MAJOR RESEARCH PROJECT

LITERATURE REVIEW: A Systematic Review of Expressive Writing for Alleviating Depressive Symptoms

EMPIRICAL PAPER: Colouring-in, a Distraction Technique? A study Looking at The Effects of Colouring-in for Adults in Reducing Negative Affect and State Rumination

Submitted by **Neil Drew**, to the University of Exeter

as a thesis for the degree of **Doctor of Clinical Psychology**, April 2018

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature: ……Neil Drew……

Table of Contents

List of Tables

Literature Review

Empirical Paper

List of Figures

Literature Review

Empirical Paper

SCHOOL OF PSYCHOLOGY

DOCTORATE IN CLINICAL PSYCHOLOGY

LITERATURE REVIEW

A Systematic Review of Expressive Writing for Alleviating Depressive Symptoms

Trainee Name:	**Neil Drew**
Primary Research Supervisor:	**Dr Nicholas Moberly** Senior Lecturer, School of Psychology, University of Exeter
Secondary Research Supervisor:	**Professor Ed Watkins** Professor of Experimental and Applied Clinical Psychology, Director of Research for Professional Doctorates, University of Exeter
Target Journal:	Clinical Psychology Review
Word Count:	6427 words (excluding abstract, tables, contents, list of figures and tables, references, appendices)

Abstract

Objective: The benefit of writing for alleviating depressive symptoms has been consistently shown in prior research. Psychological models of therapy often encourage writing in an expressive, emotionally reflective manner, although there is a paucity of research around expressive writing being effective for adults with clinical depression. There has been no previous systematic review that aimed to summarise the research investigating expressive writing in samples consisting of adults with clinical depression.

Method: The PRISMA guidelines were followed for inclusion and exclusion purposes. A search across the following databases was conducted: Medline EBSCOhost, PubMed, Web of Science, OVID (Journals@OVID, PsycINFO and PsycARTICLES), and Creativity Research Journal. A progressive screening procedure yielded 11 relevant studies for full review.

Results: The findings in this review suggest that expressive writing can alleviate depressive symptoms in participants who experience high levels of these symptoms. This was largely consistent across all studies with variations in number of conditions, control groups (e.g., active, inactive), format of expressive writing (e.g., typing, writing) and outcome measures (e.g., BDI, PHQ, CES-D). Expressive writing particularly outperformed neutral, non-emotive writing in alleviating depressive symptoms.

Conclusions: Despite considerable variability and limitations in the studies appraised in this review, overall findings indicate support for expressive writing being effective at alleviating depressive symptoms when compared to a control group. This provides evidence that emotional processing helps reduce depressive symptoms. To

enhance strength of clinical implications for expressive writing, future research should focus on samples consisting of adults with clinical depression.

Keywords: Expressive writing, depression, control group

Introduction

Expressive Writing is a formal therapeutic technique developed by Pennebaker and Beall (1986) through experimental research in which individuals are instructed to write about their emotional experiences, thoughts and feelings around these and traumatic distressing memories. Traditionally, in experimental research on expressive writing, individuals write in this manner for 20 minutes during three writing sessions held on consecutive days whilst a control group also write for this amount of time with neutral and non-emotive content such as writing about the layout of the person's bedroom (Sloan, Marx, & Greenberg, 2011). Emerging research on the benefits of expressive writing has consistently been demonstrated over the last 20 years for both physical and psychological health. Expressive writing has been shown to help psychosis symptoms by improving the immune system via increasing T – Helper cells (Pennebaker, Kiecolt-Glaser, & Glaser, 1988; Petrie, Booth, & Pennebaker 1998), decreasing physical pain problems (Smyth, 1998; Campbell & Pennebaker, 2003), reducing intrusive negative thoughts and indications of depression (Gortner, Rude, & Pennebaker, 2006; Lepore, 1997), post-traumatic stress symptoms (Sloan, Marx, Epstein, & Lexington, 2007), as well as increasing working memory capacity (Klein & Boals, 2001).

Despite substantial empirical evidence supporting expressive writing as a therapeutic intervention which enhances psychological well-being, the mechanisms explaining this effect are unknown (Graf, Gaudiano & Geller, 2008). Pennebaker and colleagues (Kacewicz, Slatcher & Pennebaker, 2007; Pennebaker, 1997; Pennebaker & Seagal, 1999; Slatcher & Pennebaker, 2007;) suggest two main theories regarding how expressive writing works. The Inhibition Theory (Pennebaker,

1989) originally proposed that withholding thoughts and feelings resulted in physiological and psychological stress. Theoretically, it argues that emotional disclosure decreases stress and its physical symptoms which then alleviates psychological problems by breaking the undesired inhibition process (Ahmadi, Abdoll, Ramezani, & Heshmati, 2010) The second theory is called Cognitive Changes Theory (Smyth, True, & Souto, 2001) which suggests that expressing emotions using words and talking about difficult experiences allows for cognitive re-structuring. This relates to an exposure-based model, meaning that repeatedly writing about the same stressful event has the capacity to lead to reducing negative affect surrounding the event (Hayes, Laurenceau, Feldman, Strauss, & Cardaciotto, 2007).

Expressive Writing and Depression

Prevalence figures suggest that 3.3 in 100 people suffer from depression in England (McManus, Bebbington, Jenkins, & Brugha, 2016). Limited research has been conducted on depression prevalence in black, Asian and ethnic minority communities in the UK (Mental Health Foundation, 2016). These communities are considered to be at an increased risk of depression (Bhui & McKenzie, 2008). Furthermore, socioeconomically disadvantaged young people are two to three times more likely to develop mental health problems than their peers (Mental Health Foundation, 2016). This suggests that UK prevalence figures do not portray the entire story with regards to who is receiving the diagnosis. The figures are also likely an underestimate due to ethnic minority communities being missed, which consequently could impact their chances of receiving treatment. Reports from England suggest only one in eight adults with mental health problems are currently receiving treatment (McManus et al., 2016; Welsh Health Survey, 2015). The

Increasing Access to Psychological Therapies (IAPT) programme was developed in the UK to support the NHS in delivering evidence-based psychological therapies for depression and anxiety. Its need in the current climate is evident with 259,016 people being referred for psychological therapies and 154,722 people entering treatment between 2012 to 2013 (IAPT Key Performance Indicators, 2013). Thus, increasing awareness of the benefits of a well-researched simple self-help method such as expressive writing may be beneficial for those who are not accessing treatment, as well as for those with a MDD diagnosis and are accessing therapy.

Despite the benefits of expressive writing being evidenced, few studies have specifically investigated the impact of expressive writing on depression (Gortner et al., 2006). An emerging evidence base is beginning to show how expressive writing can be helpful in alleviating self-reported depressive symptoms. A meta-analysis of the impact of emotional disclosure found a small but reliable effect of expressive writing on depressive symptoms compared to non-emotive writing (Frattaroli, 2006). More specifically, expressive writing has been found to be useful at reducing depressive symptoms in students preparing for exams (Lepore, 1997). A significant decline was observed in depressive symptoms measured by the SCL-90-R in the college students assigned to the expressive writing condition compared to participants in the control group who wrote about a trivial topic. Similar findings have been demonstrated for adult females with a diagnosis of depression who have suffered from domestic abuse (Koopman et al., 2005), and women with PTSD such that their depressive symptoms reduced following an expressive writing intervention compared to a non-emotive writing control group (Sloan & Marx, 2004). As opposed to expressive writing being merely a reactive strategy for depression, research has

also looked at its ability to function as a preventative intervention for those who are at risk for depression (Sloan et al., 2007). Gortner et al.'s (2006) study selected participants with a self-reported history of depression but who were currently euthymic on the Beck Depression Inventory (BDI). Results showed that expressive writing was associated with lower depressive symptoms compared to a control group in a 6-month follow up among participants high in trait emotional suppression. It was hypothesised that expressive writing allowed participants to express their emotions in ways they would not normally do, which potentially led to the prevention of a depression relapse. In addition, expressive writing has been shown to increase a person's perceived level of social support (Pennebaker, 1997), which could be a key factor in alleviation of depressive symptoms.

Although positive, the above findings have limitations. Most of the studies use student populations and do not test the effects of expressive writing on severely depressed persons. Studies often measure expressive writing in a lab setting over one session without follow-up, thus the evidence cannot challenge the possibility that the benefits of expressive writing are only momentary. Also, little is known about the moderators (e.g., depression severity, sex) and mediators (e.g., emotional content, number of expressive writing sessions, length of writing) of the effect of expressive writing. This is because the majority of studies only look to see whether expressive writing reduces symptoms when compared to a control. However, this does not show how and why expressive writing has positive effects.

Expressive Writing and Psychotherapy

Within different models of psychotherapy, it is often encouraged that clients are provided with adaptive, accessible self-care techniques that they can use in-

between sessions (Kazantzis, Lampropoulos, & Deane, 2005). It is suggested that expressive writing can be used as a supplement to existing treatment for depression (Krpan et al., 2013), with it being an evidence-based effective approach that is time and cost efficient (Kazdin & Blase, 2011). For instance, Graf et al. (2008) found that psychotherapy outpatients who completed two written self-disclosure homework exercises displayed significantly greater reductions in depressive symptoms compared to controls (treatment as usual). Taking this one step further, Hayes et al. (2007) developed a novel psychotherapy program for depressed patients called exposure-based cognitive therapy (EBCT), which included components such as stress management, exposure to depressive content, and personal growth tasks. Each patient who underwent this therapy was also asked to engage in expressive writing for 20 minutes between each psychotherapy session, to encourage emotional processing throughout the week. A preliminary investigation of EBCT in 29 patients showed that higher levels of emotion processing content in individuals' expressive writing predicted more improvement in depression and hopefulness, whereas higher levels of avoidance in the narratives predicted less improvement in symptoms and more hopelessness (Hayes, Beevers, Feldman, Laurenceau, & Perlman, 2005). Sloan and Marx (2004) argue that expressive writing may act as a form of exposure therapy for the individual, which is a known method used within psychological therapies (e.g., exposure to distress to process and relieve associated symptoms; Hofmann et al., 2012). In relation to this, some existing treatment approaches for depression such as Narrative Exposure Therapy (NET; Schauer, Neuner, & Elbert, 2011) and the reformulation of problems via Cognitive Analytic Therapy (CAT; Ryle & Kerr, 2002) incorporate written trauma narratives in practitioner manuals. So, it appears that expressive writing is already used in some models of therapy. However,

it is not explicitly labelled as expressive writing which could be a reason as to why it is not commonly recognised as a potential primary self-help intervention strategy for depression.

The gap in the literature around expressive writing being an effective intervention for a specific population group (e.g., people with severe depression) is substantial. The meta-analysis conducted by Frattaroli (2006) indicated that there is a paucity of studies investigating the impact of expressive writing on participants with severe depression. Until expressive writing is shown to be consistently effective for people with severe depression or self-reported depressive symptoms, it is unlikely to be recommended as a self-help intervention or an addition to psychotherapy. Thus, as 12 years have passed since the last review on expressive writing there is a need to update the review. This review contributes to the literature as it specifically focuses on studies with samples of participants with a diagnosis of depression or self-reported depressive symptoms as determined by scoring within the mild to severe range on standardized psychometric measures.

Aim of Systematic Review

This systematic literature review aims to explore whether expressive writing is an effective intervention for alleviating depressive symptoms in clinical depression.

Literature Review Question. Is writing about the self in an emotive and reflective manner an effective therapeutic intervention for reducing depressed mood in persons with high levels of depressive symptoms?

Method

Eligibility Criteria

The eligibility criteria are presented following recommendations by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Moher, Liberati, Tetzlaff, & Altman, 2009) five-component structured known as "PICOS" (Table 1) where each letter refers to a component: population (P), the interventions (I), the comparator group (C), the outcome (O), and the study design (S). The numbers will be presented in a flowchart (see Figure 1).

Population. The studies that were selected consisted of adult populations, aged 18-65. Studies recruiting clinical and non-clinical samples were included. Studies were only included if samples had a diagnosis of depression or scored above the cut-off criteria for clinical depression on standardized self-report measures, as determined by the developers of the scale. Samples were not omitted if they had co-morbid mental health conditions (e.g., generalized anxiety disorder), although samples where depression was not the primary diagnosis were excluded, which included psychosis, personality disorders, brain injury and bipolar disorder.

Table 1.
Study Selection Process via PICOS

PICOS components	Inclusion criteria	Exclusion criteria
Population	Aged 18-65 years, meeting criteria for depression (as determined via structured interview, mental health diagnosis or above cut-off on standardized self-report measures which meet the criteria for mild, moderate, severe depression and dysthymia)	Participants with brain injury, personality disorder, psychosis, bi-polar disorder
Intervention	Writing (including typing, blogging, etc.) about the self in an autobiographical expressive form (e.g., about one's own life and their thoughts and feelings around it). Content can include retrospective, present or future-oriented expressive writing. Writing could be one-off or a repeated task.	Non-self-focussed writing
Comparator group	Active control group (e.g., writing about other people, distraction, etc.) or do-nothing control group. Alternative intervention (e.g., drawing, expression through speech)	Studies that do not use a comparison group (e.g., repeated measures, single case designs)
Outcome	Standardized validated measures of depressed mood/symptoms, diagnosis of depression	Studies using general well-being measures or measures of positive and/or negative mood without an additional measure of depression
Study design	Randomized controlled experiments	Qualitative studies, retrospective and prospective studies not involving experimental manipulation, review papers and meta-analyses, case studies, non-peer-reviewed studies, articles not written in English

Interventions (Exposure). The only studies selected were those that used expressive writing as an intervention and compared it to a control group (e.g., non-emotive writing) or an alternative intervention group. To determine whether a study used expressive writing as an intervention technique, the key was to assess whether

the content of text was of an expressive and self-reflective nature. For example, participants should write a narrative about the self and reflect on thoughts, feelings, memories and personal experiences. This could be performed through handwriting or typing. The writing had to be focussed on the person's feelings or thoughts in the experience, rather than merely recording daily activities or writing lists with no emotional content. The expressive writing completed by the participant could be a one-off or repeated intervention.

