

Urban and natural contexts differentially modulate attention bias towards threat

Eszter Toth¹, Ali Mazaheri¹, Jane E. Raymond¹

University of Birmingham

¹ *School of Psychology, University of Birmingham, Birmingham, B15 2 TT, UK*

Corresponding author:

Jane Raymond

School of Psychology

University of Birmingham

Edgbaston, Birmingham

B15 2TT

Email: J.Raymond@bham.ac.uk

Key words: attention, threat bias, environmental psychology, fearful faces, video exposure

Abstract

Recent research showing that amygdala hyperactivity is associated with urban living raises the possibility that attention bias to fear-related stimuli, a cognitive response thought to be underpinned by amygdala hyperactivity, may be enhanced by exposure to urban environments. To investigate, we asked young adults to complete a face attention task involving threatening and non-threatening distractors before and after watching a 25-minute video showing an immersive walk in city streets or a nature reserve. A crossover design was used so that all participants viewed both videos. In the face attention task, participants made a speeded gender identification of spatially cued target faces (neutral expression), ignoring a concurrent distractor that was either a face bearing a neutral, happy, or fearful expression or a scrambled face image. Although performance in the face attention task was unaffected by distractor-type after viewing the nature video, fearful faces specifically slowed responses after viewing the urban video, an effect that was independent of mood and stress level before or after viewing the video. These results show for the first time that exposure to urban stimuli increases attention sensitivity to threatening face stimuli. Such finding suggest that urban environments may heighten vigilance to sensory stimuli that are not directly pertinent to on-going tasks, a process that may underpin the association between cities and anxiety disorders.

Urban versus natural contexts determines attention bias towards threat

Humans preferentially orient their attention toward threatening or highly arousing negative stimuli, even when this is counterproductive. Indeed, numerous studies have shown that irrelevant fearful and angry faces are especially effective at interfering with performance on simple, speeded visual search tasks (Eastwood, Smilek & Merikle, 2001; Fox et al., 2000), including flanker (Barratt & Bundesen, 2012; Grose-Fifer, Rodrigues, Hoover & Zottoli, 2013) and dot-probe tasks (Fox, Derakshan & Shoker, 2008; Pourtois, Grandjean, Sander & Vuilleumier, 2004). Importantly, attention biases to threat related stimuli are widely reported to be linked to anxiety (Bar-Haim et al., 2007). Numerous studies suggest that the neural basis for this link involves the amygdala (Vuilleumier, 2005), a core midbrain structure that is part of a larger neural circuit involving thalamus and orbital-frontal cortex. Not only have neuroimaging studies shown that anxious participants who show threat biases also show amygdala hyperactivity (Carlson, Cha & Mujica-Parodi, 2013; Etkin et al., 2004), selective serotonin reuptake inhibitors (SSRIs) known to suppress amygdala activity (Browning, Reid, Cowen, Goodwin, & Harmer, 2007) are effective at reducing threat-related attention bias (Sheline et al., 2001). Additionally, amygdala damage is associated with reduced recognition of fear expressions in faces (Broks et al., 1998; Gamer et al., 2012), adding weight to the contention that amygdala plays an important role in attention to threat stimuli, especially threatening face stimuli.

Hyperactivity in the amygdala during social stress has been linked to living in urban environments (Lederbogen et al., 2011), a finding consistent with a meta-analysis showing that urban versus rural dwelling increases risk for anxiety by 21% (Peen, Schoevers, Beekman, & Dekker, 2010). Suggesting a specific role of the physical environment, Kim et al. (2010) reported greater amygdala activation when viewing urban versus rural scenes.

Moreover, exposure to urban versus natural environments (Song et al., 2014) or sounds (Jo et al., 2019) has been causally linked to elevated anxiety. Considering the relationship between amygdala activity and anxiety and that between anxiety and threat-related attention bias, these findings suggest that exposure to urban environments may enhance threat-related attention bias.

In the present study we used a within-subject crossover design to assess threat biases before and after viewing an immersive 25-minute video of walking through city streets or a nature reserve. Threat bias assessment required participants to report as quickly and accurately as possible the gender of a pre-cued target face that was presented simultaneously with another (distractor) image. Distractors were fearful, happy or neutral faces or scrambled meaningless images. Differences in performance with fearful versus neutral face distractors were used to index threat bias. We predicted and found enhanced threat biases after watching the urban video.

