Cognitive Therapy and Research (2019) 43:893–909
https://doi.org/10.1007/s10608-019-10017-y

ORIGINAL ARTICLE

Concurrent and Prospective Relations Between Attentional Biases for Emotional Images and Relapse to Depression

Kristin Newman[1] · Leanne Quigley[2] · Amanda Fernandez[1] · Keith Dobson[1] · Christopher Sears[1]

Published online: 30 March 2019

Abstract

This study examined concurrent and prospective associations between attentional biases for emotional images and relapse to depression. Previously depressed ($n = 121$) and never depressed ($n = 28$) women completed an eye-tracking task to measure attentional biases for emotional images (face images and naturalistic images) and were then followed for 6 months to assess for relapse to depression. Participants returned for a follow-up session that included the eye-tracking task after a relapse or after 6 months. Previously depressed women who experienced a relapse to depression during the study period showed the hypothesized pattern of decreased attention to positive images and increased attention to negative images, relative to previously depressed women who did not experience a relapse and never depressed women. This was true at the initial visit for naturalistic images and at the follow-up visit for both face and naturalistic images. Women who relapsed had greater attentional biases for some image types at the follow-up visit (when in a state of relapse) than the initial visit (when in a state of remission). Contrary to hypothesis, non-relapsed previously depressed women did not exhibit attentional biases for emotional images relative to never depressed women. Reduced attention to positive images prospectively predicted relapse to depression among the previously depressed women. The results clarify how attentional biases manifest and change from a remitted to relapsed state and provide preliminary evidence for reduced attention to positive information as a risk factor for depression recurrence.

Keywords Depression · Relapse · Recurrence · Attentional biases · Eye-tracking

Introduction

Depression is a highly recurrent disorder. The majority of individuals with depression experience multiple episodes (Solomon et al. 2000), and those who have had more than one episode of depression may experience five to nine additional episodes during their lives (Burcusa and Iacono 2007). The recurrent nature of depression suggests that after remission of a depressive episode, vulnerability factors persist and maintain risk for future episodes. Effective relapse prevention therefore requires that such vulnerability factors be identified and targeted.

Cognitive models of depression propose that biased attention is a key vulnerability factor (Gotlib and Joormann 2010). Several studies have found that depressed individuals attend to sad information more and to positive information less than never depressed individuals (for reviews see Armstrong and Olatunji 2012; Gotlib and Joormann 2010; Peckham et al. 2010; Yiend 2010). According to cognitive models, attentional biases for emotional information are not merely symptoms of depression but may also be an important trait-like cognitive vulnerability factor for depression recurrence (Gotlib and Joormann 2010). It has been proposed that attentional biases may lead to an increase in negative affect and decrease in positive affect, ultimately resulting in long-term distortions in cognition that contribute to the onset, maintenance, and recurrence of depression (Armstrong and Olatunji 2012; DeRaedt and Koster 2010).

✉ Christopher Sears
sears@ucalgary.ca

[1] Department of Psychology, University of Calgary, 2500 University Drive N.W., Calgary, AB T2N 1N4, Canada

[2] Ferkauf Graduate School of Psychology, Yeshiva University, Bronx, NY, USA

Attentional Biases in Previously Depressed Individuals

Relative to the larger body of research that has documented attentional biases in currently depressed samples (Armstrong and Olatunji 2012; Gotlib and Joormann 2010; Peckham et al. 2010; Yiend 2010), much less is known about biased attention in previously depressed samples. An important question regarding attentional biases in these individuals is whether attentional biases are trait-like characteristics that contribute to susceptibility to future depressive episodes, as opposed to a state-effect present only during an active episode (DeRaedt and Koster 2010; Gotlib and Joormann 2010).

Several studies have used cross-sectional designs to investigate the attentional biases of previously depressed individuals at a single point in time, and to determine if previously depressed individuals attend to emotional stimuli similarly to never depressed or currently depressed individuals (e.g., Elgersma et al. 2018; Isaac et al. 2014; Newman and Sears 2015; Sears et al. 2011; Soltani et al. 2015). The results of these studies suggest that biased attention for emotional information persists in remitted depression. However, there have been inconsistencies with respect to the nature of the biases observed in previously depressed individuals. A few studies have found that previously depressed individuals selectively attend to sad or depression-related stimuli similar to currently depressed individuals (e.g., Joormann and Gotlib 2007; Newman and Sears 2015; Soltani et al. 2015). Some studies have also found evidence of the reduced attention to positive stimuli that is characteristic of currently depressed individuals (e.g., Joormann and Gotlib 2007; Sears et al. 2011; Soltani et al. 2015). On the other hand, several studies have not found evidence for either increased attention to sad stimuli (e.g., Elgersma et al. 2018; Isaac et al. 2014) or decreased attention to positive stimuli (e.g., Elgersma et al. 2018; Isaac et al. 2014; Woody et al. 2016) in previously depressed individuals.

Cross-sectional designs are limited in their ability to delineate the nature and role of attentional biases as a factor that contributes to depression vulnerability. Potential clinically significant differences between samples (e.g., number of prior depressive episodes, treatment history) and lack of a temporal relationship between variables make it difficult to draw conclusions about the relationship between attentional biases and depression recurrence. These limitations highlight the need for longitudinal research that examines changes in attention over time in the same individuals. Longitudinal designs allow researchers to study how attentional biases manifest from the period of remission to relapse in the same individuals and can therefore elucidate the specific changes in biases that are associated with, and may contribute to, relapse.

Attentional Biases as Predictors of Relapse

Prospective longitudinal research on attentional biases as predictors of depression recurrence is also limited. Beevers and Carver (2003) tested whether attentional biases for word stimuli in a dot-probe task predicted an increase in depressive symptoms in a mixed group of never depressed and previously depressed individuals. They found that change in biases for negative words following a negative mood induction interacted with self-reported occurrence of stressful life events to predict increased depressive symptoms 7 weeks later. Baseline attentional biases for negative words (i.e., prior to the negative mood induction) and both baseline and change in attentional biases for positive words did not predict depressive symptoms. In contrast, in a large sample of previously depressed individuals, Elgersma et al. (2019) did not find attentional biases for positive or negative words (as measured by the exogenous cueing task) to be predictive of either relapses of major depressive disorder (MDD) or severity of depression symptoms at 2- or 4-year follow-ups.

Woody et al. (2016) used a dot-probe task in combination with eye-tracking to examine attentional biases in previously and never depressed individuals. Pairs of face images (a neutral face paired with an angry, happy, or sad face) were displayed for 1000 ms during the dot-probe trials, and eye gaze was tracked simultaneously. Greater attention to angry faces, as measured by eye-tracking at the initial session, predicted a shorter time to depression recurrence over a 2-year follow-up, whereas response latencies to angry and sad faces in the dot-probe task at the initial session did not predict depression recurrence. Thus, the results from the eye-tracking task suggested that heightened attention to angry faces may be a stable risk factor for depression recurrence, but the results from the dot-probe task did not support this conclusion. Woody et al. argued that the discrepancy in the dot-probe and eye-tracking results could be related to the psychometric properties of these tasks; studies have reported an absence of a relationship between these two attentional bias indices, as well as the superior reliability of eye-tracking indices (e.g., Lazarov et al. 2018; Sanchez et al. 2017; Sears et al. 2018; Waechter et al. 2014). Consistent with this possiblity, in the Elgersma et al. (2019) study that reported null predictive effects of attentional biases on relapses and depression severity, the internal consistency of the attentional bias indices from the exogenous cueing task (a reaction-time based task similar to the dot-probe task) was reported to be near zero (estimates were between −0.13 and 0.004).

The Present Study

The present study used a prospective longitudinal design to examine attentional biases as a cognitive vulnerability factor in depression relapse. Previously depressed and never depressed participants completed an eye-tracking task to assess attentional biases for emotional images. Participants were then followed for 6 months, during which they were assessed every 2 weeks for the emergence of depressive symptoms. Participants attended a follow-up visit either if they reported clinically significant depressive symptoms, or at the end of the 6-month study period, whichever occurred first. At the follow-up visit, participants again completed an eye-tracking task, but with a different set of images. Based on their clinical status during the study follow-up period, previously depressed participants were re-categorized into a relapsed group or a non-relapsed group. The study evaluated: (1) whether attention to emotional images differs in previously depressed individuals who go on to experience a relapse compared to those who do not, as well as compared to individuals with no depression history, both before and following relapse, (2) whether there are changes in attention in previously depressed individuals from the period of remission to relapse, and (3) whether attentional biases predict future relapse to depression.

Based on the theoretical and empirical literature that suggests that depression-vulnerable individuals show increased attention for depression-related stimuli and reduced attention for positive stimuli, it was hypothesized that previously depressed participants (both relapsed and non-relapsed) would attend more to the sad images and less to the positive images than never depressed participants, at both the initial and follow-up visits (Hypothesis 1). It was further hypothesized that this pattern of biased attention would be more pronounced for the relapsed group relative to the non-relapsed group. This hypothesis is based on the idea that previously depressed individuals who go on to relapse may have greater cognitive vulnerability, as manifested by stronger attentional biases when attending to emotional information, than previously depressed individuals who do not go on to relapse. In addition, it was hypothesized that the relapsed group would attend more to the sad images and less to the positive images at the follow-up visit than the initial visit, due to their current depressive symptomatology at the follow-up visit (Hypothesis 2). The non-relapsed and never depressed participants were not expected to attend to these images differently between the initial and follow-up visits. Finally, for the previously depressed participants, it was hypothesized that greater attention to sad images and reduced attention to positive images at the initial visit would be predictive of relapse to depression during the follow-up period (Hypothesis 3).