Comparator. Studies were included if they used either a control group (e.g., non-emotive writing, distraction, rumination, wait-list) or an alternative intervention group (e.g., emotional expression through speech, cognitive therapy). Studies that had no comparator group were excluded.

Outcomes. The main outcome of interest is a validated measure of depressive symptoms including the Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983), the Patient Health Questionnaire-9 (PHQ-9; Kroenke, Spitzer, & Williams, 2001), Beck Depression Inventory (BDI; Beck & Beamesderfer, 1974), Beck Depression Inventory-II (BDI-II; Beck, Steer & Brown, 1996). The Centre for Epidemiologic Studies Depression Scale (CES-D; Radloff, 1977), Depression Anxiety Stress Scales (DASS; Lovibond & Lovibond, 1995), and the Zung Self-Rating Depression Scale (SDS; Zung, 1986). To determine whether samples met criteria for depression, cut-off scores created by the authors of the standardized psychometrics were used. Psychometric measures which only looked generally at negative affect were not used due to them not specifically measuring depression.

Study Design. Only randomized controlled studies were included. Qualitative studies, review studies, meta-analyses, or studies that did not involve experimental manipulation were not included.

Information Sources

Following a preliminary search to generate relevant terms, the following electronic databases were searched from first available entries until 28th September 2017: Medline EBSCOhost, PubMed, Web of Science, OVID (Journals@OVID, PsycINFO and PsycARTICLES), and Creativity Research Journal. Searches were conducted in the Creativity Research Journal due to the creative nature of the intervention under review in this paper. Grey literature was not searched due to limitations of time. Use of key author as a search tool (e.g., Pennebaker) was not used due to the large number of published articles on expressive writing across a range of conditions. The word “expressive” was not used in the search terms due to the large number of yielded studies its inclusion generated, and thus could not be screened because of resource limitations.

Search Terms

Table 2 displays search terms and key words entered into the databases.

Table 2.
Search Terms Entered in Databases

Research question area	Search terms
Population	(depress* OR unipolar OR dysthymi*) AND
Intervention	(writ* OR journal* OR diar* OR blog* OR handwrit* or lifelogging OR narrat*) AND
Outcome	(depress* OR dysphor* OR sad* OR despond* OR affect* OR anhedoni* OR deject* OR melanchol* OR unipolar OR dysthymi*) AND
Design	(arm* OR random* OR allocate* OR condition* OR experiment OR RCT OR group* OR manipulat* OR interven* OR treat* OR trial OR control*)

Screening Procedure

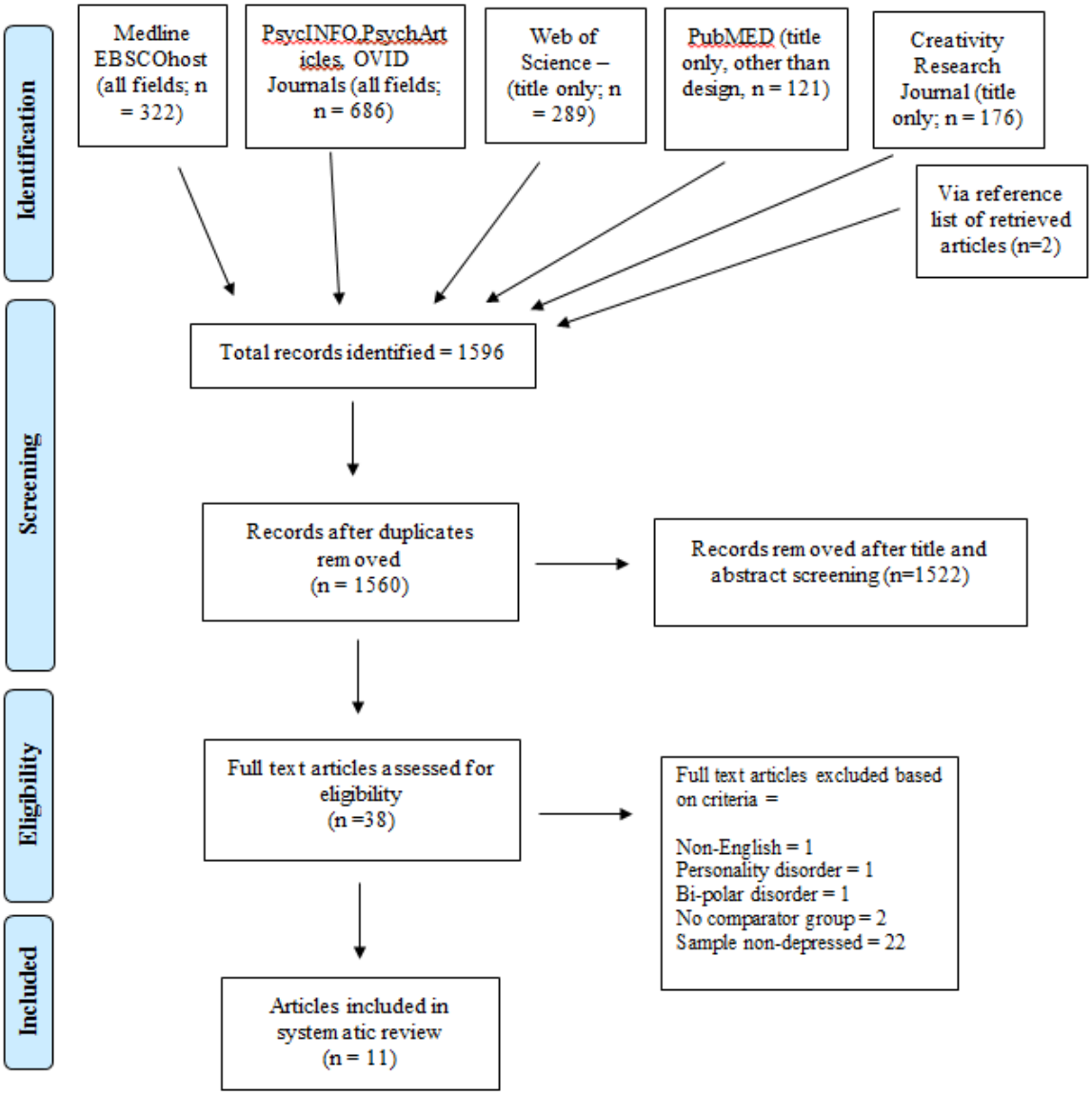
The titles and abstracts of the records generated by the search terms were initially screened for eligibility using PECOS criteria (Population, Exposure, Comparator, Outcome, Study Design; Higgins & Green, 2011). Figure 1 displays the screening flowchart. Following the database search, duplicate records were deleted and titles and abstracts of selected studies were screened for inclusion. Full texts of the remaining records were then screened, noting the reasons for any exclusions. To ensure articles were not missed, additional time was put into screening the reference lists of each full text article which was read, as recommended by the NICE (2012) guidelines for compiling systematic reviews

(Moher, 2015). Review papers were removed only after further review of reference lists. Some adjustments were made when searching on certain databases (Web of Science, PubMed, Creativity Research Journal) due to the large number of results (approx. 4,000) when searching in ‘all fields’ and time limitations necessary to complete the review. Therefore, on these databases, the search consisted of title only. An independent reviewer assessed 10 of the 38 full-text studies for eligibility, yielding 100% agreement with regards to meeting inclusion and exclusion criteria.

Data Extraction

Data were extracted from the online databases and results were entered into a pre-prepared Excel spreadsheet. Information extracted consisted of article title, authors, year, journal title, sample size, sample characteristics, abstract, outcome measures and results. For all eleven remaining studies, data were summarized using a proforma based upon population, exposure, comparator and outcomes (PECO; O’Connor, Higgins & Green 2011) following the categories: (a) study design, (b) sample characteristics, (c) exposure (interventions), (d) measures, (e) main findings, (f) evaluation and (g) quality rating.

Figure 1.
PICOS PRISMA flow chart



Risk of Bias in Studies

The Quality Assessment Tool for Quantitative Studies (QATQS; Appendix 1; Effective Public Health Practice Project, 2009; Thomas, Ciliska, Dobbins, & Micucci, 2004) was used to evaluate quality of articles as strong, moderate or

weak. This tool encompasses a standardised method of evaluating the quality of studies. Selected articles were evaluated on the six categories of the QATQS: selection bias, study design, confounders, blinding, data collection method, and withdrawals and dropouts. Studies were not excluded based on quality, but the quality assessment was used to inform the synthesis of evidence. For purposes of quality evaluation, inter-rater reliability judgements were gathered independently and blind for six (55%) of the final eleven studies yielding 100% agreement.

Results

Critical Summary

Eleven studies met the eligibility criteria for inclusion in the review, which are summarised in Table 3. All studies investigate the role of expressive writing in alleviating depressive symptoms by either comparing it to one or two conditions (e.g., active or inactive control group, alternative intervention). All studies were published between 2002 and 2016 indicating that this research is still relatively in its infancy. All participants scored above the cut-off for mild depression as determined by standardized self-report measures. Participants are drawn from non-clinical populations (seven studies) and clinical populations (four studies).

Table 3.

Summary of articles included for analysis, ordered alphabetically by author

Study & location	Design	Sample	Expressive Writing (EW) condition(s)	Comparator condition(s)	Measures	Main Findings	Evaluation	QATQS ratings*
1. Ahmadi et al. (2010) Iran	RCT	33 undergraduate students, 4 male and 29 female (*M* age = 21) with mild depression determined via self-report measures (BDI).	EW – written emotional disclosure about their depression. One session of EW in a lab setting.	Emotional expression through speech – this involved talking about depressive symptoms. One session of this, measures completed pre and post.	BDI	In EW condition, depression scores showed a significantly greater reduction from baseline (*M* = 19.7, *SD* = 5.7) to post-test (*M* = 11.5, *SD* = 5.5), Cohen *d* = 1.46, compared to the speech condition baseline (*M* = 18.3, *SD* = 3.5) to post-test scores (*M* = 18.0, *SD* = 3.8), Cohen *d* = .08.	**Strengths:** Standardised measures. **Limitations:** Small sample. Inadequate analysis and reporting of results (e.g., no effect sizes). Generalisability - Iranian population. No description of randomization process. Control condition not matched for writing.	A – Weak B – Strong C – Weak D – Moderate E – Weak F – Moderate **Overall: Weak**
2. Baum & Rude (2013) United States	RCT	218 students, age 18-30 (*M* = 21), 158 female, 60 male. Low-mild depression levels (score 16 or more on CES-D).	Two intervention conditions: i) Traditional EW (e.g. instructed to write about deepest thoughts and feelings about an emotional event) for 20 minutes ii) EW augmented by emotion-acceptance instructions influenced by mindfulness and	Non-emotional expressive writing group instructed to write objectively and without emotion about how they spend their time for 20 minutes.	CES-D	EWEA significantly more beneficial than control writing in reducing depressive symptoms and traditional EW more beneficial than control writing. EW performed significantly worse than control at alleviating depressive symptoms in participants with an initial CESD score greater or equal to 20.36 (e.g., moderate depression). EW performed significantly better	**Strengths:** Analysis looking at EW effectiveness on depression severity via simple slopes. Large sample. **Limitations:** Effect sizes, means and standard deviations not reported. No description of	A – Moderate B – Strong C – Strong D – Moderate E – Strong F – Moderate **Overall: Strong**

			compassion (e.g., participants encouraged to write in a "emotionally friendly" manner and "staying present") for 20 minutes.			then control for initial CESD scores of 16.97 or less (mild depression).	randomization process. Lack of control and standardization due to all interaction and data collection being online (e.g., environment, time taken).	
3. Dean, Potts & Barker (2016) England	RCT	204 adults (female = 157, male = 47) mild depression determined via self-report measures (score 16 or more on CES-D).	Online EW (write about upsetting experience for 5 minutes every two weeks).	Internet Support Groups hosted by Psych Central. Participants were advised to start their own discussion threads and join existing ones Measurement points were baseline, 3, and 6 months.	CES-D	3 (time) x 2 (condition) mixed ANOVA found statistically significant effect of time for depression, social support and anxiety, but none of the interaction effects were significant, suggesting there was no differences in treatment effects dependent on condition; EW or use of internet support groups. Expressive writing baseline scores (*M* = 30.2, *SD* = 12.2) compared to scores at 6 months (*M* = 21.5, *SD* = 12.7) showed a medium to large effect size for expressive writing (Cohen *d* = 0.7). A medium effect size (Cohen *d* = 0.5) was found for internet support groups baseline scores (*M* = 28.3, *SD* = 12.5) compared to 6 months scores (*M* = 21.8, *SD* = 13.3).	**Strengths:** Large sample. Qualitative component. **Limitations:** Majority of participants female. No traditional control group (e.g., non-emotive writing). Low engagement rate for the Internet support group intervention – high attrition. Engagement measured via retrospective self-reports, prone to bias.	A – Moderate B – Strong C – Moderate D – Moderate E – Strong F – Weak **Overall: Moderate**
4. Graf et al. (2008)	RCT	44 adult participants (14 male, 30 female) with a mean age of	Adapted Pennebaker EW task. Two weekly 20 minute writing sessions, to write about the most	Neutral writing about participants' plans for the week. Two	DASS	Participants in the EW condition showed statistically significant greater reductions in depressive symptoms from session 1 (*M* = 17.45, *SD* =	**Strengths:** Clinical sample. 100% retention rate.	A – Moderate B – Strong C – Weak D – Moderate E – Moderate