Methods

Participants

Complete data sets were collected from thirty-five young university-associated adults (23 females, mean age = 24.1 years, SD = 3.19) in exchange for cash. All were fluent in English, reported normal or corrected-to-normal vision and no history of neuropsychological or psychiatric disorders. The study was approved by the University of Birmingham Ethics Committee.

Power analysis (GPower 3.1; Faul, Erdfelder, Lang & Buchner, 2007) using Ter Huurne et al.'s (2015) result indicated that a large effect size of $d = -1.324$ with power of .95 required 24 participants. We tested 40 to cover potential dropouts.

Apparatus

A 17.5" Dell laptop (refresh rate, 59 Hz) displayed video, controlled presentation of experimental stimuli, and recorded data using MatLab (R2016a, Mathworks, 2007) running Psychtoolbox. Video presentation used Windows Media Player (Microsoft). Seated participants viewed the screen from 50 cm away in a quiet, well lit room and entered response using a gaming keyboard (Razer).

Stimuli

Face Task. Stimuli comprised a spatial cue, a centrally presented white fixation cross (0.5° in diameter), and two face stimuli. The spatial cue was a small ($1^\circ \times .7^\circ$), centrally presented white arrow pointing left or right. Each face in the face array subtended $8.8^\circ \times 10.2^\circ$ and was set into a $9.1^\circ \times 10.8^\circ$ rectangle. The centre of each face was presented 6.8° laterally to the left and right of centre and 3.1° below the horizontal meridian. A total of 32 colour photos of Caucasian adult faces with head hair but without glasses, make-up, or facial hair were used (Karolinska Directed Emotional Faces dataset; Lundqvist, Flykt, & Öhman, 1998). Teeth were visible in all fearful and happy faces. For each of four Face Tests, a unique set of 8 photos (4 females) was used. Two of each gender had a neutral expression and served as targets. Remaining faces serve as distractors, with each presented equally often having a neutral, fearful, or happy expression or were used to create a scrambled non-face image (made by pixelating the image into 13,984 elements, then randomly repositioning every element).

Videos were sourced from the internet: Nature video, <https://www.youtube.com/watch?v=b4AVn8mTuJw&t=2777s>, starting at 30 seconds; Urban video, <https://www.youtube.com/watch?v=LwYGxjdnZ84&t=1535s>, starting at minute 22. Both were 25-minutes long, presented without sound, and were filmed from a walker's perspective and speed at eye height. The nature video was filmed along a footpath, showing showed vegetation, a mountain and river, but no people or buildings. The urban video navigated busy streets in Vancouver, Canada, showing tall buildings, open urban spaces, signs in English, 335 neutral and three smiling faces (61% of faces were male), and little vegetation.

Procedure

The experiment consisted of two sessions conducted one week apart at the same time of day. On both days, participants completed two Face Test phases: A pre-video and a post-video phase, each comprising 12 blocks. Additionally, the pre-video phase on Day 1 contained 30 practice trials. After completing the pre-video phase, participants completed the Positive and Negative Affect Scale (PANAS; Watson, Clarke, & Tellegan, 1988); rated anxiety using a 6-point Likert scale (1 = very stressed, 6 = very relaxed), then viewed the Nature or Urban video (alternate video on Day 2; Day 1 video was counterbalanced across participants) with the instruction to imagine being present in the environment, following the route shown (De Kort, Meijenders, Sponselee, & IJsselsteijn, 2006). Immediately thereafter, PANAS and relaxation ratings were again obtained. Then, the post-video Face Test phase was conducted. On Day 2, participants also completed the Depression, Anxiety and Stress Scale (DASS; Lovibond & Lovibond, 1995; 21 items) reporting feelings over the last week and

then rated perceived pleasantness of urban and natural environments (4 items; 6-point scale) at the session's end.

Face attention task. The procedure used in each trial in the Face Task is illustrated in *Figure 1*. The task was to report the target's gender as quickly and accurately as possible by pressing either the "Z" or "A" key using the index and middle finger, respectively, of the dominant hand. Response keys were counterbalanced. The trial ended with response or after 1500 ms; no performance feedback was provided, and the next trial began immediately. Response time (interval between face array onset and response) and accuracy was recorded. Each combination of target location and gender was equally likely to occur. Target and distractor were 90% likely to be mismatched in gender. (Target-matched trials were used to detect strategies based on distractor selection and were not analysed.) The distractor was 25% likely to be happy, fearful, neutral, or scrambled. Participants completed 12 blocks (576 trials) for each pre- and post-video test episode. Rest intervals between blocks were self-paced. The order of face set used in each of the four test phases was counterbalanced across participants.