Method

Participant Recruitment and Eligibility

Two groups of participants were recruited: a previously depressed group and a never depressed group. Recruitment prioritized previously depressed participants to permit evaluation of attentional biases as prospective predictors of relapse to depression. Only female participants were recruited due to the documented gender differences in the prevalence, etiology, course, and symptomatology of depression (e.g., Kessler et al. 1993; Piccinelli and Wilkinson 2000), as well as evidence of gender differences in the processing of emotional stimuli (Donges et al. 2012; Kemp et al. 2004; Montagne et al. 2005). Participants were recruited through a variety of methods, including an internet-based research participation system, media advertisements, and community flyer postings.

Group definitions and eligibility were based on diagnostic criteria from the *Diagnostic and Statistical Manual of Mental Disorders—Fourth Edition* (DSM-IV; APA 2000) and determined with the Structured Clinical Interview for DSM-IV Axis I disorders (SCID-I; First et al. 1995). Participants in the previously depressed group met the diagnostic criteria for past, but not current (over the past month), major depressive disorder (MDD). Participants in the never depressed group had no current or previous history of MDD. Exclusion criteria for both groups included a history of or current presentation of psychosis or mania, current MDD, or current alcohol or drug abuse/addiction.

Measures

Diagnosis

The SCID-I was administered by master's-level students in clinical psychology and trained undergraduate research assistants to determine eligibility and group assignment. The SCID-I is a commonly used semi-structured interview that determines the presence of current and lifetime Axis I disorders based on diagnostic criteria outlined in the DSM-IV (SCID-I; First et al. 1995). Interrater reliability has been determined to be in the moderate to excellent range for Axis I disorders (Lobbestael et al. 2011), and inter-rater reliabilities for current versus lifetime diagnosis have ranged from $r = 0.75$ to $r = 0.85$ (Hook et al. 2007).

Depressive Symptoms

The Beck Depression Inventory-Second Edition (BDI-II; Beck et al. 1996) was used to assess depression symptom severity. The BDI-II consists of 21 items rated on a scale from zero to three. Total scores range from 0 to 63 and can categorize minimal depression (0–13), mild depression (14–19), moderate depression (20–28), and severe depression (29–63), according to the criteria of Beck et al. (1996). The BDI-II has excellent internal consistency and test–retest reliability (Beck et al. 1996).

Participants also completed the Patient Health Questionnaire-9 (PHQ-9; Spitzer et al. 1999), a 9-item measure of depression symptomatology based on the DSM-IV diagnostic criteria for MDD. Items are rated from zero ("*Not at all*") to three ("*Nearly every day*"). Total scores range from 0 to 27, where scores of 5–9 indicate minimal symptoms of depression, scores of 10–14 indicate minor/mild major depression, scores of 15–19 indicate moderately severe major depression, and scores greater than 20 indicate severe major depression. Categorical scoring can also be used to establish a diagnosis of MDD, based on meeting one of the two cardinal criteria (depressed mood, anhedonia), as well as endorsing at least four of the other criteria. The PHQ-9 can be used to assess lifetime history of depression by modifying the respondent instructions from: "Over the last 2 weeks, how often have you been bothered by any of the following problems?" to "For the 2 weeks in your life that you were the most blue, sad, or depressed, how often were you bothered by any of the following problems?" (PHQ-9L; Cannon et al. 2007). Both the current (PHQ-9) and lifetime (PHQ-9L) versions were used in the present study.

Eye Tracking Task

Apparatus

Attention to emotional images was assessed using eye gaze tracking and recording. Eye movements were recorded using an EyeLink 1000 eye-tracking system (SR Research Ltd., Ottawa, ON), which uses infrared video-based tracking technology. The system has a 1000 Hz sampling rate (allowing for a temporal resolution of 2 ms) and a typical gaze accuracy of 0.25–0.50 degrees of visual angle. Stimuli were shown on a 21-inch monitor positioned approximately 60 cm away from the participant. Participants used a chin rest to minimize head movements and increase tracking accuracy.

Protocol

During each of the two laboratory visits, participants' eye movements were tracked and recorded while they viewed sets of images presented on a computer display. Participants were shown two separate blocks of 30 trials (one block with face image stimuli and one block with naturalistic image stimuli), with a short rest period between blocks. The order of presentation of these blocks was randomized across participants. Different face and naturalistic images were shown at the initial and follow-up visits, to prevent familiarity with the images from affecting participants' viewing behavior.

Four images were presented on each trial, with one image presented in each of the four quadrants of the computer display. The images were randomly assigned to the four display locations, and each image type was equally likely to appear in each quadrant across all trials. At the start of each trial the participants fixated on a black dot in the centre of the display to ensure accurate gaze measurement. The images were then presented for 8 s, and participants' eye gaze was tracked and recorded while they examined the images. Participants were instructed to view the images as they wished, similar to other free-viewing paradigms (e.g., Arndt et al. 2014; Kellough et al. 2008). Each block of trials was approximately 12 min in duration.

Stimuli

Previous studies of depression-related attentional biases have typically used either face images or naturalistic images as stimuli. As differences in stimuli could contribute to the variability in previously reported findings, in the present study we examined attention to both image types, which allowed us to evaluate the consistency of the results across the image types.

The face images consisted of 240 photographs of male and female faces from the NimStim Facial Expressions database (Tottenham et al. 2009). Four different types of facial expressions were used: sad, threat, happy, and neutral. There were 60 images of each type of face expression, and each category was comprised of an equal number of faces with open and closed mouths. The threat image category was comprised of an equal number of angry and frightened faces, as both were deemed to potentially evoke a threat-related response.

The naturalistic images consisted of 240 images depicting people, places, and objects. There were 60 of each of four image types, selected to correspond with the face image categories: sad, threat, positive, and neutral. The sad images included scenes of people appearing sad and unhappy, neglected animals, scenes of poverty and illness, and gloomy landscapes. The threat images included themes of threat and injury, such as people being threatened with weapons, people with physical injuries (e.g., a burn on an arm), dangerous situations (e.g., a person walking along a cliff), vehicle accidents, and menacing animals. The positive images showed people smiling and laughing, children playing, rabbits and kittens, and vacation activities and destinations (e.g., a beach at a tropical resort). The neutral images included people in various activities and had no obvious positive or negative theme (e.g., a woman talking on the telephone; a group of people having a meeting), as well as pictures of objects (e.g., a bicycle, a computer) and a variety of landscapes (e.g., office buildings). To identify an optimal set of images, prior to the study an online survey was used to collect ratings for a large number of images. For each image, undergraduate student participants chose one of four categories that best described the image: (1) sad/depressing/gloomy, (2) threatening/dangerous/fearful, (3) positive/happy, or (4) neutral/no emotion. Participants also provided a valence rating using a 11-point scale, with -5 representing "very negative" and $+5$ representing "very positive", with a midpoint of zero representing "neutral." Each of the 240 images was rated by at least 44 participants (range 44–54) and an image was used in the present study only if at least 80% of the raters agreed to its category. The mean valence ratings for the sad, threat, positive, and neutral images were -2.72 (SD 0.66), -2.84 (SD 0.77), 2.71 (SD 0.47), and 0.08 (SD 0.13), respectively. The mean valence ratings were significantly different from one another ($p < 0.001$), with the exception of the difference between the ratings for sad images and threat images, $t(118) = 1.12$, $p = 0.263$.

Procedure

Prospective participants first completed an online screening measure including the BDI-II, PHQ-9, and PHQ-9L. Participants who appeared to meet the study criteria based on the online screening were invited to participate in the study. Participants attended two separate lab visits. In the initial visit, participants were administered the SCID-I by a trained interviewer to assess for inclusion and exclusion criteria and to establish diagnoses. Participants completed a computerized battery of self-report questionnaires (which again included the BDI-II, PHQ-9, PHQ-9L) and other measures unrelated to the present hypotheses (e.g., measures of anxiety, emotion regulation), and then completed the eye tracking task.

Participants were contacted every 2 weeks by phone for 6 months following their initial study visit. Depression status was assessed using the BDI-II and the PHQ-9 at each contact. Participants who scored > 13 on the BDI-II or who met the categorical criteria for depression on the PHQ-9 were scheduled to return to the lab for the follow-up visit as soon as possible. Participants who did not score > 13 on the BDI-II or meet the categorical criteria for depression on the PHQ-9 during any of the phone contacts were scheduled for a follow-up visit at the end of the 6-month period. At the follow-up visit, participants were re-administered the SCID-I to confirm the presence or absence of a major depressive episode (MDE), as well as to assess for the presence of exclusion criteria since the initial visit. Participants then completed a battery of self-report questionnaires and the eye tracking task. Previously depressed participants who were diagnosed with an MDE during the follow-up period were re-categorized as a member of the relapsed group. Previously depressed participants who did not experience an MDE during the follow-up period were re-categorized as a member of the non-relapsed group. Never depressed participants remained categorized as such, as none developed an MDE during the study period. Participants received a \$25 CAN gift card for each laboratory visit.

Data Processing

The fixation data were processed using the EyeLink Data Viewer analysis software (SR Research) to filter for blinks, missing data, and other recording artifacts (using the default settings). To be included in the analyses, a fixation had to be at least 100 ms in duration; sequential, adjacent fixations that were less than 100 ms were merged into a single fixation. The dependent variable was the total dwell time for each image type (in ms) during the 8 s presentation. Total dwell time was computed for each image type on a trial-by-trial basis by summing the individual fixation times to each image over the 8 s presentation and then averaging over the 30 trials.