United States		33. All participants were from a university-based outpatient psychotherapy unit currently undergoing therapy with depression and anxiety.	stressful and upsetting experiences of the person's entire life.	weekly 20 minute sessions.		10.40) to session 3 (*M* = 11.32, *SD* = 10.25), Cohen *d* = 0.5, compared with the neutral writing control condition session 1 (*M* = 17.91, *SD* = 12.30) to session 3 scores (*M* = 17.45, *SD* = 12.18), Cohen *d* = 0.03.	**Limitations:** Lack of demographic details due to confidentiality, thus unable to compare some group differences. Small sample. Completion rate of EW not interpedently verified. No description of randomization process.	F – Strong **Overall: Moderate**
5. Koopman et al. (2005) United States	RCT	47 females (*M* age = 36) who had been victim to intimate partner violence (IPV) and experienced depression as determined by BDI (score greater than 14).	EW for 20 minutes around traumatic events. Participants advised to write about the most traumatic experience of their lives. Four writing sessions across four weeks.	Neutral, non-emotive, writing about daily routines for 20 minutes. Four writing sessions across four weeks.	BDI	Expressive writing baseline (*M* = 18.3, *SD* = 12.9) to follow-up (*M* = 12.5, *SD* = 12.2) scores (Cohen *d* = 0.4) showed no significant difference in reduction of depressive symptoms compared to baseline neutral writing condition scores (*M* = 17.1, *SD* = 7.3) to follow-up (*M* = 14.4, *SD* = 10.8) scores (Cohen *d* = 0.2). Moderation analysis showed that women who were more depressed at baseline demonstrated significantly greater reductions in depression when assigned to EW compared to neutral writing.	**Strengths:** High control (lab setting). Moderation analysis to further explore findings. **Limitations:** Small sample and female only. No description of randomization process. Effect sizes not reported.	A – Moderate B – Moderate C – Strong D – Moderate E – Moderate F – Strong **Overall: Strong**
6. Kovac & Range (2002)	RCT	121 undergraduates (88 female, 33 male) ranging in age from 18-42 (*M*	Two intervention conditions: i) EW with cognitive change instructions (analysing thoughts around a distressing	Writing about the layout of their bedroom. Wrote for 20 minutes on four	Zung SDS	No significant reduction in Zung SDS for the traditional EW condition from baseline (*M* = 41.80, *SD* = 7.15) to follow-up (*M* = 41.24, *SD* = 6.72; Cohen *d* = 0.08), EW with	**Strengths:** Novel task in EW with cognitive instructions.	A – Moderate B – Strong C – Moderate D – Moderate E – Moderate F – Strong

United states		= 23.12, *SD* = 5.44) who scored 6 or higher on SBQ and met criteria for depression on Zung SDS.	time in your life) ii) Traditional EW task about a difficult time in your life. Wrote for 20 minutes on four days over two weeks.	days over two weeks.		cognitive change baseline (*M* = 44.08, *SD* = 9.29) to follow-up (*M* = 41.16, *SD* = 10.95; Cohen *d* = 0.2), or the control condition baseline (*M* = 42.50, *SD* = 8.70) to follow-up (*M* = 41.33, *SD* = 9.09; Cohen *d* = 0.13).	Good range of outcome measures. **Limitations:** Experimental groups too similar. Lack of control with the task being over 4 days. No description of randomization process. Effect sizes not reported Large drop out (n=47).	**Overall: Strong**
7. Kraaij et al. (2009) Holland	RCT	44 adults (*M* age = 49) with HIV and mild depression (score above 7 on HADS). 39 male and 5 female participants.	Structured Writing Intervention (SWI). Four weekly 30-minute writing sessions over a 4-week period. Write about their deepest thoughts and feelings around an emotionally significant topic and/or HIV-status. ii)	i)Wait-list control group. ii) CBT computerised self-help program (CBS). This included relaxation exercises, identifying irrational cognitions and goal planning. 4 days a week, 1 hour each day, for a period of 4 weeks.	HADS-D	Significant difference observed in CBS condition baseline depression scores (*M* = 7.31, *SD* = 4.53) to post-test depression scores (*M* = 4.69, *SD* = 4.05; Cohen *d* = 0.6) compared to WLC baseline (*M* = 8.0, *SD* = 3.32) to post-test depression scores (*M* = 7.73, *SD* = 3.88; Cohen *d* = 0.07). No significant difference was found between SWI pre-scores (*M* = 8.13, *SD* = 4.22) to post depression scores (*M* = 7.06, *SD* = 4.81; Cohen *d* = 0.2) and WLC, or SWI and CBS.	**Strengths:** Comparing EW to an evidence-based CBT intervention. Clinical sample. **Limitations:** Wait-list control. Small sample. Generalisability low with regards to only HIV population and male. Only one measure of mood. Time spent on CBT program and writing not matched. Effect sizes not reported.	A – Weak B – Strong C – Moderate D – Moderate E – Strong F – Weak **Overall: Weak**

8. Krpan et al. (2013) United States	RCT	40 students, mean age 28.5 (31 females, 9 males). Diagnosis of Major Depressive Disorder (MDD) as determined via sample inclusion criteria and BDI and PHQ-9.	EW – write about their deepest thoughts and feelings about an extremely important emotional issue that had affected them in their life. Three 20 minute writing sessions over three days.	Control condition: write about how they organise their day in a non-emotional manner. Three 20 minute writing sessions over three days.	BDI and PHQ-9	Repeated measures ANOVA on BDI revealed significant effect of time and condition (η2 = .21), but no significant time x group interaction. Repeated measures ANOVA revealed significant improvement on PHQ-9 scores pre to post in both conditions (η2 = .27), but no significant time x group interaction. EW group significantly less depressed than the control group on PHQ-9 scores at post-test (η2 = .17) and follow-up (η2 =.12).	**Strengths:** Clinical sample . Results having clinical implications due to sample. Use of two standardized measures of depression. Researchers using two attempts at finding an effect without adjusting critical alpha level for p-value. **Limitations:** Small sample, mostly female. No description of randomization process. No check of the content of writing.	A – Moderate B – Strong C – Strong D – Moderate E – Strong F – Strong **Overall: Strong**
9. Lee et al. (2016) South Korea	RCT	68 adult students (male = 41, female = 27) aged between 19-33 (*M* = 24) with moderate depression determined via self-report measures (CES-D score higher than 15).	Online EW about emotions and thoughts from their day (e.g., “please write about today’s events that mainly affected your mood and your feelings around this event”). This was performed daily for 20 minutes over a two-week period.	Control condition: online neutral writing about randomly selected news articles from that day covering politics, entertainment and sports. This was performed daily for 20 minutes over a two-week period.	CES-D	EW significantly reduced depressive symptoms compared to control condition. Changes in CES-D scores were significantly negatively correlated with number of emotional words (*r* = .33), negative words (*r* = -.32), and cognitive words (*r* = .38).	**Strengths** Novel use of online EW via an Emotional Diary App within Facebook. Daily writing as opposed to weekly – increase in validity. **Limitations** Topics participants write about in control condition could be emotive (e.g., current affairs, politics, news). Drop-out rate of 35% (24	A – Weak B – Strong C – Moderate D – Moderate E – Weak F – Moderate **Overall: Weak**

							out of 68). No description of randomization process. No report of validity or reliability of measures, post-intervention means and standard deviations, or effect sizes in results.	
10. Meston et al. (2013) United States	RCT	91 female adults (*M* age = 33) with history of childhood sexual abuse, PTSD and depression via self-report measures (BDI-II score greater than 14). 16.5% of the sample also had a clinical diagnosis of depression.	Two intervention conditions: i) Experimental treatment: Sexual Schema Focussed EW (e.g., typing content around how their abuse affected their sexuality, sexual interest, self-beliefs and their feelings about this). ii) Known active treatment: trauma-focussed EW via typing on a computer (e.g., write about deepest thoughts and feelings from a traumatic event in your life). Two 30-minute writing sessions each week in a lab setting, for 6 weeks.	No traditional control group .	BDI-II	Depression (BDI-II scores) improved significantly from pre-treatment (*M* = 18.0) to post-treatment (*M* = 12.0). No significant difference between conditions.	**Strengths:** Large and clinical sample. Longer-term follow up (6 month). **Limitations:** Analysis weak and report of statistics poor (e.g., no effect sizes, means or standard deviations). No control condition. Limited generalisability to wider depressive population (e.g., use of female, survivors of sexual abuse). No description of randomization process.	A – Weak B – Strong C – Moderate D – Moderate E – Weak F – Weak **Overall: Weak**

11. Sloan, Marx, & Greenberg (2011) United States	RCT	42 undergraduate students (*M* age = 18.9) with mild depression via self-report measure outcomes (BDI-II score greater than 14).	Traditional EW. Participants asked to write about the most traumatic experience of their lives with as much emotion and feeling as possible. Participants advised to write about the same event at each session. Three 20 minute writing sessions over three days.	Control condition: neutral writing about how participants spent their time each day without describing any emotion or opinions. Three 20 minute writing sessions over three days.	BDI-II	No significant main effect of time or interaction between time and condition was found for the depression outcome measure between the expressive writing group (*M* = 17.2, *SD* = 10.8; *M* = 17.8, *SD* = 13.3; Cohen *d* = 0.04) and the control group (*M* = 19.0, *SD* = 11.9; *M* = 15.4, *SD* = 12.1; Cohen *d* = 0.2).	**Strengths** Followed traditional Pennebaker study set-up. Use of follow-up. **Limitations** Small sample, no sex demographics. BDI-II only completed at baseline and 1 month follow-up. Mood measure immediately following EW required. No description of randomization process.	A – Moderate B – Strong C – Strong D – Moderate E – Weak F – Strong **Overall: Moderate**

Note: *QATQS ratings: A = Selection Bias, B = Study Design, C = Confounders, D = Blinding, E = Data Collection Method, F = Withdrawals and Dropouts. EW = Expressive Writing, PHQ-9 = Patient Health Questionnaire, BDI = Beck Depression Inventory, SDS = The Zung Self-Rating Depression Scale, SCID-I = Structured Clinical Interview for the DSM-IV-TR assessment for depression, DAAS = Depression Anxiety Stress Scales, CES-D = Centre for Epidemiologic Studies Depression Scale.

Design of Studies

All eleven studies were Randomized Controlled Trials (RCT), although differed more clearly on quality, intervention, use of control group and sample. Interventions consisted mostly of traditional expressive writing (e.g., write for 20 minutes across three sessions). Control conditions consisted largely of non-emotive writing or a wait-list control. Other alternative intervention conditions consisted of expressive writing with emotional instructions, sexual schema focussed expressive writing, emotional expression through speech, cognitive informed therapy, and online forum discussions. Studies will be sub-divided and discussed based on these differences.

Quality of Studies

The QATQS score ranged from weak (total of 4), moderate (3) and strong (4) overall ratings. A clear area of strength was that all studies were randomized. The QATQS rating system only allows studies to receive overall strong ratings if the design used was randomised or a controlled clinical trial. However, the quality within an RCT can differ greatly, thus some of these scores may mask the underlying limitations of the study. For example, nine studies did not specify their randomization process. The tool revealed a wide range of bias in studies with seven scoring as moderate and four scoring as weak. There was a lack of attention to the blinding process when carrying out experiments, with no studies receiving a strong score. This could be an issue as experimenters may have behaved differently to participants in different conditions. With regards to methodologies and analyses, use of statistical testing was appropriate, although only four studies reported effect sizes in

their results. Effect sizes for use of expressive writing in reducing depressive symptoms were moderate to large. In comparison, do-nothing control (e.g., wait-list) and active-control conditions (e.g., writing about a non-emotive neutral topic) displayed a small/low effect size at reducing depressive symptoms, with no clear relationship between type of control condition and effect size.

Article Reviews

This section will review the articles in brief, grouping them together to first look at the studies which compared expressive writing with a single control condition (expressive writing vs. control), followed by studies which included an additional alternative intervention condition (expressive writing vs. control vs. alternative intervention), followed by studies which did not use a control condition (expressive writing vs. alternative intervention).

Traditional Expressive Writing Manipulation (**expressive writing vs. control).** Graf et al.'s (2009) and Krpan et al.'s (2013) studies demonstrated statistically significant differences with a moderate to large effect size when comparing the control condition to expressive writing at reducing depressive symptoms (DASS, PHQ-9, BDI) pre to post treatment in a clinical population of undergraduate students. expressive writing was shown to be significantly more effective at reducing depressive symptoms compared to neutral writing about daily plans in a non-emotive manner. However, these studies only measured undergraduate students with a diagnosis of depression, thus leading to a problem with generalisability to the wider depressive population. Koopman et al. (2005) carried out a similar study but

on a population of female survivors of domestic violence but found no significant difference in reductions in depression between expressive writing and neutral writing. However, interestingly, women who were more depressed at baseline demonstrated significantly greater reductions in depression when assigned to expressive writing compared to neutral writing. This could suggest that the benefits gained from expressive writing may correlate with level of depressive symptoms (e.g., higher depression leading to expressive writing showing a greater effect). Although, it should be acknowledged that depressive symptoms were high overall in this sample prior to intervention, which may explain why no significant differences were found, in that expressive writing is instead less effective for women who are severely depressed. This study used a small, female only sample and intervention was carried out in a lab setting, restricting its ecological validity.