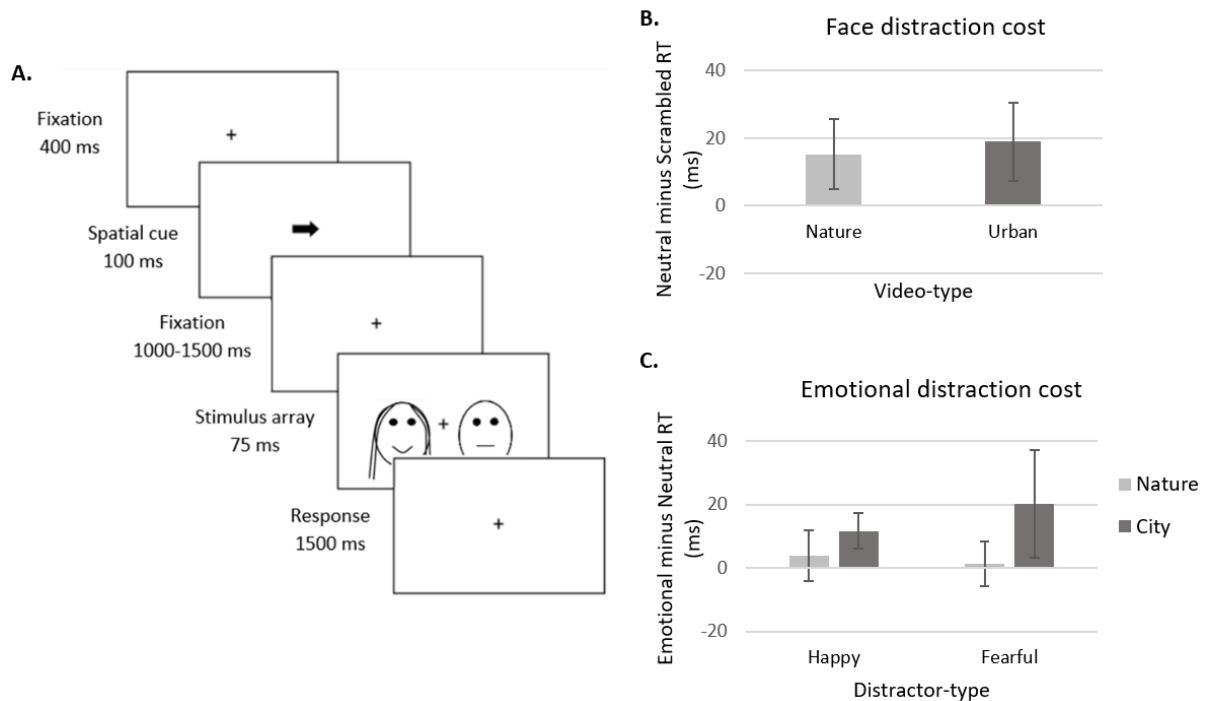


Figure 1. A. An illustration of an example trial in the Face task. Faces are shown as cartoon for illustration purposes only. On each trial a fixation cross was visible throughout except when replaced by a brief 100% valid spatial cue (100 ms). This occurred 400 ms after trial onset. A face array (75 ms) comprising target and distractor appeared between 1000 and 1500 ms after cue offset (randomly jittered interval), followed by a 1500 ms response window. A speeded report of the target's gender was required. **B.** Group mean face distraction cost measured after viewing each video type. Face distraction cost is calculated as response time (RT) for neutral minus that for scrambled face distractors. **C.** Group mean emotional distraction cost after viewing each video type. Emotional distraction cost is calculated as RT for happy or fearful distractors minus that for neutral distractors. Error bars show 95% confidence intervals.

Data analysis

Response time (RT) data were excluded for all gender-match trials, all incorrect trials, anticipation errors (RT < 200 ms) and when an RT exceeded the individual's condition mean by more than three s.d.'s. Individual proportion correct trials and average RT for each distractor condition (neutral, happy, fearful, scrambled), test phase (pre-video, post-video), video-type (nature, urban) and session order (Day 1 = nature, Day 2 = urban) were analysed using a 4 x 2 x 2 x 2 mixed-design ANOVA with distractor-type, test phase, video-type as within-subjects factors and session order as a between group factor. As session order

neither interacted with the other three factors ($p = .445$), nor had a significant main effect on RT ($p = .361$) or accuracy ($p = .108$), it was excluded from all subsequent analyses.