Data Analysis Plan

To simplify interpretation, data from the face and naturalistic images were analyzed separately. To test Hypothesis 1, we examined group differences in the total dwell times for face images and naturalistic images using separate 3 (Group: relapsed, non-relapsed, never depressed) × 4 (Image Type: sad, threat, happy/positive, neutral) mixed-model ANOVAs of dwell times at the initial and follow-up visits. One-way ANOVAs and Tukey-corrected pairwise comparisons were used to follow up statistically significant main effects and interactions. These analyses allowed us to determine if the groups differed from each other in their attention to the images at the initial and follow-up visits, and specifically, to test whether previously depressed participants exhibit attentional biases relative to never depressed participants, and whether biases are more pronounced for relapsed versus non-relapsed previously depressed participants. Next, the total dwell time data for each group was analyzed separately using a 2 (Visit: initial, follow-up) × 4 (Image type: sad, threat, happy/positive, neutral) repeated measures ANOVAs to evaluate whether total dwell times to the face and naturalistic images changed from the initial to follow-up visits. Significant interactions were followed up with paired samples *t*-tests to test for differences in total dwell time between the initial and follow-up visit for each image type separately. These analyses tested Hypothesis 2, that the relapsed group would attend more to the sad images and less to the positive images at the follow-up visit than the initial visit, whereas the never depressed and non-relapsed groups would not exhibit significant changes in attention to the images between the two visits.

Finally, we used hierarchical logistic regression analyses with bootstrapping ($n = 5000$) to test Hypothesis 3, that attention to the images at the initial visit would predict relapse status during the follow-up period among the previously depressed participants ($N = 121$).[1] Relapse status (relapsed versus non-relapsed) was the dependent variable in each regression model, while standardized total dwell times for each image type at the initial visit were the predictors. Standardized dwell times were used instead of raw fixation times for these analyses to facilitate interpretation of the parameters. Guidelines for logistic regression suggest that 10 to 15 events per predictor variable is desired for a stable model (Babyak 2004). Given that only 14 participants relapsed, the number of predictors that could be entered simultaneously into a model was limited. Therefore, total dwell times for each image type at the initial visit were analyzed in separate regression models. The total dwell times for each image type were entered in the first step of each model, with initial BDI-II score entered in the second step, to test whether total dwell times were predictive of relapse after controlling for depressive symptoms present at the initial visit. Given the relatively small number of relapses ($n = 14$) observed, we conducted additional hierarchical linear regression models to evaluate whether total dwell times for each image type at the initial visit predicted depression symptoms at the second visit. Similar to the logistic regressions, standardized total dwell times for each image type at the initial visit were analyzed in separate regression models, with the dwell times entered in the first step and initial BDI-II scores entered in the second step of the models. BDI-II score at the second study visit was the dependent variable in each model.

Results

Participant Characteristics

A total of 3155 women completed an online survey to determine eligibility for the study (2797 students and 358 community members). Eligibility was based on meeting criteria for a history of depression or no history of depression as per the online measures (PHQ-9, PHQ-9L, and BDI-II). A sample of 289 eligible participants attended the initial study visit; of them, 140 were excluded for the following reasons: (a) current or prior mania, current substance abuse, current depression, psychosis, relocation out of the city, did not meet depression criteria as per the SCID-I ($n = 62$); (b) inability to collect eye-tracking data due calibration errors ($n = 18$); (c) attrition/drop out ($n = 45$); and (d) failure to identify depression symptoms during phone screenings, but endorsement of criteria for an MDE during the study period at the follow-up visit ($n = 15$).[2] The final sample ($N = 149$) consisted of 28 never depressed participants and 121 previously depressed participants who were in remission at baseline. Among the previously depressed participants, 14 relapsed during the study period (relapsed group), and 107 maintained remission (non-relapsed group).[3] For the 14 participants who relapsed, the mean time to relapse was 105.14 days (SD 58.78; range = 25–205). Time to relapse was not significantly correlated with any of the dwell time indices for face and naturalistic images at the initial or follow-up visit, or with changes in dwell times from the initial to follow-up visit.

Table 1 presents the characteristics of the participant groups. There were some differences in age, ethnicity, and marital status of the previously depressed (relapsed and non-relapsed) and never depressed participants, which was due to a greater percentage of students versus community members in the never depressed sample (89% students) relative to the previously depressed sample (44% students). Relapsed participants were significantly older and more likely to be divorced or separated than never depressed participants, and non-relapsed participants were more likely to be Caucasian and less likely to be never married than never depressed participants (Table 1). Relapsed and non-relapsed participants did not differ significantly on any demographic variable. As expected, never depressed participants had a lower mean BDI-II score at the initial visit than the relapsed and non-relapsed participants. At the follow-up visit, relapsed participants had the highest mean BDI-II score, followed by the non-relapsed participants, and then the never depressed participants. Relapsed and non-relapsed participants did not differ significantly in the number of prior depressive episodes reported.

[1] Three participants were missing eye tracking data for the face images and four participants were missing eye tracking data for the naturalistic images, resulting in samples sizes for the regression analyses of $N = 118$ and $N = 117$ for the face images and naturalistic images, respectively.

[2] Previously depressed participants who dropped out or were lost to follow-up did not differ from those who attended both study visits on age or depression symptoms at the initial visit (BDI-II, PHQ-C, PHQ-L), all $p > 0.10$.

[3] For the face image set, three participants from the non-relapsed group were missing data (due to technical issues with the eye-tracking system or calibration errors), resulting in a sample of 28 never depressed, 104 non-relapsed, and 14 relapsed participants for the analyses of dwell times for the face images. For the naturalistic image set, one participant from the never depressed group and four participants from the non-relapsed group were missing data (for the same reasons), resulting in a sample of 27 never depressed, 103 non-relapsed, and 14 relapsed participants for the analyses of dwell times for the naturalistic images.

Table 1 Participant characteristics and questionnaire scores for the relapsed, non-relapsed, and never depressed groups at the initial visit (baseline) and follow-up visit

	Relapsed ($n = 14$)	Non-relapsed ($n = 107$)	Never depressed ($n = 28$)
Age in years, M (SD)	29.5 (12.1)$_a$	29.3 (12.2)$_{a,b}$	22.7 (10.6)$_b$
Ethnicity			
Caucasian	79%$_{a,b}$	76%$_a$	46%$_b$
Marital status			
Married or living with someone as if married	21%$_a$	22%$_a$	11%$_a$
Divorced or separated	21%$_a$	13%$_{a,b}$	0%$_b$
Never married	57%$_{a,b}$	65%$_a$	89%$_b$
Current mood (+5 to − 5)	1.4 (2.2)$_a$	2.5 (1.7)$_{a,b}$	2.8 (1.5)$_b$
BDI-II—initial visit	10.7 (7.3)$_a$	7.5 (6.1)$_a$	1.7 (2.4)$_b$
BDI-II—follow-up visit	25.3 (10.3)$_a$	7.8 (7.0)$_b$	1.1 (1.6)$_c$
PHQ-C9—initial visit	6.2 (4.9)$_a$	4.0 (3.5)$_a$	0.78 (1.1)$_b$
PHQ-9—follow-up visit	15.5 (4.6)$_a$	4.2 (4.1)$_b$	0.68 (1.0)$_c$
PHQ-9L	21.5 (7.0)$_a$	19.5 (5.5)$_a$	1.2 (2.2)$_b$
# Previous MDEs	8.0 (8.0)$_{a,b}$	5.7 (12.8)$_a$	0 (0)$_b$

Standard deviations in parentheses. Means in the same row that do not share a subscript differ at $p < 0.05$ based on Tukey pairwise comparisons following a significant one-way ANOVA or based on Chi square tests, as appropriate

Current mood mood rating on a 11-point scale, from very positive (+5) to very negative (−5), BDI-II Beck depression inventory-II, *PHQ-9* patient health questionnaire (current). *PHQ-9L* patient health questionnaire (lifetime), *# previous MDEs* number of previous episodes of depression

Table 2 Mean total dwell times (ms) for the face images and naturalistic images for the relapsed, non-relapsed, and never depressed groups, at the initial and follow-up visits

Image type	Relapsed		Non-relapsed		Never depressed	
	Initial	Follow-up	Initial	Follow-up	Initial	follow-up
Face images						
Sad	1533 (94)	1624 (94)	1411 (34)	1373 (34)	1377 (66)	1332 (67)
Threat	1483 (80)	1564 (98)	1388 (29)	1417 (36)	1454 (56)	1445 (69)
Happy	1960 (183)	1384 (210)	2054 (67)	2109 (77)	2151 (130)	2133 (149)
Neutral	1506 (70)	1558 (97)	1564 (25)	1564 (35)	1508 (49)	1542 (69)
Naturalistic images						
Sad	1953 (126)	2068 (136)	1780 (46)	1576 (50)	1579 (91)	1372 (97)
Threat	1870 (126)	2151 (141)	1656 (46)	1764 (52)	1677 (90)	1798 (101)
Positive	1588 (174)	1375 (220)	2020 (64)	2133 (81)	2193 (125)	2190 (158)
Neutral	881 (83)	754 (92)	1019 (30)	980 (34)	1100 (60)	1072 (66)

Standard errors in parentheses. One participant in the non-relapsed group and one participant in the never depressed group were missing data for the naturalistic images

Hypothesis 1: Group Differences in Total Dwell Time at the Initial and Follow-Up Visits

Mean total dwell times for the face images and the naturalistic images at the initial and follow-up visits are listed in Table 2 and shown in Figs. 1 and 2.