Sloan, Marx, and Greenberg (2011) used a non-clinical undergraduate sample, but who scored in the mild to moderate range for depression on the BDI-II, to measure the effects of three 20-minute expressive writing sessions, compared with a neutral writing control, completed at participants' homes over three days. Contrary to previous findings, no significant difference was found in reduction of depressive symptoms (BDI-II) for expressive writing compared to control at baseline to follow-up. This could be explained by a weak QATQS score for data collection. BDI-II was only completed at baseline and at one-month follow-up, whereas findings could look different if mood measures were completed immediately following intervention. This also

illustrates the limitations of previous studies, in that the effect of expressive writing could be very short-lived.

Ahmadi et al. (2010) compared expressive writing to expressing emotion through speech at reducing depression scores (BDI) in undergraduate students, a less traditional comparison. Depression scores in the expressive writing condition showed a significantly greater reduction from baseline to post-test compared to the speech condition baseline to post-test scores. This shows there may be something specific about emotional expression through the act of writing, rather than simply expression alone accounting for change. This claim could have been explored further if the authors had included a non-emotive writing control condition as well. Ahmadi et al. (2010) failed to report the effect sizes in their results. Also, the study only used 33 participants, all of which were undergraduates and mostly female, both impacting generalizability and uncertainty about the magnitude of effect and whether it would replicate in larger samples. Findings from this section were consistent in that expressive writing outperformed an active control condition in alleviating depressive symptoms.

Additional Manipulation Intervention (expressive writing vs. control vs. alternative intervention). Kraaij et al. (2009) compared an evidence-based intervention for depression, CBT computerised self-help (CBS), to traditional online expressive writing and a wait-list control group within a male population with HIV and mild depression (assessed using the HADS). Expressive writing and CBS were significantly more effective at reducing depression compared to wait-list control with a medium effect size,

although no significant difference was observed between expressive writing and CBS. This study shows how expressive writing can compare to an evidence-based CBT intervention in alleviating depressive symptoms in a clinical sample, and likely be considerably less expensive to deliver. However, the study's use of a wait-list control merely shows that these interventions are more effective than essentially doing nothing (e.g., expressive writing did not outperform an active control). Also, time spent on CBS program and expressive writing was not matched, with participants in the CBS program spending more time on task, which makes the expressive writing significant results more impressive considering less time was spent on the task.

Kovac and Range (2002) carried out a similar study on undergraduate students with an active control group (write about bedroom layout), compared to expressive writing with cognitive change instructions and traditional expressive writing. Unlike previous research, no significant reduction in depression scores (on the Zung SDS) was observed in any of the three conditions. These results may be due to the severity of depression in the measured sample, considering the lack of research around whether expressive writing can be helpful for individuals with severe depression. Baum and Rude (2013) also carried out a study on undergraduates, which was highly similar to Kovac and Range's (2002) research; however, instead of additional cognitive instructions in one of the expressive writing conditions, they incorporated additional emotional-acceptance and compassion instructions. Findings showed that expressive writing with emotion-acceptance instructions and traditional expressive writing to be more effective

than control (writing about daily routine). Also, expressive writing with emotional acceptance was shown to be significantly more effective at reducing depression scores (CES-D) than traditional expressive writing. This study used a large sample (*N* = 218), although ratio to female (158) to male (60) was very unbalanced, which appears to be a recurring limitation of expressive writing samples. Also, the study procedures lacked control as all interactions with participants and data collection were carried out online, thus researchers were unable to standardize time spent on the task or the environment it was carried out. Overall, these results show expressive writing to compare to already established interventions (e.g., CBS), and again to outperform control conditions in alleviating depressive symptoms. Interestingly, it also shows how traditional expressive writing can be improved by involving emotional-acceptance instructions (Kovac & Range, 2002) which are governed by self-compassion theory (Leary & Tate, 2007) and mindfulness (Teasdale, Segal, & Williams, 2006).

No traditional control condition (expressive writing vs. alternative intervention). Two studies did not use a traditional active control group (e.g., non-emotive writing) or an inactive control group (e.g., wait-list, do-nothing) but instead compared traditional expressive writing with an alternative intervention. Meston et al.'s (2013) study looked at a clinical sample of female survivors of childhood sexual abuse and complex trauma, with self-reported mild to moderate depression (BDI-II). They found depression scores to significantly improve for both the expressive writing and sexual schema focussed expressive writing condition. However, no significant difference was

found between the two conditions, showing that when expressive writing is more structured around the individual's presentation, it does not appear to have any added benefit. This study has limited generalizability to the wider depressive population. A recent study by Dean, Potts, and Barker (2016) divided participants into two conditions: i) online expressive writing twice a week, and ii) use of an internet support group twice a week (e.g., writing in forums, starting own threads) to see the effect on depressive symptoms over a six-month period. Participants' depression scores in both conditions significantly improved, but there was no significant difference between expressive writing and internet support groups. A strength of this study was that it used a large sample (*N* = 204). However, experimental control was poor, considering the level of engagement in the expressive writing and internet support groups task was measured via retrospective self-reports, thus prone to bias. There was also a large attrition rate in the internet support group condition, which may imply its effectiveness being misrepresented in results due to missing data from those who dropped-out and no intention-to-treat analysis conducted.

Discussion

Summary of Evidence

The findings in this review suggest that expressive writing can alleviate depressive symptoms in participants who meet criteria for significant levels of depressive symptoms. This was largely consistent across all studies with

variations in number of conditions, control groups, format of expressive writing (e.g., typing, writing) and outcome measures (e.g., BDI, PHQ, CES-D). Expressive writing particularly outperformed neutral, non-emotive writing in alleviating depressive symptoms. The findings' generalisability extends to the samples used in the reviewed studies, which predominantly consisted of female undergraduate students. A small selection of the studies used a clinical sample of participants attending outpatient psychotherapy, where expressive writing demonstrated positive benefits compared to a control condition. Furthermore, findings from the studies discussed also demonstrated additional benefits of expressive writing, such as significantly reducing anxiety, suicidality and post-traumatic stress symptoms. Studies failed to test mediators of expressive writing's effect, such as the person's state rumination, level of mindfulness, and level of restraint/avoidance of emotions. Studies should explore such mediators further to see whether findings are consistent with theories on the effect of expressive writing. For example, Pennebaker's (1989) Inhibitions Theory could have been supported by study data, as it may have been the relaxation of constraint of thoughts and feelings by participants in expressive writing that led to the significant change in depressive symptoms. By not testing mediators of expressive writing, studies were unable to link findings to theory, or rule out other possible theoretical interpretations for the efficacy of expressive writing. For instance, people may have had expectations that expressive writing should help and thus reported fewer symptoms purely because of this rather than therapeutic change.

As opposed to expressive writing simply being more effective than non-emotive writing at reducing depressive symptoms, it also performed well when compared to evidence-based interventions such as a computerised CBT self-help program (Kraaij et al., 2009). A number of studies asked participants to type their expressive writing online. It could be argued that this moves away from the traditional, well researched, Pennebaker (1991) method of expressive writing. However, expressive writing interventions taking place online could increase accessibility for individuals. The content of words used could also be more easily monitored on computers, which is often analysed in expressive writing research (Lee et al., 2016). An overall finding from this review was that expressive writing alleviated depressive symptoms consistently, regardless of whether the expressive piece was hand-written or typed. In relation to the quality of designs of studies as assessed by the QATQS, strength of design was not associated with the strength of evidence for expressive writing. Thus, it does not seem as though the apparent efficacy of expressive writing could be attributable to poor quality designs (e.g., unclear randomisation procedure, high attrition rates) with threatened internal validity.

Generalizability of the findings is limited given the small number of studies reviewed and lack of variation in samples used. Almost all studies used a selected sample consisting of Western, Educated, Industrialized, Rich and Democratic participants (WEIRD; Henrich, Heine, & Norenzayan, 2010). For instance, only 2 of the 11 studies were conducted in non-western locations e.g., Iran and South Korea). Although findings in the two eastern

locations were consistent with the western studies findings (e.g., expressive writing reduced depressive symptoms), inferences made from this review to other populations should be avoided as it remains unclear whether expressive writing is effective cross-culturally. Across samples, some participants were clinically diagnosed with depression, whereas in other studies individuals were included if they met the cut-off for depression when completing a self-report measure (e.g., BDI). This may lead to a large difference in severity of symptoms across the samples used, although the strength of evidence across the review did not appear dependent on severity of symptoms or whether the population was clinical or non-clinical. Future research may look to expand on sample size and diversity, whilst looking for differences in expressive writing effects around severity of condition (e.g., studies looking at severe depression), as additional problems come with severe depression (e.g., poor concentration, fatigue) which may hinder the efficacy of expressive writing.

Limitations

This review has several limitations which should be acknowledged when making any inferences from its findings. An initial problem consisted of the adjustments that had to be made when searching on certain databases (Web of Science, PubMed, Creativity Research Journal) due to the large number of results (approx. 4,000) when searching in ‘all fields’. Thus, on these databases, the search consisted of title only. Although the search had to be carried out in this manner due to resource limitations, it was at the risk of potentially missing relevant articles. To account for this, additional time was

put into analysing the reference lists of the 38 full-text articles read. Similarly, due to limited resources, grey literature was not searched. There is arguably an inherent risk of bias in published papers towards reporting positive results. Future reviews could be improved by including unpublished papers and involve a research team to help the screening of a large number of results when searching in ‘all fields’ on the databases, although this would incur additional costs in money and time.

The findings of this review are limited due to the quality of studies being mixed as measured by the QATQS rating system, with scores of weak, moderate and strong across the eleven articles. With the QATQS favouring RCT designs, some studies received overall scores of strong which arguably misleads the reader to the quality of the article. An area not covered by the QATQS rating system is that of effect sizes in results of studies, which were often not published in reviewed articles, limiting the ability to make inferences about the importance of findings. Other areas of weaknesses the QATQS identified were unclear randomization processes: use of blinding in studies, awareness of potential confounds and drop-out scores. Thus, the use of a methodologically sound design in an RCT across all articles should not blind the reader to other concerns.

A further limitation of this evidence base is its inability to provide any deepened understanding of the mechanisms through which expressive writing leads to positive change. For example, the Cognitive Changes Theory (Smyth, True, & Souto, 2001) suggests that expressive writing allows for cognitive restructuring, thus alleviating psychological distress. The studies in

this review did not look at how or why expressive writing reduced depressive symptoms, but instead merely whether it did or not. Rather than traditional experiments in which expressive writing is compared to non-emotive writing, research could look deeper into the functions of expressive writing by comparing different forms of expressive writing against one another. For instance, Hayes et al. (2007) suggest that repeatedly writing about the same stressful event reduces negative affect. However, some studies in this review did not specify whether participants were asked to write about the same event or a different event, or even whether this was an instruction they gave, thus generally lacking control over whether content of expressive writing is novel or repeated. This is important, as Hayes et al. (2007) argues that expressive writing is effective as an exposure-based model, in that repeatedly writing about the same event leads to the reduction of negative affect around said event, by breaking the cycle of avoidance. To measure this, research could simply compare two groups, one which writes about the same event over consecutive days, and the other where participants write about different events. This would be a start to understanding mechanisms of change.

Future Implications

The findings from this review support the notion that expressive writing can help alleviate depressive symptoms in people with clinically significant depression either because they meet diagnostic criteria or score above the cut-off on standardized psychometric measures (e.g., BDI). Findings were consistent across studies and therefore clinical implications could involve

routine guidance by clinicians to ask clients to engage in expressive writing as an intervention used in-between therapeutic sessions, although caution in this approach is required considering many studies were not conducted in clinical populations undergoing therapy. Some therapeutic models already incorporate writing of past trauma and emotions into their structure such as NET (Schauer, Neuner, & Elbert, 2011) and CAT (Ryle & Kerr, 2002). Thus, expressive writing may be helpful for those undergoing therapy involving processing trauma. Expressive writing could be made more explicit within practitioner manuals, considering it is an easy-to-follow and cost-effective measure.

It could be argued that expressive writing is even more accessible in modern times due to computer access. For example, although traditional expressive writing is not concerned with spelling or grammar and even reassures the participant of this, some may still find hand-writing and spelling difficult, whereas typing may overcome this potential barrier via the spell-check function. Furthermore, expressive writing could be offered as a guided self-help tool for clients on a wait-list for psychological therapy as the research suggests that expressive writing is significantly more efficacious than being on a wait-list (Kraaij et al., 2009)**.** One reason this may have not been taken up as of yet is due to the majority of expressive writing research being conducted in an undergraduate student population. Expressive writing literature should focus on building its findings on adults with a diagnosis of depression.

A potential barrier for expressive writing being used across settings is issues of confidentiality and privacy. Some patients may fear that a family member or friend would find their expressive writing. More specifically, secondary-care mental health settings where psychosis and paranoia are highly prevalent may lead to resistance from the patient due to the paranoia and uncertainty around how the expressive writing piece would be used. Additionally, there are a number of people who find it difficult to read and write (e.g., learning disabilities and dyslexia). This could also act as a barrier to accessing expressive writing. Therefore, clinicians should first gauge whether expressive writing is suitable for the clientele they are working with, and then ensure a collaborative discussion around how expressive writing can be used.