Separate repeated measures ANOVAs were conducted on pre-video and post-video data sets using distractor condition and video-types as factors. Mauchly's Test of Sphericity was used for all ANOVAs. Face distraction costs were calculated by subtracting RT for scrambled distractors from RT for neutral distractors; threat and happy distraction costs were calculated by subtracting RT for neutral distractors from RT for fearful and happy distractors, respectively. Mean comparisons used paired samples t -tests (2-tailed). Here and in all other analyses, Bonferroni corrections applied where necessary and alpha levels were set at .05.

Table 1. Mean pre- and post-video RT (ms) and accuracy (% correct) for each video-type and distractor-type. S.d. are shown in parentheses.

Video-type		Urban				Nature			
Distractor-type		Scrambled	Neutral	Happy	Fearful	Scrambled	Neutral	Happy	Fearful
RT	Pre-video	598 (118)	621 (126)	622 (119)	619 (124)	601 (98)	615 (104)	622 (102)	618 (99)
	Post-video	563 (94)	582 (100)	594 (105)	602 (108)	559 (91)	574 (98)	578 (94)	575 (93)
Accuracy	Pre-video	93 (.05)	90 (.06)	91 (.05)	91 (.06)	92 (.06)	90 (.07)	91 (.05)	90 (.06)
	Post-video	94 (.04)	92 (.06)	92 (.05)	93 (.06)	93 (.06)	92 (.07)	91 (.08)	91 (.08)

Ratings and scales. PANAS ratings for positive and negative items (10 each) were summed separately to produce two scores. Pre- and post-video relaxation ratings and each PANAS subscale scores were analysed using Wilcoxon Sign Ranks Tests (2-tailed). DASS scores for depression, anxiety, and stress items (7 items each) were summed separately to

produce three scores. Possible differences between those viewing the nature versus urban video on Day 1 on these subscales were assessed using Mann-Whitney U tests (2-tailed). Whether DASS subscales, post-video relaxation ratings and post-video PANAS subscales are predictive of threat bias was analysed using linear regression.

Data can be accessed at <https://reshare.ukdataservice.ac.uk/>.

Results

Threat bias was greater after viewing the urban video

Omnibus analyses of RT showed a significant interaction between phase (pre-video, post video), video-type, and distractor-type ($F_{3,99} = 3.121, p = .044$). Subsequent analyses of pre-video RTs found no interaction between upcoming video-type and distractor-type ($F_{3,102} = .670, p = .572$), indicating that distractor effects on performance were similar prior to either video. To focus on post-video distractor effects, face distraction as well as threat and happy faces biases were compared between video-types. Here, the group average threat bias (i.e. slowing of RT when distractor was a fearful face) was significantly larger after viewing the urban (20 ms, s.d. = 21) versus nature video (1 ms, s.d. = 21 ms; $t(34) = -4.584, p < .001$; see *Fig. 1*). In contrast, the happy face bias was modest and unaffected by video-type (nature: 4 ms, s.d. = 24; urban: 12 ms, s.d. = 17; $t(34) = -1.464, p = .152$). Similarly, mean face distraction cost was only 4 ms larger after the urban (19 ms, s.d. = 36) than the nature video (15 ms, s.d. = 32; $t(34) = -.668, p = .509$), indicating that effects on attention were threat specific. Although analysis of proportion correct (see *Table 1*) showed higher accuracy for scrambled (93%, s.d. = 4) versus face distractors (91%, s.d. = 5; $t(34) = 3.987, p < .001$; distractor main effect $F(3,99) = 8.085, p < .001, \eta^2_p = .197$), neither phase nor

distractor-type interacted with video-type (all p 's > .273), indicating that threat bias effects reported here are not due to speed-accuracy trade-offs.

Relaxation rating was lower but positive mood was higher after the urban video

Analysis of pre-video mood measures revealed no significant difference between the urban and nature videos ($p > .23$). Subsequent analyses of the post-video mood measures revealed that relaxation ratings were significantly lower after the urban (4.6, s.d. = 1) versus nature video (5.2, s.d. = .87; $Z = -2.457$, $p = .014$). In comparison, post-video score on the positive subscale of the PANAS were significantly higher after the urban (24.5, s.d. = 10) versus nature video (21.2, s.d. = 9; $Z = -2.875$, $p = .004$). PANAS negative subscale scores were unaffected by video-type ($p = .075$). Participants who watched the nature versus urban video on Day 1 did not differ on any of the DASS subscales ($p > .351$). Finally, neither DASS subscales nor post-city-video mood measures predicted post-city-video threat bias ($p > .084$), indicating that these measures did not contribute to threat bias.