Analysis of the Face Images Data

For the face images at the initial visit, there was a main effect of Image Type, $F(3, 429) = 28.95$, $p < 0.001$, partial $\eta^2 = 0.17$, such that participants had longer dwell times to the happy faces than the other face types (threat, sad, and

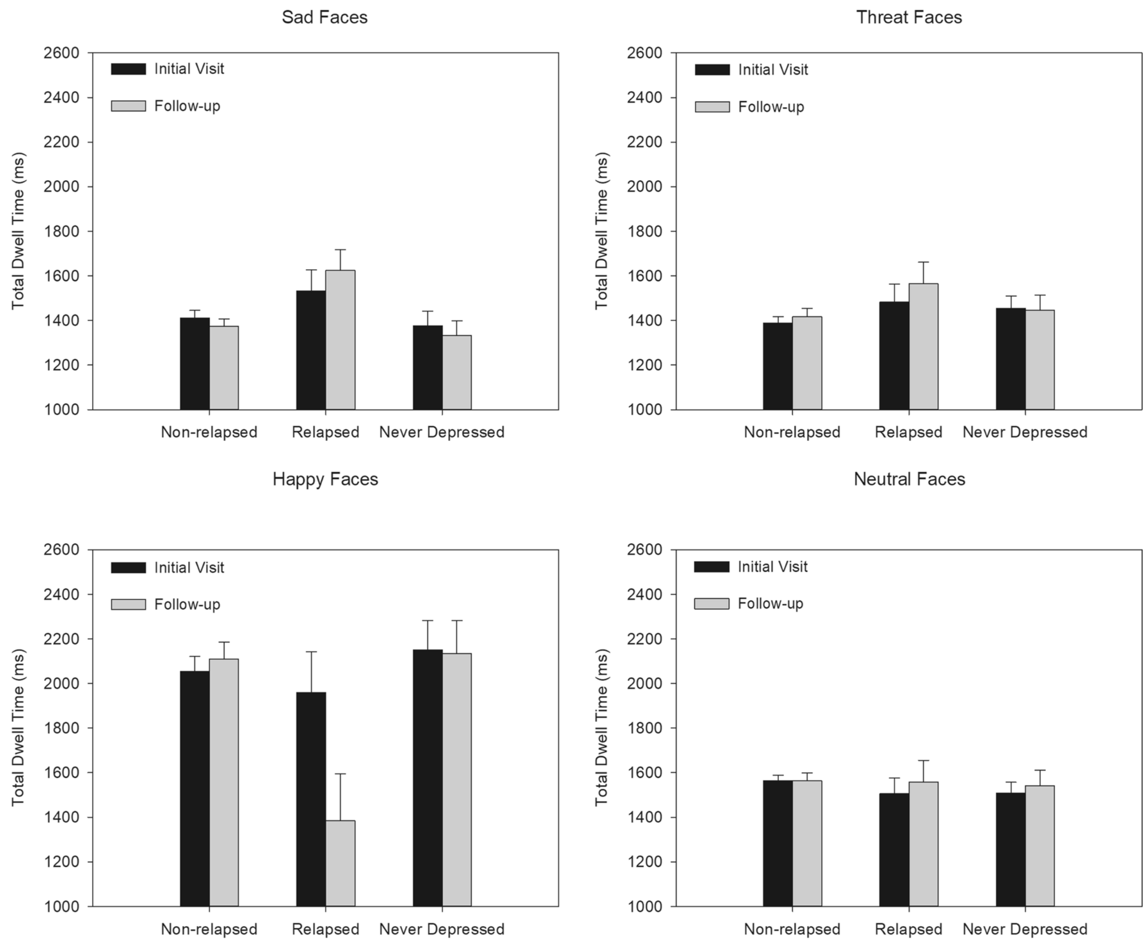

Fig. 1 Total dwell times for the face images at the initial and follow-up visits for non-relapsed, relapsed, and never depressed participants. Error bars represent one standard error

neutral). There were no other significant effects in the analysis, including the Group × Image Type interaction ($p > 0.10$). This indicated that the relapsed, non-relapsed, and never depressed participants did not differ significantly in how they attended to the faces at the initial visit (see Fig. 1).

At the follow-up visit, there was a main effect of Image Type, $F(3, 429) = 9.57$, $p < 0.001$, partial $\eta^2 = 0.06$, which was qualified by a significant Group × Image Type interaction, $F(6, 429) = 4.01$, $p = 0.001$, partial $\eta^2 = 0.05$, reflecting group differences in attention to the images. The interaction was explored using one-way ANOVAs, which indicated that the groups differed in their dwell times to sad faces, $F(2, 143) = 3.56$, $p = 0.03$, partial $\eta^2 = 0.05$, and happy faces, $F(2, 143) = 5.44$, $p = 0.005$, partial $\eta^2 = 0.07$, but not to threat faces or neutral faces (ps > 0.10).[4] Relapsed participants attended to sad faces significantly more than both non-relapsed participants (mean difference = 251.60 ms, $SE = 100.87$, $p = 0.036$), and never depressed participants (mean difference = 292.23 ms, $SE = 115.98$, $p = 0.034$). There was no significant difference between non-relapsed and never depressed participants in their dwell times to sad faces (mean difference = 40.63 ms, $SE = 75.44$, $p = 0.852$). Relapsed participants had significantly shorter total dwell times to happy faces than both non-relapsed participants (mean difference = 725.23 ms, $SE = 224.46$, $p = 0.004$), and never depressed participants (mean difference = 749.37 ms, $SE = 258.08$, $p = 0.012$), which reflected their reduced attention to happy faces. There was no significant difference between non-relapsed participants and never depressed participants in their total dwell times to happy faces (mean difference = 24.14 ms, $SE = 167.87$, $p = 0.989$).

[4] Levene's test of homogeneity of variances was significant for the happy faces, $p < 0.05$. The Brown-Forsythe robust test confirmed that the groups differed in their dwell times for happy faces, $F(2, 69.29) = 9.87$, $p < 0.001$.

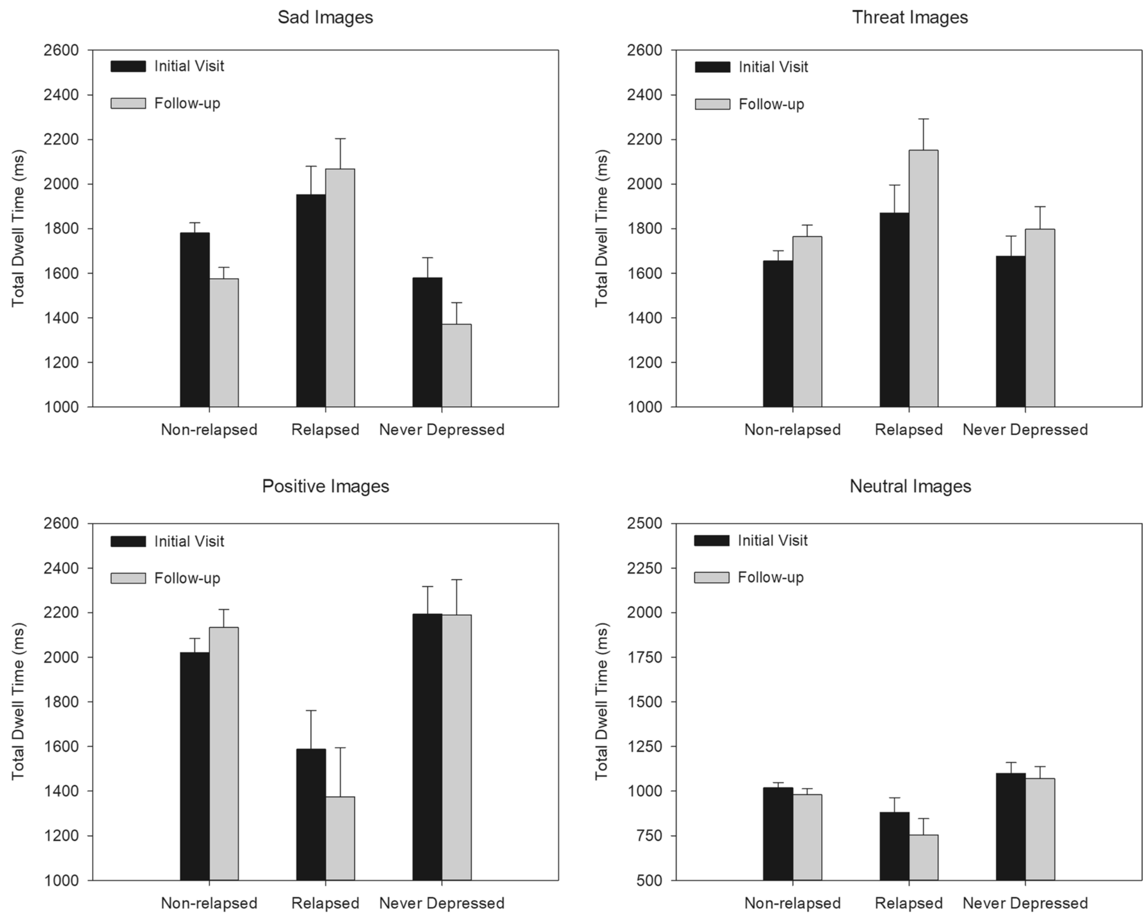

Fig. 2 Total dwell times for the naturalistic images at the initial and follow-up visits for non-relapsed, relapsed, and never depressed participants. Error bars represent one standard error

Analysis of the Naturalistic Image Data

The results for the naturalistic images are shown in Fig. 2. For the naturalistic images at the initial visit, there was a main effect of Image Type, $F(3, 423) = 42.36$, $p < 0.001$, partial $\eta^2 = 0.23$, which was qualified by a significant Group × Image Type interaction, $F(6, 423) = 3.07$, $p = 0.006$, partial $\eta^2 = 0.04$. The interaction was followed up using one-way ANOVAs, which indicated that the groups differed in their dwell times for sad images, $F(2, 141) = 3.22$, $p = 0.043$, partial $\eta^2 = 0.04$, and positive images, $F(2, 141) = 4.02$, $p = 0.020$, partial $\eta^2 = 0.05$, but not for threat or neutral images ($ps > 0.10$).[5] For sad images, relapsed participants had significantly longer total dwell times than never depressed participants (mean difference = 373.81 ms, $SE = 156.09$, $p = 0.047$). The difference between relapsed participants and non-relapsed participants was not significant (mean difference = 172.74 ms, $SE = 135.00$, $p = 0.409$), nor was the difference between non-relapsed and never depressed participants (mean difference = 201.07 ms, $SE = 102.47$, $p = 0.125$). For positive images, relapsed participants had dwell times that were marginally shorter than non-relapsed participants (432.46 ms, $SE = 185.61$, $p = 0.055$), and significantly shorter than never depressed participants (mean difference = 605.25 ms, $SE = 214.60$, $p = 0.015$), which reflected their reduced attention to positive images. There was no significant difference between non-relapsed and never depressed participants (mean difference = 172.79 ms, $SE = 140.88$, $p = 0.440$).