In relation to a lack of long-term studies on expressive writing, or even the use of follow-ups, there appears to be a gap in the literature around what individuals are meant to do with the expressive writing piece after they have written it. For instance, it is unknown whether individuals should shred it, store it, or frequently re-read it to help process emotions or re-structure cognitions (Hayes et al., 2007). It would be interesting to examine whether the benefits of expressive writing are dependent on the amount of times the person re-reads their journal entries. One could hypothesise that increasing exposure to the traumatic memories through reflective reading may help process emotions further and acknowledge potential cognitive biases which are trying to be resolved. However, the opposite may occur, in that too much emotional processing leads to rumination. This may be particularly relevant for clients who are more resistant to expressing and processing emotions, and therefore

come to therapy with strong defences. This is important, as it may be the very nature of how expressive writing is offered and delivered to clients with regards to whether it leads to positive change.

Conclusion

This review revealed that expressive writing consistently alleviated depressive symptoms for adults with self-reported depression via standardised measures, and for adults with a clinical diagnosis of MDD. All articles used an RCT design to compare expressive writing against an alternative intervention or control group. Expressive writing was found to outperform active and inactive control groups and at least match alternative evidence-based interventions. Expressive writing appears to be an efficient, easy-to-use tool that can be utilised during psychological therapy and beneficial in alleviating depressive symptoms in a non-clinical population. However, limitations exist around the generalisability and external validity of expressive writing findings due to samples predominately consisting of students in laboratory settings. The long-term benefits of expressive writing are yet to be determined due to the lack of follow-up within research. Further research in this area will help provide clues about the psychological mechanisms involved in depression and help broaden the potential interventions for alleviating symptoms.

References

Ahmadi, M., Abdollahi, M. H., Ramezani, V., & Heshmati, R. (2010). The impact of written emotional expression on depressive symptoms and working memory capacity in Iranian students with high depressive symptoms. *Social and Behavioural Sciences, 5*, 1610-1614. doi:10.1016/j.sbspro.2010.07.334

Baum, E. S., & Rude, S. S. (2013). Acceptance-enhanced expressive writing prevents symptoms in participants with low initial depression. *Cognitive Therapy and Research, 37*, 35-42. doi:10.1007/s10608-012-9435-x

Beck, A. T., & Beamesderfer, A. (1974). Assessment of depression: The depression inventory. *Modern Trends in Pharmacopsychiatry Psychological Measurements in Psychopharmacology, 7,* 151-169. doi:10.1159/000395074

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

Bhui, K., & McKenzie, K. (2008). Rates and risk factors by ethnic group for suicides within a year of contact with mental health services in England and Wales. *Psychiatric Services, 59*, 414–420.

Campbell, R. S., & Pennebaker, J. W. (2003). The secret life of pronouns. *Psychological Science, 14*, 60-65. doi:10.1111/1467-9280.01419

Dean, J., Potts, H. W., & Barker, C. (2016). Direction to an internet support group compared with online expressive writing for people with depression and anxiety: A randomized trial. *Journal of Medical Internet Research Mental Health, 3*, (2):e12. doi:10.2196/mental.5133

Frattaroli, J. (2006). Experimental disclosure and its moderators: A meta-analysis. *Psychological Bulletin, 132*, 823-865. doi:10.1037/0033-2909.132.6.823

Gortner, E., Rude, S. S., & Pennebaker, J. W. (2006). Benefits of expressive writing in lowering rumination and depressive symptoms. *Behaviour Therapy, 37*, 292-303. doi:10.1016/j.beth.2006.01.004

Graf, M. C., Gaudiano, B. A., & Geller, P. A. (2008). Written emotional disclosure: A controlled study of the benefits of expressive writing homework in outpatient psychotherapy. *Psychotherapy Research, 18*, 389-399. doi:10.1080/10503300701691664

Hayes, A. M., Beevers, C. G., Feldman, G. C., Laurenceau, J., & Perlman, C. (2005). Avoidance and processing as predictors of symptom change and positive growth in an integrative therapy for depression. *International Journal of Behavioural Medicine, 12*, 111-122. doi:10.1207/s15327558ijbm1202_9

Hayes, A. M., Laurenceau, J., Feldman, G., Strauss, J. L., & Cardaciotto, L. (2007). Change is not always linear: The study of nonlinear and discontinuous patterns of change in psychotherapy. *Clinical Psychology Review, 27*, 715-723. doi:10.1016/j.cpr.2007.01.008

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioural and Brain Sciences, 33,* 61-83. doi: 10.1017/S0140525X0999152X

Higgins, J. P., & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions*. Wiley-Blackwell.

Hofmann, S. G., Asnaani, A., Vonk, I. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioural therapy: A review of meta-analyses. *Cognitive Therapy and Research, 36*, 427-440. doi:10.1007/s10608-012-9476-1

Improving Access to Psychological Therapies (2013). Key performance indicators. Retrieved from: https://digital.nhs.uk/data-and-information/publications/statistical/improving-access-to-psychological-therapies-key-performance-indicators-iapt-kpis/improving-access-to-psychological-therapies-key-performance-indicators-iapt-kpis-final-q4-2012-13.

Kazantzis, N., Lampropoulos, G. K., & Deane, F. P. (2005). A national survey of practicing psychologists' use and attitudes toward homework in psychotherapy. *Journal of Consulting and Clinical Psychology, 73*, 742-748. doi:10.1037/0022-006x.73.4.742

Kazdin, A. E., & Blase, S. L. (2011). Rebooting psychotherapy research and practice to reduce the burden of mental illness. *Perspectives on Psychological Science, 6*, 21-37. doi:10.1177/1745691610393527

Kacewicz, E., Slatcher, R. B., & Pennebaker, J. W. (2007). Expressive writing: An alternative to traditional methods. In L. L'Abate(Ed.), *Low-cost interventions to promote physical and mental health: Theory, research, and practice* (pp. 271-284). New York: Springer.

Klein, K., & Boals, A. (2001). Expressive writing can increase working memory capacity. *Journal of Experimental Psychology: General, 130*, 520-533. doi:10.1037//0096-3445.130.3.520

Koopman, C., Ismailji, T., Holmes, D., Classen, C. C., Palesh, O., & Wales, T. (2005). The effects of expressive writing on pain, depression and posttraumatic stress disorder symptoms in survivors of intimate partner violence. *Journal of Health Psychology, 10*, 211-221. doi:10.1177/1359105305049769

Kovac, S. H., & Range, L. M. (2002). Does writing about suicidal thoughts and feelings reduce them? *Suicide and Life-Threatening Behaviour, 32*, 428-440. doi:10.1521/suli.32.4.428.22335

Kraaij, V., Shelley M. C., Van Der Veek, G. N., Schroevers, M., Witlox, R., & Maes, S. (2009). Coping, goal adjustment, and psychological well-being in HIV-infected men who have sex with men. *AIDS Patient Care and STDs, 22*, 395-402. doi:10.1089/apc.2007.0145

Krpan, K. M., Kross, E., Berman, M. G., Deldin, P. J., Askren, M. K., & Jonides, J. (2013). An everyday activity as a treatment for depression: The benefits of expressive writing for people diagnosed with major depressive disorder. *Journal of Affective Disorders, 150*, 1148-1151. doi:10.1016/j.jad.2013.05.065

Leary, M. R., & Tate, E. B. (2007). The multi-faceted nature of mindfulness. *Psychological Inquiry, 18,* 251-255. doi:10.1080/10478400701598355

Lee, S. W., Kim, I., Yoo, J., Park, S., Jeong, B., & Cha, M. (2016). Insights from an expressive writing intervention on Facebook to help alleviate depressive symptoms. *Computers in Human Behaviour, 62*, 613-619. doi:10.1016/j.chb.2016.04.034

Lepore, S. J. (1997). Social–environmental influences on the chronic stress process. In B. H. Gottlieb (Ed.), *The Plenum series on stress and coping. Coping with chronic stress* (pp. 133-160). New York: Plenum Press. doi: 10.1007/978-1-4757-9862-3_5

Lovibond, P., & Lovibond, S. (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*, 335-343. doi:10.1016/0005-7967(94)00075-u

Mehl, M. R., Vazire, S., Ramirez-Esparza, N., Slatcher, R. B., & Pennebaker, J. W. (2007). Are women really more talkative than men? *Science, 317*, 82. doi:10.1126/science.1139940

Mental Health Foundation. (2016). Fundamental facts about mental health. Retrieved from: https://www.mentalhealth.org.uk/sites/default/files/fundamental-facts-about-mental-health-2016.pdf.

Meston, C. M., Lorenz, T. A., & Stephenson, K. R. (2013). Effects of expressive writing on sexual dysfunction, depression, and PTSD in women with a history of childhood sexual abuse: Results from a randomized clinical trial. *The Journal of Sexual Medicine, 10*, 2177-2189. doi:10.1111/jsm.12247

Moher, D. (2015). *Reporting health research*. New York: John Wiley.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151*, 264-269. doi:10.7326/0003-4819-151-4-200908180-00135

NICE (2012). Identifying the evidence and literature searching. Retrieved from: https://www.nice.org.uk/process/pmg6/chapter/identifying-the-evidence-literature-searching-and-evidence-submission.

Nielson, D. M., Smith, T. A., Sreekumar, V., Dennis, S., & Sederberg, P. B. (2015). Human hippocampus represents space and time during retrieval of real-world memories. *Proceedings of the National Academy of Sciences, 112*, 11078-11083. doi:10.1073/pnas.1507104112

Pan, A., Lucas, M., Sun, Q., Dam, R. M., Franco, O. H., Willett, W. C., & Hu, F. B. (2011). Increased mortality risk in women with depression and diabetes mellitus. *Archives of General Psychiatry, 68*, 42-50. doi:10.1001/archgenpsychiatry.2010.176

Pennebaker, J. W., & Francis, M. E. (1996). Cognitive, emotional, and language processes in disclosure. *Cognition and Emotion, 10*, 601-626. doi:10.1080/026999396380079

Pennebaker, J. W. (1989). Confession, inhibition, and disease. *Advances in Experimental Social Psychology, 22,* 211-244. doi:10.1016/s0065-2601(08)60309-3

Pennebaker, J. W., & Beall, S. K. (1986). Emotional disclosure protocol. *PsycTESTS Dataset*. doi:10.1037/t08134-000

Pennebaker, J. W., Kiecolt-Glaser, J. K., & Glaser, R. (1988). Disclosure of traumas and immune function: Health implications for psychotherapy. *Journal of Consulting and Clinical Psychology, 56*, 239-245. doi:10.1037//0022-006x.56.2.239

Pennebaker, J. W., & Seagal, J. D. (1999). Forming a story: The health benefits of narrative. *Journal of Clinical Psychology, 55*, 1243-1254. doi:10.1002/(sici)1097-4679(199910)55:103.0.co;2-n

Pennebaker, J. W. (1997). Writing about emotional experiences as a therapeutic process. *Psychological Science, 8*, 162-166. doi:10.1111/j.1467-9280.1997.tb00403.x

Petrie, K. J., Booth, R. J., & Pennebaker, J. W. (1998). The immunological effects of thought suppression. *Journal of Personality and Social Psychology, 75*, 1264-1272. doi:10.1037//0022-3514.75.5.1264

Radloff, L. S. (1977). The CES-D scale. *Applied Psychological Measurement, 1*, 385-401. doi:10.1177/014662167700100306

Ryle, A., & Kerr, I. B. (2002). *Introducing cognitive analytic therapy: Principles and practice*. New York: Wiley.

Teasdale, J. D., Segal, Z. V., & Williams, J. M. (2006). Mindfulness training and problem formulation. *Clinical Psychology: Science and Practice, 10*, 157-160. doi:10.1093/clipsy.bpg017

Schauer, M., Neuner, F., & Elbert, T. (2011). *Narrative exposure therapy: A short term treatment for traumatic stress disorders* (2nd ed.). Cambridge, MA: Hogrefe Publishing.

Slatcher, R. B., & Pennebaker, J. W. (2007). Emotional expression and health. In: S. Ayers, A. Baum, C. McManus, S. Newman, K. Wallston, J. Weinman & R. West (Eds.), *Cambridge Handbook of Psychology, Health and Medicine* (pp. 84-88). Cambridge: Cambridge University Press.

Sloan, D. M., & Marx, B. P. (2004). A closer examination of the structured written disclosure procedure. *Journal of Consulting and Clinical Psychology, 72*, 165-175. doi:10.1037/0022-006x.72.2.165.

Sloan, D. M., Marx, B. P., Epstein, E. M., & Lexington, J. M. (2007). Does altering the writing instructions influence outcome associated with written disclosure? *Behaviour Therapy, 38*, 155-168. doi:10.1016/j.beth.2006.06.005

Sloan, D. M., Marx, B. P., & Greenberg, E. M. (2011). A test of written emotional disclosure as an intervention for posttraumatic stress disorder. *Behaviour Research and Therapy, 49*, 299-304. doi:10.1016/j.brat.2011.02.001

Smyth, J., True, N., & Souto, J. (2001). Effects of writing about traumatic experiences: the necessity for narrative structuring. *Journal of Social and Clinical Psychology, 20*, 161-172. doi:10.1521/jscp.20.2.161.22266

Smyth, J. M. (1998). Written emotional expression: Effect sizes, outcome types, and moderating variables. *Journal of Consulting and Clinical Psychology, 66*, 174-184. doi:10.1037//0022-006x.66.1.174

Spiers, N., Qassem, T., Bebbington, P., Mcmanus, S., King, M., Jenkins, R., & Brugha, T. S. (2016). Prevalence and treatment of common mental disorders in the English national population, 1993–2007. *British Journal of Psychiatry, 209*, 150-156. doi:10.1192/bjp.bp.115.174979

Spitzer R., Kroenke, K., & Williams, J. (1999). Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *Journal of the American Medical Association, 282,* 1737-1744.