Discussion

We contrasted the effects of nature and urban scene exposure on attention bias towards fearful faces. To assess attention bias, participants engaged in a visuospatial attention task in which target faces were flanked by faces displaying either a neutral, happy or fearful emotion. This assessment was conducted, before and after viewing a 25-minute video of an immersive walk through a nature reserve or a city. Distraction RT costs for neutral and happy face expressions were unaffected by video-type, suggesting that attention biases towards faces in general or towards faces with specifically happy expressions were not influenced by video content. Importantly, distraction costs for fearful

faces were greater after viewing the urban versus nature video, indicating that city exposures specifically induced attention bias towards fearful facial expressions.

One explanation for the threat attention bias seen after the urban video may be that viewing it heightened anxiety levels, a state associated with enhanced vigilance towards threat-related stimuli (see Bar-Haim et al., 2007 for review). Arguing against this possibility is that relaxation levels indicated lower anxiety levels after versus before viewing either video, and that no threat-related attention bias was detected before video viewing. This is supported by a regression analysis showing that participants' relaxation ratings (PANAS or DASS scores) did not contribute to their distractibility by fearful faces after the urban video.

An alternative explanation for our results may be that the urban video presented many faces whereas the nature video had none. As faces seen in the video were never fearful (largely neutral), some form of priming by fearful faces cannot account for these effects. However, the urban video showed crowds of people in close proximity to the camera (participant's viewpoint), creating a stimulus configuration known to activate the amygdala (Kennedy, Gläscher, & Adolphs, 2009), a region closely linked to threat-related attention bias (see Vuilleumier, 2005 for review). If so, such activation could underpin the threat bias observed here and would suggest that crowds of strangers are critical urban features for heightened threat bias. Although a city environment without people might therefore not elicit threat bias, overlearned associations between urban spaces and strangers are likely sufficient to produce this effect even in the absence of faces.

Sensitization to threat may be an advantageous adaptation in cities as it could help to avoid dangerous situations (Blumstein, 2016). However, threat biases are associated with

elevated anxiety levels (Cavanagh, Urry, & Shin, 2011) and may contribute to the development of anxiety disorders (Mathews & MacLeod, 2002). Our novel finding that virtual urban environments enhance threat attention bias suggests a basis for the relationship between city living and increased risk of anxiety disorders (Peen et al., 2010).

Despite anecdotal accounts that city dwellers easily ignore people around them, our results suggest instead that they, unlike their rural or suburban counterparts, may be hypersensitized to faces displaying fear, and quite possibly to other threat-related stimuli. Not only is our direct demonstration that virtual environments can modulate attention biases consistent with a handful of correlational studies linking threat biases to violent video-game play (Kirsh, Olczak, & Mounts, 2005), it raises the distinct possibility that exposure to nature, even virtually, could protect the mental health of urban dwellers.

References

- Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & Van Ijzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: a meta-analytic study. *Psychological Bulletin*, 133(1), 1.
- Barratt, D., & Bundesen, C. (2012). Attentional capture by emotional faces is contingent on attentional control settings. *Cognition & Emotion*, 26(7), 1223-1237.
- Blumstein, D. T. (2016). Habituation and sensitization: new thoughts about old ideas. *Animal Behaviour*, 120, 255-262.
- Broks, P., Young, A. W., Maratos, E. J., Coffey, P. J., Calder, A. J., Isaac, C. L., ... & Roberts, N. (1998). Face processing impairments after encephalitis: amygdala damage and recognition of fear. *Neuropsychologia*, 36(1), 59-70.

Browning, M., Reid, C., Cowen, P. J., Goodwin, G. M., & Harmer, C. J. (2007). A single dose of citalopram increases fear recognition in healthy subjects. *Journal of Psychopharmacology*, 21(7), 684-690.

Carlson, J. M., Cha, J., & Mujica-Parodi, L. R. (2013). Functional and structural amygdala–anterior cingulate connectivity correlates with attentional bias to masked fearful faces. *Cortex*, 49(9), 2595-2600.

Cavanagh, S. R., Urry, H. L., & Shin, L. M. (2011). Mood-induced shifts in attentional bias to emotional information predict ill-and well-being. *Emotion*, 11(2), 241.

de Kort, Y. A., Meijnders, A. L., Sponselee, A. A., & IJsselsteijn, W. A. (2006). What's wrong with virtual trees? Restoring from stress in a mediated environment. *Journal of Environmental Psychology*, 26(4), 309-320.