For the follow-up visit, there was a main effect of Image Type, $F(3, 423) = 37.46$, $p < .001$, partial $\eta^2 = 0.21$, which was qualified by a significant Group × Image Type interaction, $F(6, 423) = 5.80$, $p < 0.001$, partial $\eta^2 = 0.08$. Additional analyses revealed similarities and differences from the pattern of group differences observed during the initial visit. The ANOVAs indicated that the groups differed in their dwell

[5] Levene's test of homogeneity of variances was significant for the positive images, $p < .05$. The Brown-Forsythe robust test confirmed that the groups differed in their dwell times for positive images, $F(2, 56.67) = 5.41$, $p < .01$.

times for sad images, $F(2, 141) = 8.69$, $p < 0.001$, partial $\eta^2 = 0.11$, and positive images, $F(2, 141) = 5.57$, $p = 0.005$, partial $\eta^2 = 0.07$, as was the case during the initial visit. For sad images, relapsed participants had significantly longer total dwell times than never depressed participants (mean difference = 696.18 ms, $SE = 167.64$, $p < 0.001$), which was also the case during the initial visit. Relapsed participants also had significantly longer total dwell times for sad images than non-relapsed participants (mean difference = 492.33 ms, $SE = 144.99$, $p = 0.003$), which was not the case during the initial visit. Non-relapsed and never depressed participants did not differ significantly in their dwell times to sad images (mean difference = 203.84 ms, $SE = 110.06$, $p = 0.157$), which was also true during the initial visit. For positive images the group differences at the follow-up visit were similar to those observed at the initial visit: Relapsed participants had significantly shorter total dwell times than both never depressed participants (mean difference = 815.87 ms, SE = 271.57, $p = .009$) and non-relapsed participants (mean difference = 758.72 ms, $SE = 234.88$, $p = .004$), whereas the non-relapsed and never depressed participants did not differ significantly (mean difference = 57.16 ms, $SE = 178.28$, $p = 0.945$).

Unlike the initial visit, the groups also differed in their attention to threat and neutral images at the follow-up visit, $F(2, 141) = 3.30$, $p = 0.040$, partial $\eta^2 = 0.05$, and $F(2, 141) = 3.90$, $p = 0.022$, partial $\eta^2 = 0.05$, respectively. For threat images, the relapsed participants had significantly longer total dwell times than non-relapsed participants (mean difference = 387.52 ms, $SE = 150.88$, $p = 0.030$). The differences in dwell times for threat images were non-significant between relapsed and never depressed participants (mean difference = 353.56 ms, $SE = 174.45$, $p = 0.110$) and between non-relapsed and never depressed participants (mean difference = 33.97 ms, $SE = 114.53$, $p = 0.953$) For neutral images, the relapsed participants' total dwell times were significantly shorter than those of never depressed participants (mean difference = 317.84 ms, $SE = 114.36$, $p = 0.017$), but did not differ significantly from non-relapsed participants (mean difference = 226.92 ms, $SE = 98.91$, $p = 0.060$). Non-relapsed and never depressed participants also did not differ significantly in their dwell times for neutral images (mean difference = 90.92 ms, $SE = 75.08$, $p = 0.449$). Taken together, these analyses indicate that relapsed participants differed from non-relapsed and never depressed participants in their attention to naturalistic the images at both visits (particularly the sad and positive images, for which group differences were observed at both visits).

Hypothesis 2: Changes in Total Dwell Time from the Initial to Follow-Up Visit

Analysis of the Face Images Data

For the face images, there was no significant interaction between Visit and Image Type for non-relapsed participants, $F(3, 309) = 0.60$, $p = 0.613$, partial $\eta^2 = 0.01$, or never depressed participants, $F(3, 81) = 0.13$, $p = 0.941$, partial $\eta^2 = 0.01$, which indicated that there were no significant differences in the way these participants attended to the faces at the initial and follow-up visits (Fig. 1). In contrast, there was a significant Visit × Image Type interaction for relapsed participants, $F(3, 39) = 6.25$, $p = 0.001$, partial $\eta^2 = 0.33$. Total dwell times for happy faces decreased at the follow-up visit relative to the initial visit (mean difference = 576.55 ms, $SE = 141.50$), $t(13) = 4.07$, $p = 0.001$. For threat, sad, and neutral faces there were no significant differences between the initial and follow-up visits (all ps > 0.10). Thus, only the relapsed participants exhibited significant changes in their attention to the images between the initial and follow-up visits, and only for the happy faces.

Analysis of the Naturalistic Image Data

For the naturalistic images, there was no significant interaction between Visit and Image Type for never depressed participants, $F(3, 78) = 2.41$, $p = 0.074$, partial $\eta^2 = 0.09$, which indicated that there were no significant changes in these participants' attention to the images from the initial to the follow-up visit (Fig. 2). In contrast, for relapsed participants, there was a significant Visit × Image Type interaction, $F(3, 39) = 3.38$, $p = 0.028$, partial $\eta^2 = 0.21$. T-tests indicated that total dwell times to threat images were longer at the follow-up visit relative to the initial visit (mean difference = 281.22 ms, $SE = 92.38$), $t(13) = 3.04$, $p = 0.009$. Conversely, for positive images total dwell times were marginally shorter at the follow-up visit than at the initial visit (mean difference = 213.34 ms, $SE = 101.32$), $t(13) = 2.11$, $p = 0.055$. There were no significant differences for sad or neutral images (ps > 0.10). For non-relapsed participants there was also a significant Visit × Image Type interaction, $F(3, 306) = 7.04$, $p < 0.001$, partial $\eta^2 = 0.07$, but the nature of the interaction was different. For threat images, total dwell times were longer at the follow-up visit than at the initial visit (mean difference = 107.90 ms, $SE = 39.64$), $t(102) = 2.72$, $p = 0.008$, similar to the relapsed participants. Unlike the marginal difference observed for relapsed participants, there was no difference for positive images for non-relapsed participants ($p > 0.10$). In addition, for sad images, total dwell times were shorter at the follow-up visit relative to the initial visit (mean difference = 204.33 ms, $SE = 49.60$), $t(102) = 4.12$, $p < 0.001$. There was no difference for neutral images ($p > 0.10$). These results indicate that only the relapsed and non-relapsed participants attended to the naturalistic images differently at the initial and follow-up visits, as shown in Fig. 2.

Hypothesis 3: Attention at the Initial Visit as a Predictor of Relapse and Change in Depression Symptoms

Analysis of the Face Images Data

For the face images, total dwell times for sad, threat, happy, and neutral faces were not predictive of depression relapse (sad faces: $B = 0.35$, $SE = 0.36$, $p = 0.275$; threat faces: $B = 0.33$, $SE = 0.36$, $p = 0.321$; happy faces: $B = -0.15$, $SE = 0.40$, $p = 0.625$; neutral faces: $B = -0.22$, $SE = 0.30$, $p = 0.405$). Total dwell times for face images remained non-significant predictors of relapse after controlling for initial BDI-II score. When predicting depression symptoms (BDI-II scores) at the second visit, dwell times for sad, threat, happy, and neutral faces were not significant predictors (sad faces: $B = 0.50$, $SE = 0.92$, $p = 0.593$; threat faces: $B = -0.06$, $SE = 0.94$, $p = 0.948$; happy faces: $B = -0.16$, $SE = 97$, $p = 0.858$; neutral faces: $B = 0.27$, $SE = 0.80$, $p = 0.732$) and remained non-significant after controlling for initial BDI-II score.

Analysis of the Naturalistic Image Data

For the naturalistic images, total dwell time to positive images significantly predicted relapse ($B = -1.01$, $SE = 0.40$, $p = 0.003$) and remained a significant predictor after controlling for initial BDI-II score ($B = -0.91$, $SE = 0.39$, $p = 0.006$). Thus, reduced attention to the positive images at the initial visit was associated with a greater risk of relapse, after depressive symptoms at the initial visit were accounted for. The odds ratio for the effect was 0.40 (when controlling for initial BDI-II score), which indicates that a one standard deviation decrease in total dwell time on the positive images corresponded to an increase in the odds of relapse of more than two times. Although significant, the confidence interval for the effect was relatively large (95% CI -1.88 to -0.36, corresponding to odds ratios of 0.15 and 0.70), indicating potential variability in the true effect size. Total dwell times for sad, threat, and neutral images did not predict depression relapse either as single predictors (sad images: $B = 0.36$, $SE = 0.30$, $p = 0.181$; threat images: $B = 0.50$, $SE = 0.31$, $p = .076$; neutral images: $B = -0.50$, $SE = 0.39$, $p = 0.143$) or after controlling for initial BDI-II score.

When examining depression symptoms (BDI-II score) at the second visit as an outcome, dwell time to positive naturalistic images was a significant predictor when it was the only predictor in the model ($B = -1.69$, $SE = 0.75$, $p = 0.023$). However, after controlling for initial BDI-II scores, dwell time to positive images no longer significantly predicted BDI-II scores at the second visit ($B = -0.64$, $SE = 0.67$, $p = 0.340$). Total dwell times for sad, threat, and neutral naturalistic images did not predict BDI-II scores at the second visit (sad images: $B = 1.40$; $SE = 0.88$, $p = 0.118$; threat images: $B = 0.49$, $SE = 0.87$; $p = 0.577$; neutral images: $B = -0.12$; $SE = 1.06$, $p = 0.914$) and remained non-significant after controlling for initial BDI-II score.

