21105/joss.03041>

Posner, M., I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3-25. <https://doi.org/10.1080/00335558008248231>

Risko, E. F., Anderson, N. C., Lanthier, S., & Kingstone, A. (2012). Curious eyes: Individual differences in personality predict eye movement behavior in scene-viewing. *Cognition, 122*(1), 86-90. <https://doi.org/10.1016/j.cognition.2011.08.014>

Schertz, K. E., & Berman, M. G. (2019). Understanding nature and its cognitive benefits. *Current Directions in Psychological Science, 28*(5), 496-502. <https://doi.org/10.1177/0963721419854100>

Stevenson, M. P., Schilhab, T., & Bentsen, P. (2018). Attention restoration theory II: A systematic review to clarify the attention processes affected by exposure to natural environments. *Journal of Toxicology and Environmental Health, Part B, 21*(4), 227-268. <https://doi.org/10.1080/10937404.2018.1505571>

Taylor, A. F., Kuo, F. E., & Sullivan, W. C. (2002). Views of nature and self-discipline: Evidence from inner city children. *Journal of Environmental Psychology, 22*(1-2), 49-63.

Tennessen, C. M., & Cimprich, B. (1995). Views to nature: Effects on attention. *Journal of Environmental Psychology, 15*(1), 77–85. [https://doi.org/10.1016/0272-4944(95)90016-0](https://psycnet.apa.org/doi/10.1016/0272-4944(95)90016-0)

Torralba, A., & Oliva, A. (2003). Statistics of natural image categories. *Network, 14*(3), 391-412.

Trawiński, T., Zang, C., Liversedge, S. P., Ge, Y., & Donnelly, N. (2023). The time-course of fixations in representational paintings: A cross-cultural study. *Psychology of Aesthetics, Creativity, and the Arts, 17*(4), 412-427. <https://doi.org/10.1037/aca0000508>

Twedt, E., Rainey, R. M., & Proffitt, D. R. (2019). Beyond nature: The roles of visual appeal and individual differences in perceived restorative potential. *Journal of Environmental Psychology, 65,* 101322. <https://doi.org/10.1016/j.jenvp.2019.101322>

Ulrich, R. S., Simons, R. F., Losito, B. D., Fiorito, E., Miles, M. A., & Zelson, M. (1991). Stress recovery during exposure to natural and urban environments. *Journal of Environmental Psychology, 11*(3), 201-230. <https://doi.org/10.1016/S0272-4944(05)80184-7>

Valtchanov, D., & Ellard, C. G. (2015). Cognitive and affective responses to natural scenes: Effects of low level visual properties on preference, cognitive load and eye-movements. *Journal of Environmental Psychology, 43,* 184-195. <https://doi.org/10.1016/j.jenvo.2015.07.001>

Van den Berg, A. E., Joye, Y., & Koole, S. L. (2016). Why viewing nature is more fascinating and restorative than viewing buildings: A closer look at perceived complexity. *Urban Forestry & Urban Greening, 20,* 397-401. <https://doi.org/10.1016/j.ufug.2016.10.001>

Wilkie, S., & Clouston, L. (2015). Environmental preference and environment type congruence: Effects on perceived restoration potential and restoration outcomes. *Urban Forestry & Urban Greening, 14,* 368-376. <https://doi.org/10.1016/j.ufug.2015.03.002>

Wilkie, S., Thompson, E., Cranner, P., & Ginty, K. (2020). Attention restoration theory as a framework for analysis of Tweets about urban green space: A case study. *Landscape Research, 45*(6), 777-788. <https://doi.org/1080/01426397.2020.1738363>

Wu, Y., Zhuo, Z., Liu, Q., Yu, K., Huang, Q., & Liu, J. (2021). The relationships between perceived design intensity, preference, restorativeness and eye movements in designed urban green space. *International Journal of Environmental Research and Public Health*, *18*(20), 10944. <https://doi.org/10.3390/ijerph182010944>

Figure 1:

A lake with a mountain in the background

Description automatically generated Background pattern

Description automatically generated with medium confidence No image

Description automatically generated

A city street with cars and buses

Description automatically generated with low confidence A picture containing background pattern

Description automatically generated No image

Description automatically generated

A B C

Figure 2:

A diagram of different types of images

Description automatically generated with medium confidence

Figure 3:

Figure 4:



Figure 5:

Figure 6:

Figure Captions:

*Figure 1:* Examples of nature (top) and urban (bottom) images used in Experiment One. All images were presented in grayscale and the original images (A) were modified to create images that only contained mid-high spatial frequencies (B) or only contained low spatial frequencies (C).

*Figure 2:* Spearman-Brown corrected reliability estimates (*r-*splits) for the number of fixations to original, mid-high, and low spatial frequency nature and urban images.

*Figure 3:* The number of fixations (a) and the mean fixation duration in milliseconds (b) to nature and urban images for the three levels of image manipulation. Error bars represent standard error.

*Figure 4:* The mean fixation duration (top panel) and the mean saccade length (bottom panel) to nature and urban images for the three levels of image manipulation across five epochs of viewing. The grey shaded ribbons indicate the mean ± standard error for each condition. *Note:* The vertical shaded area highlights the viewing epochs where significant differences between nature and urban images were observed.

*Figure 5:* Showing the total distance travelled (in pixels) on the nature and urban scenes measured using the total length of saccades in an image.

*Figure 6:* Image ratings for pleasantness (a) and ease of identification (b) for nature and urban images according to image manipulation. Error bars represent standard error.

1. Valtchanov and Ellard (2015) also analysed blink rate as a measure of cognitive load (increased blink rate is associated with increased mental workload (Holmquist et al., 2011)) and found a lower blink rate for nature scenes compared to urban scenes. However, this was only for original images and images with low spatial frequencies intact, there was no difference in blink rate in nature and urban scenes when low spatial frequencies were removed. From this, they concluded that low spatial frequency is an important component in the cognitive benefits gained from restorative environments. [↑](#footnote-ref-1)
2. To create the low spatial frequency images a Gaussian blur was performed on each scene using the GIMP Image Processing app (σ = 15). To create the mid-high spatial frequency images a Fast Fourier Transform was used to transfer the spatial domain into the frequency domain, the lower end of the frequency scale was removed, leaving the mid-to-high frequency content of the images intact, and the images were then whitened to minimise low spatial frequencies and enhance mid-high spatial frequencies. [↑](#footnote-ref-2)