Thomas, B., Ciliska, D., Dobbins, M., & Micucci, S. (2004). A process for systematically reviewing the literature: Providing the research evidence for public health nursing interventions. *Worldviews on Evidence-Based Nursing, 1*, 176-184. doi:10.1111/j.1524-475x.2004.04006.x

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica, 67*, 361-370.

Zung, W. W. (1986). Zung Self-rating Depression Scale and Depression Status Inventory. *In: T. A. Ban & N. Sartorius (Eds.), Assessment of depression* (pp. 221–231). New York: Springer Verlag. doi: 10.1007/978-3-642-70486-4_21

Appendices

Appendix 1: Quality Assessment Tool for Quantitative Studies

QUALITY ASSESSMENT TOOL FOR QUANTITATIVE STUDIES

COMPONENT RATINGS

A) SELECTION BIAS

(Q1) Are the individuals selected to participate in the study likely to be representative of the target population?

1 Very likely
2 Somewhat likely
3 Not likely
4 Can't tell

(Q2) What percentage of selected individuals agreed to participate?

1 80 - 100% agreement
2 60 – 79% agreement
3 less than 60% agreement
4 Not applicable
5 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

B) STUDY DESIGN

Indicate the study design

1 Randomized controlled trial
2 Controlled clinical trial
3 Cohort analytic (two group pre + post)
4 Case-control
5 Cohort (one group pre + post (before and after))
6 Interrupted time series
7 Other specify _____
8 Can't tell

Was the study described as randomized? If NO, go to Component C.

No Yes

If Yes, was the method of randomization described? (See dictionary)

No Yes

If Yes, was the method appropriate? (See dictionary)

No Yes

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

C) CONFOUNDERS

(Q1) Were there important differences between groups prior to the intervention?

1 Yes
2 No
3 Can't tell

The following are examples of confounders:

1 Race
2 Sex
3 Marital status/family
4 Age
5 SES (income or class)
6 Education
7 Health status
8 Pre-intervention score on outcome measure

(Q2) If yes, indicate the percentage of relevant confounders that were controlled (either in the design (e.g. stratification, matching) or analysis)?

1 80 – 100% (most)
2 60 – 79% (some)
3 Less than 60% (few or none)
4 Can't Tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

D) BLINDING

(Q1) Was (were) the outcome assessor(s) aware of the intervention or exposure status of participants?

1 Yes
2 No
3 Can't tell

(Q2) Were the study participants aware of the research question?

1 Yes
2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

E) DATA COLLECTION METHODS

(Q1) Were data collection tools shown to be valid?

1 Yes
2 No
3 Can't tell

(Q2) Were data collection tools shown to be reliable?

1 Yes
2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

F) WITHDRAWALS AND DROP-OUTS

(Q1) Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?

1 Yes
2 No
3 Can't tell
4 Not Applicable (i.e. one time surveys or interviews)

(Q2) Indicate the percentage of participants completing the study. (If the percentage differs by groups, record the lowest).

1 80 -100%
2 60 - 79%
3 less than 60%
4 Can't tell
5 Not Applicable (i.e. Retrospective case-control)

RATE THIS SECTION	STRONG	MODERATE	WEAK	
See dictionary	1	2	3	Not Applicable

G) INTERVENTION INTEGRITY

(Q1) What percentage of participants received the allocated intervention or exposure of interest?

1 80 -100%
2 60 - 79%
3 less than 60%
4 Can't tell

(Q2) Was the consistency of the intervention measured?

1 Yes
2 No
3 Can't tell

(Q3) Is it likely that subjects received an unintended intervention (contamination or co-intervention) that may influence the results?

4 Yes
5 No
6 Can't tell

H) ANALYSES

(Q1) Indicate the unit of allocation (circle one)

community organization/institution practice/office individual

(Q2) Indicate the unit of analysis (circle one)

community organization/institution practice/office individual

(Q3) Are the statistical methods appropriate for the study design?

1 Yes
2 No
3 Can't tell

(Q4) Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?

1 Yes
2 No
3 Can't tell

GLOBAL RATING

COMPONENT RATINGS
Please transcribe the information from the gray boxes on pages 1-4 onto this page. See dictionary on how to rate this section.

A	**SELECTION BIAS**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
B	**STUDY DESIGN**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
C	**CONFOUNDERS**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
D	**BLINDING**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
E	**DATA COLLECTION METHOD**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	
F	**WITHDRAWALS AND DROPOUTS**	**STRONG**	**MODERATE**	**WEAK**	
		1	2	3	Not Applicable

GLOBAL RATING FOR THIS PAPER (circle one):

1	STRONG	(no WEAK ratings)
2	MODERATE	(one WEAK rating)
3	WEAK	(two or more WEAK ratings)

With both reviewers discussing the ratings:

Is there a discrepancy between the two reviewers with respect to the component (A-F) ratings?

No Yes

If yes, indicate the reason for the discrepancy

1	Oversight
2	Differences in interpretation of criteria
3	Differences in interpretation of study

Final decision of both reviewers (circle one):

1	**STRONG**
2	**MODERATE**
3	**WEAK**

Quality Assessment Tool for Quantitative Studies Dictionary

The purpose of this dictionary is to describe items in the tool thereby assisting raters to score study quality. Due to under-reporting or lack of clarity in the primary study, raters will need to make judgements about the extent that bias may be present. When making judgements about each component, raters should form their opinion based upon information contained in the study rather than making inferences about what the authors intended.

A) SELECTION BIAS

(Q1) Participants are more likely to be representative of the target population if they are randomly selected from a comprehensive list of individuals in the target population (score very likely). They may not be representative if they are referred from a source (e.g., clinic) in a systematic manner (score somewhat likely) or self-referred (score not likely).

(Q2) Refers to the % of subjects in the control and intervention groups that agreed to participate in the study before they were assigned to intervention or control groups.

B) STUDY DESIGN

In this section, raters assess the likelihood of bias due to the allocation process in an experimental study. For observational studies, raters assess the extent that assessments of exposure and outcome are likely to be independent. Generally, the type of design is a good indicator of the extent of bias. In stronger designs, an equivalent control group is present and the allocation process is such that the investigators are unable to predict the sequence.

Randomized Controlled Trial (RCT)

An experimental design where investigators randomly allocate eligible people to an intervention or control group. A rater should describe a study as an RCT if the randomization sequence allows each study participant to have the same chance of receiving each intervention and the investigators could not predict which intervention was next. If the investigators do not describe the allocation process and only use the words ‘random’ or ‘randomly’, the study is described as a controlled clinical trial.

See below for more details.
Was the study described as randomized?
Score YES, if the authors used words such as random allocation, randomly assigned, and random assignment. Score NO, if no mention of randomization is made.

Was the method of randomization described?

Score YES, if the authors describe any method used to generate a random allocation sequence.

Score NO, if the authors do not describe the allocation method or describe methods of allocation such as alternation, case record numbers, dates of birth, day of the week, and any allocation procedure that is entirely transparent before assignment, such as an open list of random numbers of assignments.

If NO is scored, then the study is a controlled clinical trial.

Was the method appropriate?

Score YES, if the randomization sequence allowed each study participant to have the same chance of receiving each intervention and the investigators could not predict which intervention was next. Examples of appropriate approaches include assignment of subjects by a central office unaware of subject characteristics, or sequentially numbered, sealed, opaque envelopes.

Score NO, if the randomization sequence is open to the individuals responsible for recruiting and allocating participants or providing the intervention, since those individuals can influence the allocation process, either knowingly or unknowingly.

If NO is scored, then the study is a controlled clinical trial.

Controlled Clinical Trial (CCT)
An experimental study design where the method of allocating study subjects to intervention or control groups is open to individuals responsible for recruiting subjects or providing the intervention. The method of allocation is transparent before assignment, e.g., an open list of random numbers or allocation by date of birth, etc.

Cohort analytic (two group pre and post)
An observational study design where groups are assembled according to whether or not exposure to the intervention has occurred. Exposure to the intervention is not under the control of the investigators. Study groups might be non- equivalent or not comparable on some feature that affects outcome.

Case control study
A retrospective study design where the investigators gather ‘cases’ of people who already have the outcome of interest and ‘controls’ who do not. Both groups are then questioned or their records examined about whether they received the intervention exposure of interest.

Cohort (one group pre + post (before and after)
The same group is pretested, given an intervention, and tested immediately after the intervention. The intervention group, by means of the pretest, act as their own control group.

Interrupted time series
A time series consists of multiple observations over time. Observations can be on the same units (e.g., individuals over time) or on different but similar units (e.g., student achievement scores for particular grade and school). Interrupted time series analysis requires knowing the specific point in the series when an intervention occurred.

C) CONFOUNDERS

By definition, a confounder is a variable that is associated with the intervention or exposure and causally related to the outcome of interest. Even in a robust study design, groups may not be balanced with respect to important variables prior to the intervention. The authors should indicate if confounders were controlled in the design (by stratification or matching) or

in the analysis. If the allocation to intervention and control groups is randomized, the authors must report that the groups were balanced at baseline with respect to confounders (either in the text or a table).

D) BLINDING

(Q1) Assessors should be described as blinded to which participants were in the control and intervention groups. The purpose of blinding the outcome assessors (who might also be the care providers) is to protect against detection bias.

(Q2) Study participants should not be aware of (i.e. blinded to) the research question. The purpose of blinding the participants is to protect against reporting bias.

E) DATA COLLECTION METHODS

Tools for primary outcome measures must be described as reliable and valid. If 'face' validity or 'content' validity has been demonstrated, this is acceptable. Some sources from which data may be collected are described below:

Self reported data includes data that is collected from participants in the study (e.g.,completing a questionnaire, survey, answering questions during an interview, etc.).

Assessment/Screening includes objective data that is retrieved by the researchers. (e.g., observations by investigators).

Medical Records/Vital Statistics refers to the types of formal records used for the extraction of the data.

Reliability and validity can be reported in the study or in a separate study. For example, some standard assessment tools have known reliability and validity.

F) WITHDRAWALS AND DROP-OUTS

Score **YES** if the authors describe BOTH the numbers and reasons for withdrawals and drop-outs. Score **NO** if either the numbers or reasons for withdrawals and drop-outs are not reported.

The percentage of participants completing the study refers to the % of subjects remaining in the study at the final data collection period in all groups (i.e. control and intervention groups).

G) INTERVENTION INTEGRITY

The number of participants receiving the intended intervention should be noted (consider both frequency and intensity). For example, the authors may have reported that at least 80 percent of the participants received the complete intervention. The authors should describe a method of measuring if the intervention was provided to all participants the same way. As well, the authors should indicate if subjects received an unintended intervention that may have influenced the outcomes. For example, co-intervention occurs when the study group receives an additional intervention (other than that intended). In this case, it is possible that

the effect of the intervention may be over-estimated. Contamination refers to situations where the control group accidentally receives the study intervention. This could result in an under-estimation of the impact of the intervention.

H) ANALYSIS APPROPRIATE TO QUESTION

Was the quantitative analysis appropriate to the research question being asked?

An intention-to-treat analysis is one in which all the participants in a trial are analyzed according to the intervention to which they were allocated, whether they received it or not. Intention-to-treat analyses are favoured in assessments of effectiveness as they mirror the noncompliance and treatment changes that are likely to occur when the intervention is used in practice, and because of the risk of attrition bias when participants are excluded from the analysis.

Component Ratings of Study:

For each of the six components A – F, use the following descriptions as a roadmap.

A) SELECTION BIAS
Strong: The selected individuals are very likely to be representative of the target population (Q1 is 1) **and** there is greater than 80% participation (Q2 is 1).

Moderate: The selected individuals are at least somewhat likely to be representative of the target population (Q1 is 1 or 2); **and** there is 60 - 79% participation (Q2 is 2). 'Moderate' may also be assigned if Q1 is 1 or 2 and Q2 is 5 (can't tell).

Weak: The selected individuals are not likely to be representative of the target population (Q1 is 3); **or** there is less than 60% participation (Q2 is 3) **or** selection is not described (Q1 is 4); and the level of participation is not described (Q2 is 5).

B) DESIGN
Strong: will be assigned to those articles that described RCTs and CCTs.

Moderate: will be assigned to those that described a cohort analytic study,a case control study,a cohort design or an interrupted time series.

Weak: willbeassignedtothosethatusedanyothermethodordidnotstatethemethodused.

C) CONFOUNDERS

Strong: will be assigned to those articles that controlled for at least 80% of relevant confounders (Q1is2);**or**(Q2is1).

Moderate: will be given to those studies that controlled for 60–79% of relevant confounders (Q1is1)**and**(Q2is2).

Weak: will be assigned when less than 60% of relevant confounders were controlled (Q1is1) **and** (Q2is3) **or** control of confounders was not described (Q1 is 3) **and** (Q2 is 4).