Discussion

The present study examined concurrent and prospective associations between attentional biases and relapse to depression. Using eye-tracking methodology within a longitudinal study design, we evaluated: (1) attention to emotional images in previously depressed individuals who went on to experience a relapse to depression compared to previously depressed individuals who did not experience a relapse and never depressed individuals, before and after the time of relapse; (2) within-group changes in attention to emotional images from the period of remission to relapse; and (3) the ability of attentional biases to prospectively predict relapse to depression. The study included both face images and naturalistic images to determine if the results would differ for these different image types.

Hypothesis 1: Group Differences in Total Dwell Time at the Initial and Follow-Up Visits

Consistent with Hypothesis 1, relapsed participants generally attended less to happy faces and positive images and more to sad faces and sad images than never depressed participants, although there were some notable differences between the study visits (initial versus follow-up visit) and image types (face images versus naturalistic images). At the initial visit, relapsed participants showed attentional biases for naturalistic images, but not for face images. Relapsed participants attended more to sad naturalistic images than never depressed participants and less to positive naturalistic images than both non-relapsed (marginally) and never depressed participants. At the follow-up visit, relapsed participants attended less to happy faces and more to sad faces, relative to the other participant groups. For the naturalistic images at the follow-up visit, relapsed participants attended more to sad images and less to positive images than both non-relapsed and never depressed participants. In addition, relapsed participants attended more to threat images than non-relapsed participants and less to neutral images than never depressed participants. Thus, relapsed participants generally exhibited the predicted pattern of decreased attention to positive stimuli and increased attention to negative stimuli, although for face stimuli these attentional biases were observed only at the follow-up visit.

It is unclear why attentional biases for face images were not observed at the initial visit. It is possible that face images are less potent stimuli for eliciting emotional biases in attention in previously depressed individuals compared to more complex emotional scenes. As can be seen by comparing the dwell time data shown in Fig. 1 (face images) and 2 (naturalistic images), the group differences were generally similar in direction, but smaller in magnitude, for the face images relative to the naturalistic images. Figure 1 shows that at the initial visit relapsed participants had longer dwell times for sad faces and shorter dwell times for happy faces relative to never depressed participants, and although these differences were not statistically significant, they are consistent with the predicted attentional biases. The attenuated group differences for sad and happy faces relative to positive and sad naturalistic images may be related to another important difference that can be discerned from Figs. 1 and 2; namely, that neutral faces received more attention than neutral naturalistic images. In fact, total dwell times to neutral faces were significantly longer (~ 1500 ms) than total dwell times to neutral naturalistic images (~ 900 ms, $p < 0.001$), which may be explained by research that suggests that neutral faces are not immediately recognized as being emotionally neutral (e.g., Lee et al. 2008; Leppänen et al. 2004). Thus, when participants were viewing the 8-s displays of four images, they attended more to neutral faces than neutral naturalistic images, leaving less time to attend to sad and happy faces than to sad and positive naturalistic images. As a result, group differences in attention to sad and happy faces would be attenuated relative to group differences for sad and positive naturalistic images. If true, this would suggest that attentional biases may be easier to detect in free-viewing tasks that use naturalistic images (when displays include neutral stimuli), which would be an important consideration for future research. To our knowledge this is the first study to examine attentional biases measured using both naturalistic and face images with the same participants, and therefore additional studies will be necessary to replicate and further understand differences in the results between these image types.

When considering the between-group differences across the initial and follow-up visits, the pattern of findings appears to indicate that attentional biases are both a trait- and state-effect. At the follow-up visit, the relapsed participants had experienced a relapse and were in a currently depressed state. The finding that participants attended less to happy faces and more to sad faces in a relapsed state relative to non-relapsed and never depressed participants is consistent with the large body of evidence for these attentional biases in currently depressed individuals (see Armstrong and Olatunji 2010; Gotlib and Joormann 2010; Peckham et al. 2010; Yiend 2010, for reviews). Further, the finding that relapsed participants exhibited attentional biases for naturalistic images, in both remitted and relapsed states, is consistent with some previous studies on attentional biases in remitted depression. The evidence for attentional biases is sparser and less consistent in remitted depression than current depression, however. As noted, a few studies have reported that previously depressed individuals attend more to negative stimuli (e.g., Joormann and Gotlib 2007; Newman and Sears 2015; Soltani et al. 2015) and less to positive stimuli (e.g., Joormann and Gotlib 2007; Sears et al. 2011; Soltani et al. 2015) than never depressed individuals, like currently depressed individuals. However, some studies have found that previously depressed individuals show both similarities and differences in their attention to emotional stimuli compared to currently depressed individuals (e.g., Soltani et al. 2015), or report attentional biases for only depression-related (e.g., Newman and Sears 2015) or positive stimuli (e.g., Sears et al. 2011). Thus, attentional biases in remitted depression appear to be weaker and/or less stable than those in current depression, which is consistent with the notion that biases show both stability and change across depressive states.

It is possible that the attentional biases exhibited by relapsed participants at baseline (when they were in a state of remission), which appear to reflect a trait-like cognitive vulnerability to depression/relapse, in fact reflect state effects.[6] For instance, although the relapsed participants did not meet criteria for current MDD, it is possible that they were at the end or the beginning of an active depressive episode. Relapsed participants may have been more recently remitted and thus experiencing more residual symptoms than non-relapsed participants. Or, relapsed participants may have been in a prodromal phase of a depressive episode and experiencing corresponding cognitive changes. Relapsed and non-relapsed participants did not differ significantly on mood state or depressive symptoms at baseline; however, the non-significant differences between the groups were in the direction of lower mood and higher symptoms in the relapsed group. Thus, we cannot rule out the possibility that the small differences in baseline mood/symptom state accounted for some of the differences in attentional biases observed between the relapsed and non-relapsed participants. In addition, as discussed below, non-relapsed previously depressed participants did not exhibit attentional biases relative to never depressed participants in this study, which may suggest that attentional biases are not an enduring vulnerability factor for a subset of previously depressed individuals.

[6] We thank an anonymous reviewer for suggesting this possibility.

To further delineate trait and state effects, future research should recruit samples of previously depressed individuals who are even more homogenous on baseline mood and depression symptoms than the current sample. However, baseline differences may still emerge between previously depressed individuals who go on to experience a relapse to depression and those who remain remitted. Another possibility is to operationalize depression dimensionally and examine cross-sectional and prospective associations between attentional biases and symptom severity, rather than associations between attentional biases and diagnostic categories. Another approach would to assess time elapsed since last MDE in previously depressed participants, and control for duration of remission at baseline in analyses examining relationships between relapse and attentional biases. Such analyses could test the possibility that individuals who go on to relapse may be more recently remitted, and thus may still be experiencing 'state-like' cognitive consequences of active depression (e.g., attentional biases). A challenge for future studies will be to assess time elapsed since last MDE reliably, as it may be difficult for participants to accurately remember and self-report this data. Studies could use a design in which currently depressed participants are followed to the point of remission, either in a treatment or naturalistic study, and then followed and assessed for relapse. An alternative would be for studies to restrict recruitment to previously depressed participants who have experienced an MDE in the prior 1–2 years to increase the likelihood of accurate self-report of time elapsed since the last MDE.

The hypothesis that the non-relapsed previously depressed participants would demonstrate biased attention for emotional images (Hypothesis 1) was not supported by the data. Although Hypothesis 1 also specified that biases would be stronger for relapsed participants, non-relapsed participants were expected to show biases relative to never depressed participants based on the prior reports of biases in previously depressed samples (Joormann and Gotlib 2007; Newman and Sears 2015; Sears et al. 2011; Soltani et al. 2015). The results showed that non-relapsed participants did not differ from never depressed participants in their attention to any of the face image or naturalistic image types, at both the initial and follow-up visits. These findings provide additional evidence that attentional biases are less consistently observed in remitted versus current depression. Newman and Sears (2015) also found inconsistent attentional biases for previously depressed individuals in contrast to the pattern of attentional biases usually observed in currently depressed samples. They reported that previously depressed individuals had increased attention to depression-related naturalistic images, but not decreased attention to positive naturalistic images, compared to never depressed individuals. Overall, the evidence may indicate that attentional biases persist following remission from depression in an attenuated or less stable form, which would account for the inconsistent findings across studies. The persistence of biases in an attenuated form may represent a stable vulnerability factor for the recurrence of depression.

The current results highlight another potential source of variation in the results of previous studies of attentional biases in remitted depression; namely, sample heterogeneity. Prior research has typically treated previously depressed samples as homogeneous and have not explored factors that could distinguish subsets of previously depressed individuals. Our results suggest that vulnerability to future depressive episodes is one factor that is associated with attentional biases, as relapsed participants exhibited stronger attentional biases for emotional images than non-relapsed participants, even prior to relapse. Of course, vulnerability to future depressive episodes cannot be known or used to distinguish subsets of previously depressed individuals in a cross-sectional design. However, there may be other important differences in sample characteristics in studies examining previously depressed individuals that may be related to or independent from vulnerability to relapse, including number of previous depressive episodes, time since last episode, duration of depression, severity of residual depressive symptoms, and treatment history and status. These and other characteristics may influence whether attentional biases are observed in a particular sample. Future research on attentional biases should consider potential sources of heterogeneity in previously depressed samples and report relevant sample characteristics, to facilitate resolution of discrepant findings and to identify subsets of previously depressed individuals who do and do not exhibit attentional biases.