Eastwood, J. D., Smilek, D., & Merikle, P. M. (2001). Differential attentional guidance by unattended faces expressing positive and negative emotion. *Perception & Psychophysics*, 63(6), 1004-1013.

Etkin, A., Klemenhagen, K. C., Dudman, J. T., Rogan, M. T., Hen, R., Kandel, E. R., & Hirsch, J. (2004). Individual differences in trait anxiety predict the response of the basolateral amygdala to unconsciously processed fearful faces. *Neuron*, 44(6), 1043-1055.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175-191.

Fox, E., Derakshan, N., & Shoker, L. (2008). Trait anxiety modulates the electrophysiological indices of rapid spatial orienting towards angry faces. *Neuroreport*, 19(3), 259-263.

Fox, E., Lester, V., Russo, R., Bowles, R. J., Pichler, A., & Dutton, K. (2000). Facial expressions of emotion: Are angry faces detected more efficiently?. *Cognition & Emotion*, 14(1), 61-92.

Gamer, M., Schmitz, A. K., Tittgemeyer, M., & Schilbach, L. (2013). The human amygdala drives reflexive orienting towards facial features. *Current Biology*, 23(20), R917-R918.

Grose-Fifer, J., Rodrigues, A., Hoover, S., & Zottoli, T. (2013). Attentional capture by emotional faces in adolescence. *Advances in Cognitive Psychology*, 9(2), 81.

Jo, H., Song, C., Ikei, H., Enomoto, S., Kobayashi, H., & Miyazaki, Y. (2019). Physiological and psychological effects of forest and urban sounds using high-resolution sound sources. *International Journal of Environmental Research and Public Health*, 16(15), 2649.

Kennedy, D. P., Gläscher, J., Tyszka, J. M., & Adolphs, R. (2009). Personal space regulation by the human amygdala. *Nature Neuroscience*, 12(10), 1226-1227.

Kim, T. H., Jeong, G. W., Baek, H. S., Kim, G. W., Sundaram, T., Kang, H. K., ... & Song, J. K. (2010). Human brain activation in response to visual stimulation with rural and urban scenery pictures: A functional magnetic resonance imaging study. *Science of the Total Environment*, 408(12), 2600-2607.

Kirsh, S. J., & Mounts, J. R. (2007). Violent video game play impacts facial emotion recognition. *Aggressive Behavior: Official Journal of the International Society for Research on Aggression*, 33(4), 353-358.

Lederbogen, F., Kirsch, P., Haddad, L., Streit, F., Tost, H., Schuch, P., ... & Meyer-Lindenberg, A. (2011). City living and urban upbringing affect neural social stress processing in humans. *Nature*, 474(7352), 498.

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, 33(3), 335-343.

Lundqvist, D. et al. (1998). The Karolinska directed emotional faces (KDEF). *CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet* **91(630)**, 2-2.

Mathews, A., & MacLeod, C. (2002). Induced processing biases have causal effects on anxiety. *Cognition & Emotion*, 16(3), 331-354.

Peen, J., Schoevers, R. A., Beekman, A. T., & Dekker, J. (2010). The current status of urban-rural differences in psychiatric disorders. *Acta Psychiatrica Scandinavica*, 121(2), 84-93.

Pourtois, G., Grandjean, D., Sander, D., & Vuilleumier, P. (2004). Electrophysiological correlates of rapid spatial orienting towards fearful faces. *Cerebral cortex*, 14(6), 619-633.

Sheline, Y. I., Barch, D. M., Donnelly, J. M., Ollinger, J. M., Snyder, A. Z., & Mintun, M. A. (2001). Increased amygdala response to masked emotional faces in depressed subjects

resolves with antidepressant treatment: an fMRI study. *Biological psychiatry*, 50(9), 651-658.

Song, C., Ikei, H., Igarashi, M., Miwa, M., Takagaki, M., & Miyazaki, Y. (2014). Physiological and psychological responses of young males during spring-time walks in urban parks. *Journal of physiological anthropology*, 33(1), 1-7.

Ter Huurne, N., Fallon, S. J., van Schouwenburg, M., van der Schaaf, M., Buitelaar, J., Jensen, O., & Cools, R. (2015). Methylphenidate alters selective attention by amplifying salience. *Psychopharmacology*, 232(23), 4317-4323.

Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends in cognitive sciences*, 9(12), 585-594.

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.

Acknowledgements: E.T. was supported by a studentship from The Leverhulme Trust (Forest Edge Project).