D) BLINDING
Strong: The outcome assessor is not aware of the intervention status of participants (Q1 is 2); **and t**he study participants are not aware of the research question (Q2 is 2).
Moderate: The outcome assessor is not aware of the intervention status of participants (Q1 is 2); **or** the study

participants are not aware of the research question (Q2 is 2); **or b**linding is not described (Q1 is 3 and Q2 is 3). **Weak:** The outcome assessor is aware of the intervention status of participants (Q1 is 1); **and** the study participants are aware of the research question (Q2 is 1).

E) DATA COLLECTION METHODS
Strong: The data collection tools have been shown to be valid (Q1 is 1); **and** the data collection tools have been shown to be reliable (Q2 is 1).
Moderate: The data collection tools have been shown to be valid (Q1 is 1); **and** the data collection tools have not been shown to be reliable (Q2 is 2) **or** reliability is not described (Q2 is 3).
Weak: The data collection tools have not been shown to be valid (Q1 is 2) **or** both reliability and validity described (Q1 is 3 and Q2 is 3).

F) WITHDRAWALS AND DROP-OUTS - a rating of:
Strong: will be assigned when the follow-up rate is 80% or greater (Q2 is 1).
Moderate: will be assigned when the follow-up rate is 60 – 79% (Q2 is 2) **OR** Q2 is 5 (N/A).

Weak: will be assigned when a follow-up rate is less than 60% (Q2 is 3) or if the withdrawals and drop-outs were not described (Q2 is 4).

SCHOOL OF PSYCHOLOGY

DOCTORATE IN CLINICAL PSYCHOLOGY

Empirical Paper

Colouring-in, a Distraction Technique? A study Looking at The Effects of Colouring-in for Adults in Reducing Negative Affect and State Rumination

Trainee Name:	**Neil Drew**
Primary Research Supervisor:	**Dr Nicholas Moberly**
	Senior Lecturer, School of Psychology, University of Exeter
Secondary Research Supervisor:	**Professor Ed Watkins**
	Professor of Experimental and Applied Clinical Psychology, Director of Research for Professional Doctorates, University of Exeter

Target Journal:	Behaviour Research and Therapy
Word Count:	8374 (excluding abstract, table of contents, list of figures and tables, references, appendices)

Abstract

Background

Adult colouring books have become increasingly popular in recent years, with suggestions that they can reduce stress and increase calmness, but there is currently a limited evidence base in this area. This study explored whether colouring-in is more effective than a neutral distraction activity and rumination in improving affect and state rumination after experiencing a laboratory stressor that involved solving difficult anagrams**.**

Method

The study was a mixed (3 x 3) experimental design, with condition as the between-subjects factor (distraction, colouring-in, rumination) and time as the repeated measures factor (baseline, post-stressor, post-manipulation). An undergraduate student sample (*N* = 90) was randomly assigned to one of three experimental conditions. Participants completed the State Rumination Questionnaire (SRQ) and Positive and Negative Affect Schedule (PANAS) to measure the impact of the stressor and the effects of each experimental condition.

Results

After experiencing the stressor, participants showed a significant increase in negative affect ($p <. 001$) and state rumination ($p <. 001$), and a significant reduction in positive affect ($p <. 001$). Participants in the colouring-in and distraction condition experienced a significantly greater improvement in negative affect than participants in the rumination condition ($p = .001$). Furthermore, the difference between the effects of the rumination versus the distraction conditions on negative affect was significantly more pronounced for people reporting higher levels of depressive

symptoms (p <. 001). No significant difference was found between colouring-in and distraction conditions on state rumination (p = .52), positive affect (p = .92) or negative affect (p = .23). There was no significant difference between conditions change in state rumination (p = .81).

Conclusion

Findings from this study suggest colouring-in is as good at alleviating negative affect as a traditional distraction intervention. Results replicate previous findings that distraction activities are associated with mood improvement compared to rumination. However, distraction activities used in this study failed to change state rumination across conditions. Further research is required to understand the mechanisms underlying colouring-in which are associated with effects on mood.

Introduction

Rumination

Rumination is a specific form of thought that is repetitive, uncontrollable, unprompted and centralized around focusing on the self (Nolen-Hoeksema, 1991). Depressive rumination is the most common form of rumination and it is known to perpetuate and intensify depressive symptoms (Morrow & Nolen-Hoeksema, 1990; Nolen-Hoeksema, 1991, 2000; Spasojevic & Alloy, 2001; Watkins, 2008) and as a trait, increase vulnerability to experiencing depression (Nolen-Hoeksema, 1991, 2000). It is therefore important to increase our understanding of the mechanisms involved in depressive rumination and the factors which heighten its severity for individuals. Previous experimental research has found depressogenic effects of rumination when compared to a distraction task for participants who were already experiencing sad mood (Nolen-Hoeksema, 2000).

Models and Theories for Rumination

The Response Styles Theory (RST) conceptualizes depressive rumination as a style of responding to stressful stimuli by passively and repetitively focusing on the self or on negative mood and emotions (Nolen-Hoeksema & Morrow, 1991). More specifically, depressive rumination involves an introspective analysis of negative thoughts and feelings as well as their causes and meanings. This behaviour leads to distress and cognitive deficits (e.g., reduced concentration) and has been found to be common in students and community adults with depression (Just & Alloy, 1997; Nolen-Hoeksema, 2000). Although RST is, arguably, the most prominent theory of rumination it does not explain some important aspects underlying its causes such as differences in how and when people ruminate. Control theories of rumination

propose a wider conceptualization of rumination as “a class of conscious thoughts that revolve around a common theme and occur in the absence of immediate environmental demands requiring the thoughts”, thus focussing on process, without specifying content (Martin & Tesser, 1996, p. 7). The “control” aspect relates to the person’s attempt to control their perception of reality so that it matches the goal they strive towards. The discrepancy between perceived and desired states of the world is believed to instigate rumination.

As well as RST and Control Theory, Stress Reactive Rumination Theory (SRRT) argues that rumination occurs when people have repetitive thoughts involving negative inferences after experiencing stressful life events (Alloy et al., 2000). This links with well-known attribution theories of depression where individuals attribute causes of negative events to personal, global and stable factors which can lead to a sense of learned helplessness (Abramson, Seligman, & Teasdale, 1978; Seligman, 1976). Alloy et al. (2000) introduced the concept of stress-reactive rumination to explain the onset and duration of depression, hypothesising that the effect of negative inferences (i.e., a negative cognitive style) on depression is more detrimental when these inferences are actively rehearsed and ruminated upon (Robinson & Alloy, 2003). Furthermore, it is argued that stress-reactive rumination is particularly maladaptive as it may worsen the effects of stressors on depressive symptoms in the context of a cognitive vulnerability-stress model (Rood, Roelofs, Bogels, & Meesters, 2012).

Distraction

Distraction is a form of emotion regulation. Emotion regulation can be defined as the extrinsic and intrinsic processes responsible for monitoring, evaluating, and modifying emotional reactions (Thompson, 1994). This includes the regulation of

one's own feelings and the regulation of other people's feelings (Niven, Totterdell, & Holman, 2009; Burman, Green, & Shanker, 2015). There are many different forms of emotion regulation when dealing with negative mood or stressful events (rumination, reappraisal, distraction, worry, etc.). Distraction can be a short-term adaptive and instrumental alternative to rumination that lifts one's mood and relieves depressive symptoms, and then if necessary, one can commence problem solving (Nolen-Hoeksema, 1991).

Distracting responses are thoughts and behaviours that divert one's attention away from one's depressed mood and its consequences and turn it to pleasant or benign thoughts and activities that are absorbing, engaging, and capable of providing positive reinforcement (Csikszentmihalyi, 2009; Nolen-Hoeksema, 1991). For example, writing a list, imagining iconic images (e.g., Empire State building), going for a run or a bike ride, seeing a movie with friends, thinking about a holiday or deadline, or concentrating on a project at work. In several experiments, Nolen-Hoeksema (1991) contrasted the effects of rumination with distraction (which involved imagining concrete details of objects) in participants who had undergone sad mood inductions, showing that distraction resulted in better mood outcomes than rumination. Effective distractions do not include inherently dangerous or self-destructive activities, such as reckless driving, heavy drinking, drug abuse, or aggressive behaviour. These may take attention away from current problems in the short-term but are harmful in the long run. However, it has been argued that when distraction is used in the right amount whilst in the difficult situation, with motivation and interest, then it can function as an 'adaptive coping strategy' (Thwaites & Freeston, 2005).

Colouring-in to Distract

One form of distraction which has become popular in recent years is colouring-in. Adult colouring books have been marketed with the suggestion that they reduce stress and increase calmness (Heussenstamm, 2014), but a review of the literature completed in advance of this study indicated a limited evidence base in this area. The adult colouring book industry make grandiose claims that they can reduce stress and increase happiness because “everyone will benefit from the stress-relieving effect that increased focus and creativity can provide” (Chapman, Merrit, & Wilde, 2015, p. 2). Some of the books also claim their designs invoke mindfulness as “the mandala designs draw your eyes towards their centres, and in doing so, focus the mind” (Heussenstamm, 2014, p. 1). The American Art Therapy Association’s stance on adult colouring books is interesting, as it argues that colouring-in is not a form of therapy, but that it does provide a safe-ground for controlled, contained use of art for self-soothing purposes whereby users can externalize their focus and shift maladaptive internal dialogue (Carolan & Betts, 2015), although these claims have barely been tested. It is currently uncertain whether colouring-in can function as an effective distraction technique for alleviating negative affect.

There are specific reasons why colouring-in may be more effective than other forms of distraction. Evidence suggests that engagement with artistic activities, either as an observer or as an initiator, can enhance one's moods, emotions, and other psychological states (Staricoff, 2003). Bell and Robbins (2007) found that student participants who spent 20 minutes freely drawing showed significantly greater reductions in negative mood and anxiety compared to those who merely viewed and sorted famous art prints for the same period of time. Similarly, Petrillo and Winner

(2005) found that participants induced into a negative mood showed significant increases in positive mood after drawing, compared to those who copied shapes. Given this preliminary evidence, it is plausible to predict that colouring-in may provide effective distraction and relief from negative mood and stress compared to traditional distraction interventions and rumination, due to the concrete external focus (Watkins, 2008) and the relaxation and flow (Csikszentmihalyi, 2009; Roston, 2016) that intricate colouring activities may provide.

The Evidence Base for Colouring-in

In one of the first studies investigating the psychological effects of colouring, Curry and Kasser (2005) induced anxiety by asking undergraduate students to write about a fearful experience. Participants were then randomly assigned to one of three conditions: colouring-in a complex circular mandala drawing, colouring-in an irregular plaid design, or drawing freely on a blank piece of paper for 20 minutes. Findings showed a greater reduction in anxiety among participants in the colouring-in conditions compared to the drawing condition, with no significant difference between the two colouring-in conditions. In a similar study with 88 psychology undergraduates, Vennet and Serice (2012) found that colouring a mandala had greater anxiety-reducing effects than free drawing. Small (2006) carried out a study with undergraduate students but with no induced stress, which found no effect of colouring or drawing on anxiety levels, perhaps because anxiety levels were already low. Muthard and Gilbertson (2016) incorporated focussed breathing alongside mandala colouring as their experimental group compared to a no-task control group. Student participants experienced a psychosocial stressor and then were randomly allocated to colouring with relaxation or the no-task control group. They found self-

reported negative affect and state anxiety were significantly lower after manipulation in the mandala-colouring experimental group. This shows how colouring-in can be used alongside other well established simple interventions, although the study's absence of an active-control group weakens its conclusions.

Flett et al. (2017) used a diary study to test the effects of colouring-in. They looked at whether a seven-day colouring intervention would lead to reductions in depressive symptoms, stress and anxiety, and increases in resilience, flourishing, and mindfulness in a population of 107 non-clinical female students. Participants either coloured a series of complex pictures or engaged in another focussed activity (logic-puzzle control group) for at least 10 minutes each day, over seven days. Participants in the colouring condition showed significant reductions in depressive symptoms and anxiety following intervention, whereas no significant reduction was found in the control group. The colouring condition showed significantly greater reductions in depressive symptoms when compared to the control condition.

The empirical evidence for colouring-in is mixed with regards to its effectiveness in reducing negative affect. Studies have not explicitly compared colouring-in with rumination to see whether it performs similar to the standard Nolen-Hoeksema distraction paradigm.

Rationale for Intervention and Aim of Current Study

This study compared colouring-in with an active distraction control and rumination condition, thus linking with the previous literature on colouring-in by including an active control that matches some of the properties of colouring-in (e.g., use of pencil, creativity), and also building on existing research by comparing colouring-in with rumination. The study examined whether the use of colouring-in

was more effective than a neutral distraction activity and rumination in reducing negative affect and state rumination that are expected to increase after participants experience a laboratory stressor that involves solving difficult anagrams**.** First, analyses explored whether both distraction conditions (colouring-in and Nolen-Hoeksema distraction) performed better than rumination, then tests were conducted to see whether colouring-in would be more effective than the Nolen-Hoeksema distraction manipulation.