Hypothesis 2: Changes in Total Dwell Time from the Initial to Follow-Up Visit

In partial support of Hypothesis 2, within-group differences in attentional biases between the initial and follow-up visits were observed for the relapsed participants. Relapsed participants attended to happy faces and positive naturalistic images less and to threat naturalistic images more at the follow-up visit compared to the initial study visit. These results indicate that attention to emotional information becomes more biased as a person moves from a remitted depressed state to a relapsed state, although it is unclear why biases for sad faces and naturalistic images did not significantly increase as well. Prior cross-sectional designs comparing previously and currently depressed samples have been limited by the third variable problem and a lack of temporal ordering of the remitted and currently depressed states. Given the noted heterogeneity of previously depressed samples, sample characteristics may influence whether differences in attentional biases are observed between previously and currently depressed individuals. Within-group comparisons of the same individuals in a remitted versus relapsed state control for individual differences in potentially relevant characteristics. Our results therefore provide clear evidence for decreased attention to positive information and, to some extent, increased attention to negative information following a relapse of depression, although additional research is needed to determine why biases increased more for threat naturalistic images than the sad images.

Consistent with Hypothesis 2, changes in attentional biases from the initial to follow-up visit were not observed for the never depressed participants, but contrary to the hypothesis, changes were observed for the non-relapsed participants. The changes in attentional biases from the initial to follow-up visit for the non-relapsed participants differed from those of the relapsed participants. Non-relapsed participants attended more to threat naturalistic images and less to sad naturalistic images at the follow-up visit compared to the initial visit. These changes appear to be unrelated to changes in depression symptoms, as BDI-II scores for the non-relapsed participants were similar at the initial and follow-up visits. Measurement error or unreliability of the eye tracking tasks also seems to be an unlikely explanation for the differences between attentional biases at the initial and follow-up visits given recent evidence for the high reliability and stability of attentional bias indices from eye tracking tasks (e.g., Lazarov et al. 2018; Sanchez et al. 2017; Sears et al. 2018; Waechter et al. 2014). Future research should examine additional predictors of change in attentional biases in previously depressed samples, beyond changes in depression status or symptoms.

Hypothesis 3: Attention at the Initial Visit as a Predictor of Relapse and Change in Depression Symptoms

We found preliminary support for Hypothesis 3, that attentional biases predict relapse to depression. Reduced attention to positive naturalistic images at the initial visit was associated with a greater likelihood of relapse during the follow-up period among previously depressed participants. This relationship remained after controlling for initial depressive symptoms, thus ruling out the possibility that residual depressive symptoms accounted for both the decreased attention to positive images and increased vulnerability to relapse. Reduced attentional engagement with positive information may therefore be a unique risk factor for depression recurrence. Although promising, this finding is limited by the small number of relapses that occurred during the follow-up period. That is, only 12% of previously depressed participants experienced a relapse during the 6-month follow-up, which limited the amount of variance in outcome that could be explained by the predictor variables. The low relapse rate compared to previous studies (e.g., Backs-Dermott et al. 2010) may be due to a shorter follow-up period (6-months versus 1-year or more in some previous studies) and differences in the recruitment of previously depressed participants. We recruited previously depressed participants who had ever experienced a prior MDE, as opposed to following currently depressed participants to the point of remission (e.g., Backs-Dermott et al. 2010) or restricting recruitment to individuals who had experienced a recent MDE. Due to the small number of relapses observed in this study, we conducted supplementary regression analyses evaluating attentional biases at the initial visit as predictors of change in depression symptoms during the course of the study. Despite the increased statistical power of these analyses, attentional biases to the different image types did not significantly predict changes in depression symptoms among the previously depressed sample. Based on the results of Beevers and Carver (2003), it is possible that attentional biases are stronger predictors of future symptoms when tested in a negative mood state and/or in combination with stressful life events. Additional research is needed to explore potential moderators of the predictive effect of attentional biases on relapse and depression symptom severity.

Strengths, Limitations, and Directions for Future Research

This study benefitted from a number of conceptual and methodological strengths. The study rationale was based on a well-validated model of depression. The initial sample size was large, which permitted the possibility of meaningful longitudinal data. The sample was rigorously collected and followed over time, and was assessed using standardized measures. The measure of attentional bias was direct and employed state of the art assessment of attention to a variety of validated stimuli. Finally, both face images and naturalistic images were used as stimuli in the eye tracking task to determine whether the results would differ by image type. We observed some differences in the results across the two image types, which provides preliminary evidence that naturalistic images may be a more potent stimulus for eliciting attentional biases in depression-vulnerable individuals compared to face images, although additional research is required to corroborate this conclusion.

Despite the above strengths, this study had limitations that should be considered when evaluating the findings. Given the small number of relapses that occurred during the 6-month follow-up period, future studies that aim to predict relapses should use a follow-up period of 1–2 years to increase the likelihood of capturing depression relapse (simultaneously creating a larger and more representative relapsed group). Another beneficial modification would be to recruit samples with less variability in time since the previous depressive episode, perhaps limiting recruitment to individuals who have experienced an episode within the prior 2-year period. This strategy could eliminate individuals who are stably recovered and less likely to experience relapse, and thus increase the power of the study to identify predictors of relapse. As noted previously, this modification may also permit assessment and examination of time elapsed since the past MDE, as participants may be able to report this data more accurately if they have experienced a relatively recent MDE.

Another limitation was that some participants who relapsed were not identified at the time of relapse but reported symptoms consistent with a relapse retrospectively at the 6-month follow-up ($n=15$). As participants were monitored by phone at bi-weekly intervals for symptoms of depression, it is unclear if their depressive symptoms were under-reported during the telephone follow-ups, their retrospective recall of depressive symptoms at the 6-month follow-up visit was incorrect, or the symptoms were inaccurately evaluated during the bi-weekly phone follow-ups. These individuals were therefore excluded from all analyses due to the potential unreliability of the depression relapse diagnosis. Future studies should take measures to avoid these issues given that the present study suggests that they are not uncommon. One possibility would be to use both self-report and researcher-administered measures of depressive symptoms to improve the sensitivity of relapse assessment.

Although the longitudinal study design allowed a test of the hypothesis that attentional biases predict future recurrences of depression, no definitive conclusions regarding causality can be made. Studies that manipulate attentional biases, such as via attentional bias modification (ABM) procedures, are necessary to test the causal role of attentional biases in relapse to depression. It also remains unclear whether attentional biases are a pre-existing vulnerability factor for depression onset or develop as a consequence of having experienced depression (i.e., the "scar hypothesis"; Lewinsohn et al. 1981). This distinction is important for understanding the etiology of attentional biases and depression. Studies that examine attentional biases in individuals before they experience an initial depressive episode, as well as studies that compare individuals who have experienced a single episode of depression versus multiple episodes, would clarify whether attentional biases are a pre-existing vulnerability factor or develop at some point during the course of depression. Regardless of whether attentional biases are pre-existing or a "scar" of depression, however, the present results suggest that previously depressed individuals who exhibit reduced attention to positive information are more likely to experience a relapse in the short-term.

Finally, it should be noted that the generalizability of the present results is limited by the use of a sample consisting of only women, and that due to differences in the proportions of students versus community members in the participant groups, the previously depressed (relapsed and non-relapsed) participants tended to be older, more likely to be Caucasian, and less likely to be never married than the never depressed participants. Ideally, the findings of this study should be corroborated in mixed-sex samples that match previously depressed and never depressed participants on relevant demographic characteristics.

Conclusions

The present study provided evidence for concurrent and prospective associations between attentional biases for emotional images and relapse to depression. Previously depressed women who went on to experience a relapse to depression showed the hypothesized pattern of decreased attention to positive images and increased attention to negative images, relative to previously depressed women who did not experience a relapse and never depressed women, at the initial visit for naturalistic images and at the follow-up visit for both face and naturalistic images. The finding that attentional biases for face images were only exhibited by the previously depressed women once they had relapsed (i.e., when they were in a currently depressed state) suggests that attentional biases are at least in part influenced by state factors such as the presence of negative mood and/or elevated depression symptoms. In contrast to hypothesis and previous research, previously depressed women who did not experience a relapse did not demonstrate significant attentional biases relative to never depressed women in this study. Taken together, these results suggest that attentional biases persist in remitted depression but may be weaker and/or less consistent compared to those characteristic of current depression. This observation highlights the importance of considering potential sample heterogeneity in research on previously depressed individuals. Another important finding was that reduced attention to positive images prospectively predicted relapse to depression over 6 months. Thus, reduced attention to positive information may reflect a cognitive vulnerability for depression recurrence, although this result requires replication in a sample with a greater number of relapses. Future research that manipulates attentional biases and/or explores mechanisms by which attentional biases lead to relapse will further advance our understanding of the role of biased attention in depression recurrence.

Acknowledgements We thank two anonymous reviewers for their excellent feedback and suggestions. This research was supported by the Canadian Institutes of Health Research (CIHR) and the Natural Sciences and Engineering Council of Canada (NSERC).

Compliance with Ethical Standards

Conflict of Interest Kristin Newman, Leanne Quigley, Amanda Fernandez, Keith Dobson, and Christopher Sears declare that they have no conflict of interest.

Ethical Approval This research was approved by the authors' institutional research ethics board (the Conjoint Faculties Research Ethics Board at the University of Calgary).

Informed Consent Informed consent was obtained from all the individuals participating in the study.

Animal Rights No animal studies were carried out by the authors for this article.

References

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th edn., text rev.) Washington, DC: Author.

Armstrong, T., & Olatunji, B. (2012). Eye tracking of attention in the affective disorders: A meta-analytic review. *Clinical Psychology Review, 32*, 704–723. https://doi.org/10.1016/j.cpr.2012.09.004.

Arndt, J. E., Newman, K. R., & Sears, C. R. (2014). An eye tracking study of the time course of attention to positive and negative images in dysphoric and non-dysphoric individuals. *Journal of Experimental Psychopathology, 5*, 399–413. https://doi.org/10.5127/jep.035813.

Babyak, M. (2004). What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic Medicine, 66*, 411–421.

Backs-Dermott, B. J., Dobson, K. S., & Jones, S. L. (2010). An evaluation of an integrated model of relapse in depression. *Journal of Affective Disorders, 124*, 60–67. https://doi.org/10.1016/j.jad.2009.11.015.

Beck, A., Steer, R., & Brown, G. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio: Psychological Corporation.

Beevers, C., & Carver, C. (2003). Attentional bias and mood persistence as prospective predictors of dysphoria. *Cognitive Therapy and Research, 27*, 619–637. https://doi.org/10.1023/A:1026347610928.

Burcusa, S., & Iacono, W. (2007). Risk for recurrence in depression. *Clinical Psychology Review, 27*, 959–985. https://doi.org/10.1016/j.cpr.2007.02.005.

Cannon, D., Tiffany, S., Coon, H., Scholand, M., McMahon, W., & Leppert, M. (2007). The PHQ-9 as a brief assessment of lifetime major depression. *Psychological Assessment, 19*, 247–251. https://doi.org/10.1037/1040-3590.19.2.24.

DeRaedt, R., & Koster, E. (2010). Understanding vulnerability for depression from a cognitive neuroscience perspective: A reappraisal of attentional factors and a new conceptual framework. *Cognitive, Affective, and Behavioral Neurosciences, 10*, 50–70. https://doi.org/10.3758/CABN.10.1.50.

Donges, U., Kersting, A., & Suslow, T. (2012). Women's greater ability to perceive happy facial emotion automatically: Gender differences in affective priming. *PLoS ONE, 7*, e41745. https://doi.org/10.1371/journal.pone.0041745.

Elgersma, H. J., Koster, E. H., Vugteveen, J., Hoekzema, A., Penninx, B. W., Bockting, C. L., & de Jong, P. J. (2019). Predictive value of attentional bias for the recurrence of depression: A 4-year prospective study in remitted depressed individuals. *Behaviour Research and Therapy, 114*, 25–34. https://doi.org/10.1016/j.brat.2019.01.001.

Elgersma, H. J., Koster, E. H. W., van Tuijl, L. A., Hoekzema, A., Penninx, B. W. J. H., Bockting, C. L. H., et al. (2018). Attentional bias for negative, positive, and threat words in current and remitted depression. *PLoS ONE, 13*(10), e0205154. https://doi.org/10.1371/journal.pone.0205154.

First, M., Spitzer, R., Gibbon, M., & Williams, J. (1995). *Structured clinical interview for DSM-IV axis I disorders*. New York: Biometrics Research Department.

Gotlib, I., & Joormann, J. (2010). Cognition and depression: Current status and future directions. *Annual Review of Clinical Psychology, 6*, 285–312. https://doi.org/10.1146/annurev/clinpsy.121208.131305.

Hook, J., Hodges, E., Whitney, K., & Segal, D. (2007). Structured and semi-structured interviews. In M. Hersen & J. Thomas (Eds.), *Handbook of clinical interviewing with adults* (pp. 24–37). Los Angeles: Sage.

Isaac, L., Vrijsen, J., Rinck, M., Speckens, A., & Becker, E. (2014). Shorter gaze duration for happy faces in current but not remitted depression: Evidence from eye movements. *Psychiatry Research, 218*, 79–86. https://doi.org/10.1016/j.psychres.2014.04.002.

Joormann, J., & Gotlib, I. (2007). Selective attention to emotional faces following recovery from depression. *Journal of Abnormal Psychology, 116*, 80–85. https://doi.org/10.1037/0021-843X.116.1.80.

Kellough, J. L., Beevers, C. G., Ellis, A. J., & Wells, T. T. (2008). Time course of selective attention in clinically depressed young adults: An eye-tracking study. *Behaviour Research and Therapy, 46*, 1238–1243. https://doi.org/10.1016/j.brat.2008.07.004.

Kemp, A., Silberstein, R., Armstrong, S., & Nathan, P. (2004). Gender differences in the cortical electrophysiological processing of visual emotional stimuli. *Neuroimage, 21*, 632–646. https://doi.org/10.1016/j.neuroimage.2003.09.055.

Kessler, R., McGonagle, K., Swartz, M., Blazer, D., & Nelson, C. (1993). Sex and depression in the National comorbidity survey I: Lifetime prevalence, chronicity, and recurrence. *Journal of Affective Disorders, 29*, 85–96. https://doi.org/10.1016/0165-0327(93)90026-G.

Lazarov, A., Ben-Zion, Z., Shamai, D., Pine, D. S., & Bar-Haim, Y. (2018). Free viewing of sad and happy faces in depression: A potential target for attention bias modification. *Journal of Affective Disorders, 238*, 94–100. https://doi.org/10.1016/j.jad.2018.05.047.

Lee, E., Kang, J., Park, I., Kim, J., & An, S. (2008). Is a neutral face really evaluated as being emotionally neutral? *Psychiatry Research, 157*, 77–85. https://doi.org/10.1016/j.psychres.2007.02.005.

Leppänen, J., Milders, M., Bell, J., Terriere, E., & Hietanen, J. (2004). Depression biases the recognition of emotionally neutral faces. *Psychiatry Research, 128*, 123–133. https://doi.org/10.1016/j.psychres.2004.05.020.

Lewinsohn, P. M., Steinmetz, J. L., Larson, D. W., & Franklin, J. (1981). Depression-related cognitions: Antecedent or consequence? *Journal of Abnormal Psychology, 90*, 213–219. https://doi.org/10.1037//0021-843X.90.3.213.

Lobbestael, J., Leurgans, M., & Arntz, A. (2011). Inter-rater reliability of the structured clinical interview for DSM-IV axis I disorders (SCID I) and axis II disorders (SCID II). *Clinical Psychology and Psychotherapy, 18*, 75–79. https://doi.org/10.1002/cpp.693.

Montagne, B., Kessell, R., Frigerio, E., deHaan, E., & Perrett, D. (2005). Sex differences in the perception of affective facial expressions: Do men really lack emotional sensitivity? *Cognitive Processing, 6*, 136–141. https://doi.org/10.1007/s10339-005-0050-6.

Newman, K. R., & Sears, C. R. (2015). Eye gaze tracking reveals different effects of a sad mood induction on the attention of previously depressed and never depressed women. *Cognitive*

Therapy and Research, 39, 292–306. https://doi.org/10.1007/s10608-014-9669-x.

Peckham, A., McHugh, K., & Otto, M. (2010). A meta-analysis of the magnitude of biased attention in depression. *Depression and Anxiety, 27*, 1135–1142. https://doi.org/10.1002/da.20755.

Piccinelli, M., & Wilkinson, G. (2000). Gender differences in depression. *British Journal of Psychiatry, 177*, 486–492. https://doi.org/10.1192/bjp.177.6.486.

Sanchez, A., Romero, N., & De Raedt, R. (2017). Depression-related difficulties disengaging from negative faces are associated with sustained attention to negative feedback during social evaluation and predict stress recovery. *PLoS ONE, 12*, e0175040. https://doi.org/10.1371/journal.pone.0175040.

Sears, C., Quigley, L., Fernandez, A., Newman, K., & Dobson, K. (2018). The reliability of attentional biases for emotional images measured using a free-viewing eye-tracking paradigm. *Behavior Research Methods*. https://doi.org/10.3758/s13428-018-1147-z.

Sears, C. R., Newman, K. R., Ference, J. D., & Thomas, C. L. (2011). Attention to emotional images in previously depressed individuals: An eye-tracking study. *Cognitive Therapy and Research, 35*, 517–528. https://doi.org/10.1007/s10608-011-9396-5.

Solomon, D. A., Keller, M. B., Leon, A. C., Mueller, T. I., Lavori, P. W., Shea, M. T., … Endicott, J. (2000). Multiple recurrences of major depressive disorder. *American Journal of Psychiatry, 157*, 229–233. https://doi.org/10.1176/appi.ajp.157.2.229.

Soltani, S., Newman, K., Quigley, L., Fernandez, A., Dobson, K., & Sears, C. (2015). Temporal changes in attention to sad and happy faces distinguish currently and remitted depressed individuals from never depressed individuals. *Psychiatry Research, 230*, 454–463. https://doi.org/10.1016/j.psychres.2015.09.036.

Spitzer, R., Kroenke, K., Williams, J., & The Patient Health Questionnaire Primary Care Study Group (1999). Validation and utility of a self-report version of the PRIME-MD: The PHQ primary care study. *Journal of the American Medical Association, 282*, 1737–1744. https://doi.org/10.1001/jama.282.18.1737.

Tottenham, N., Tanaka, J., Leon, A., McCarry, T., Nurse, M., Hare, T., Marcus, D., Westerlund, A., Casey, B., & Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research, 168*, 242–249. https://doi.org/10.1016/j.psychres.2008.05.006.

Waechter, S., Nelson, A., Wright, C., Hyatt, A., & Oakman, J. (2014). Measuring attentional bias to threat: Reliability of dot probe and eye movement indices. *Cognitive Therapy and Research, 38*, 313–333. https://doi.org/10.1007/s10608-013-9588-2.

Woody, M., Owens, M., Burkhouse, K., & Gibb, B. (2016). Selective attention toward angry faces and risk for major depressive disorder in women: Converging evidence from retrospective and prospective analyses. *Clinical Psychological Science, 4*, 206–215. https://doi.org/10.1177/2167702615581580.

Yiend, J. (2010). The effects of emotion on attention: A review of attentional processing of emotional information. *Cognition and Emotion, 24*, 3–47. https://doi.org/10.1080/02699930903205698.