A written version of Nolen-Hoeksema's (2000) rumination condition task was used (Appendix 2) so that all conditions were matched with respect to the use of motor activity relating to writing and drawing. The active control condition consisted of a modified version of the Nolen-Hoeksema (1991) distraction activity (Appendix 1). This involved imagining a series of objects (e.g., the Empire State Building), and to help match this condition for motor activity, participants drew what they were imagining. Regarding mood change following the intervention after the stressor, it was hypothesised that participants in the colouring-in condition and distraction condition combined will experience greater improvement in affect and state rumination than participants in the rumination condition. It was also hypothesised that participants in the colouring-in condition will experience greater improvement in affect and state rumination than participants in the distraction condition. This because, colouring-in has been proposed to have additional positive elements (i.e., creativity, flow, etc.) that are not present in the Nolen-Hoeksema distraction condition. Also, it was hypothesised that the benefit of colouring-in over distraction will be more pronounced for participants with more depressive symptoms. This is because distraction is known to be a short-term intervention (Nolen-Hoeksema,

1991) and thus more dysphoric participants are likely to return to rumination and distressed mood after this activity in comparison with colouring.

Hypotheses

1. After experiencing the stressor, participants in the distraction and colouring-in conditions will experience greater improvement in positive and negative affect (H1a) and state rumination (H1b) than participants in the rumination condition.
2. After experiencing the stressor, participants in the colouring-in condition will experience greater improvement in positive and negative affect (H2a) and state rumination (H2b) than participants in the distraction condition.
3. The contrasts in H1 and H2 will be more pronounced for participants higher in depressive symptoms then participants lower in depressive symptoms, as determined by PHQ-8 scores.

Method

Design

The study was a mixed (3 x 3) experimental design, with condition as the between-subjects factor (distraction, colouring-in, rumination) and a repeated-measures factor of time (baseline, post-stressor, post-manipulation). Participants were randomly assigned to one of the three experimental conditions: (1) colouring-in (n = 30); (2) drawing distraction (n = 30) and; (3) written rumination (n = 30). Participants' state rumination and negative affect (dependent variables) were measured before the stressor and before and after each of the manipulation tasks.

Sample

The sample consisted of 90 participants (80 females, 10 males) all of whom were University of Exeter students. Ages ranged from 18 to 40 years old (M = 19.0, SD = 3.6). Participants were recruited via advertisements on the University of Exeter online Research Participation System. This was a mandatory study sign-up system in which first-year psychology undergraduates were required to partake in research studies in exchange for course credits. Once signed up, participants were also entered into a prize draw to win Amazon gift vouchers.

Inclusion/Exclusion Criteria

Participants were required to be between 18 to 65 years of age and a student at the University of Exeter. Although participants who scored within the severe range of depressive symptoms (score > 18) based on the Patient Health Questionnaire-8 (PHQ-8) would have been excluded from the study for ethical reasons, no participants met any of the exclusion criteria. Ethical approval was sought and granted from the School of Psychology Committee at the University of Exeter (Appendix 3).

Measures and Materials

Depressive Symptoms. The PHQ-8 (Kroenke & Spitzer, 2002) was used as a screening measure of depressive symptoms (Appendix 4). The PHQ-8 measures depressive symptoms over the last two weeks using eight items. Validity and reliability for diagnosing depression has been well established for this measure (Kroenke, Spitzer, & Williams, 2001).

State Rumination. The State Rumination Questionnaire (SRQ; LeMoult, Arditte, D'Avanzato, & Joormann, 2013) is a 10-item measure assessing the extent

to which people are engaging in state rumination in response to a stressor. LeMoult et al. (2013) found an internal consistency of α = .70 for the SRQ. Questions from the SRQ were adapted based on both the definition of rumination from the Response Styles Theory (RST; Nolen-Hoeksema, Wisco, & Lyubomirsky, 2008). Questions were adapted to assess people's repetitive thoughts about the anagram stressor and the emotions evoked by the stressor (Appendix 5), e.g., "To what extent did you think about how sad or upset you felt?" Responses were made on a 5-point Likert-type scale ranging from 1 (*not at all*) to 5 (*a lot*). This scale was completed following the stressor and then following the manipulation to assess the extent to which participants' thoughts were focused on symptoms of distress.

State affect. The short version of the Positive and Negative Affect Schedule (PANAS; Appendix 6; Crawford & Henry, 2004; Mackinnon et al., 1999) was used to assess participants' state affect levels at three points during the study. The scale consisted of 10 adjectives that describe different feelings and emotions (five positive and five negative) with regards to how the person was feeling in the present moment. For example, a word would appear (e.g., distressed, alert, upset) and then the participant had to rate on a scale of 1 (*very slightly or not at all)* to 5 (*extremely)* how well the word represented their current feeling. The shorter version of the PANAS that was included has shown good concurrent validity (Crawford & Henry, 2004).

Stressor. An anagram stressor task previously shown to induce negative mood (MacLeod, Rutherford, Campbell, Ebsworthy, & Holder, 2002; Watkins, Moberly, & Moulds, 2008) was used to induce negative mood and state rumination (Appendix 7). This consists of 15 hard-to-solve anagrams, each five letters long, presented sequentially on a computer monitor via E-Prime Software. Participants

were asked to solve as many anagrams as possible into real words. Each anagram was presented on the screen for 10 seconds, after which an answer had to be typed within 10 seconds. Before commencing the task, participants were told (erroneously) that they were expected to answer 60-70% of anagrams correctly and that scores on this task are a reliable indicator of future academic and career success. All participants were told the percentage of anagrams they solved correctly automatically by the computer program once they had completed all anagrams, informing them that they scored well below average. Overall, no participant solved more than 30% of the anagrams correctly.

Rumination manipulation. After the stressor, participants in the rumination condition were instructed to reflect on a number of statements such as: “think about why you turned out this way”, encouraging an abstract, verbal-analytical style of thinking seen in depressive rumination (Nolen-Hoeksema, 1991; Watkins & Teasdale, 2001, 2004). Participants were asked to do this for 10 minutes. To control for physical activity, participants were instructed to write down prominent words that came to mind whilst thinking of these statements (Appendix 2).

Colouring-in manipulation. Participants in the colouring-in condition were given a diagram of a mandala, gathered from Google Images, with brief instructions asking them to colour in the diagram for 10 minutes (Appendix 8), stating that there was no right or wrong way to colour in the design and that they were not expected to finish colouring in the entire picture. Twenty assorted coloured pencils were provided for use.

Distraction manipulation. Participants in the distraction condition were instructed to complete Nolen-Hoeksema’s (1991) distraction activity (Appendix 1).

Guided by a list of physical objects and scenes (e.g., the Empire State Building), participants were instructed to imagine these scenes for 10 minutes. Participants were advised to spend approximately 20 seconds thinking about each object, although they were left to pace themselves through the list. To control for motor activity, participants were instructed to make rough, basic sketches of the objects on the paper provided.

Procedure

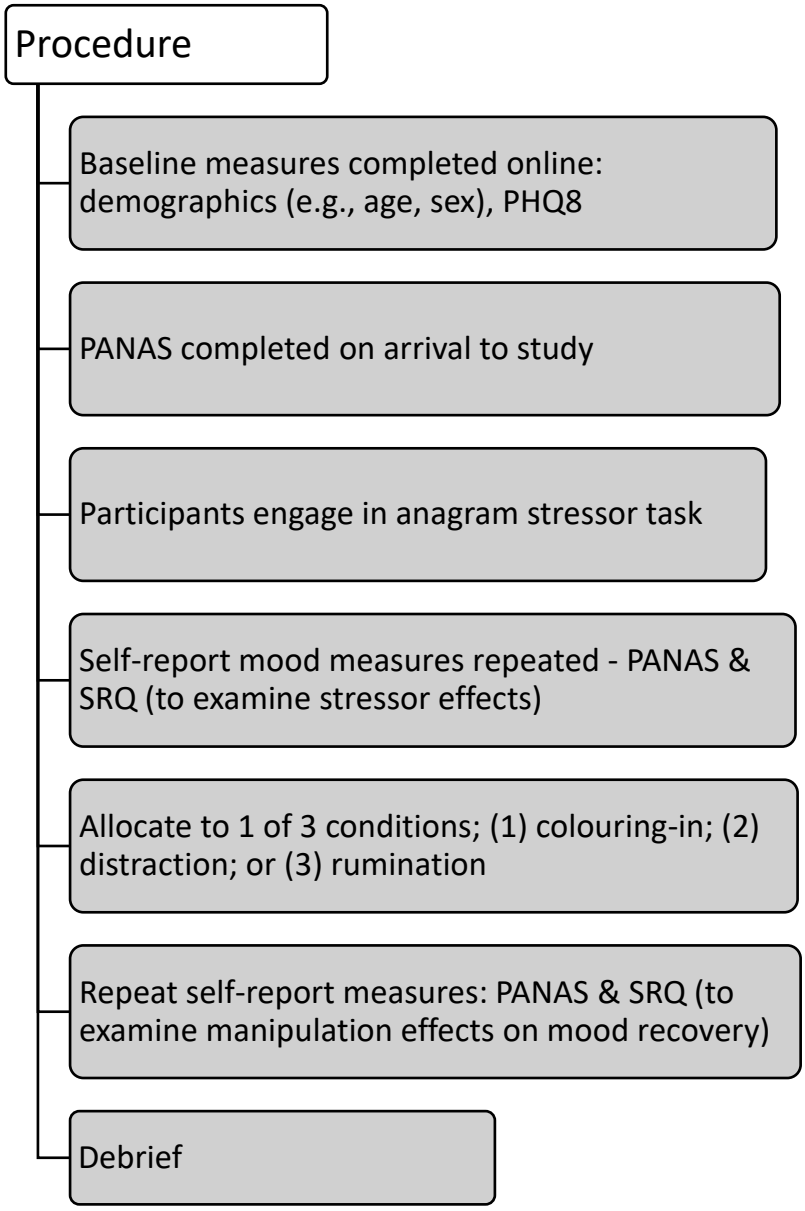
A schematic outline of the procedure can be seen in Figure 1. First, after participants signed up to the study by booking a slot via the online SONA system, they were emailed an information pack describing the rationale for the study (Appendix 9). The email included a link to complete the PHQ-8 online, as well as participant demographic questions (e.g., sex, age).

On arrival in the testing room, participants signed an informed consent form (Appendix 10) and completed the first PANAS to gather baseline mood data. They then began a computerized version of the anagram stressor task. Following the stressor, the second PANAS was completed to measure the stressor's effects on mood, and the first SRQ, to measure the extent to which they were ruminating during the stressor task. The participant was then randomly allocated to one of the three experimental conditions using an online computer-generated randomizer tool: (1) colouring-in; (2) distraction; or (3) rumination, which they engaged in for 10 minutes.

Finally, participants in all conditions repeated the state mood and rumination measures (PANAS and SRQ) to assess whether mood and rumination changed during and at the end of the manipulation period. All participants were then fully debriefed (Appendix 11) and offered contact details of local organisations that could offer mental

health advice and support (Appendix 12). Overall, the session lasted approximately 30 minutes.

Figure 1. *Procedure of study*



Data Analysis Plan

Data were analysed using IBM SPSS Statistics (Version 23) for Windows. Descriptive statistics were used to examine the presence of outliers and univariate normality. Transformations were applied to potential outliers (greater than ± 3

standard deviations from the mean) such that they were reduced to 3 SDs from the mean. Randomization tests via ANOVA were conducted to establish equivalence of conditions on dependent variables at baseline.

Mixed ANOVAs (3 [condition] x 2 [time]) with Helmert contrasts on the effect of condition were used to test whether there was statistically significant change in participants' affect from before to after the stressor and whether there was significant change in affect and state rumination from before to after the post-distraction activity. Each condition was coded such that the Helmert contrasts resulted in the test of the hypotheses: (1) rumination vs (distraction + colouring), (2) distraction vs colouring. Interactions between time and these contrasts of condition provided the crucial hypothesis tests. The alpha level was set at .05 for all analyses. Regression analysis was also used to test whether PHQ-8 moderates the effect of condition (contrast-coded as above) on negative affect and state rumination.

Power calculation

Power calculations using G*Power software for the rumination vs. both distraction conditions were based on the effect size reported in Lyubomirsky and Nolen-Hoeksema's (1995) study of rumination versus distraction on mood in dysphoric and non-dysphoric participants, which revealed a medium-sized effect for the interaction between condition and time. Thus, a medium effect size is expected for both mood and rumination outcomes. Further power calculations for the colouring-in vs. distraction condition were based on the effect size reported in a study by Flett et al. (2017), which compared colouring-in to an active control (e.g., daily puzzles), which revealed a small to medium effect size for the interaction between condition and time. It is also assumed the magnitude of effect on state rumination will be similar to these findings.

Manipulation check: to detect a medium effect size (effect of anagrams on affect and state rumination) for the effect of time in repeated-measures ANOVA, f = .25, power = .80, correlation between repeated measures = .5, a total sample size of 34 was required.

Hypothesis 1: to detect a medium effect size for the time x condition interaction on affect and state rumination in mixed ANOVA, f = .25, power = .80, correlation between repeated measures = .5, a total sample size of 51 is required (with at least 26 in the rumination condition).

Hypothesis 2: to detect a small-medium effect size for the time x condition interaction on affect and state rumination in mixed ANOVA, f = .19, power = .80, correlation between repeated measures = .5, required total sample size of 87 (at least 44 per condition).

Results

Consultation Findings

Prior to full empirical testing, the current study was piloted with three participants, and alongside this consultation was made with members of the Lived Experience Group (LEG) within the Mood Disorders Centre at the University of Exeter. This was to discuss each of the study conditions and their personal experiences of colouring-in and use of distraction. One member of the LEG described her experience of colouring-in as *"OK and quite relaxing, however I'm not sure if I could do it in a moment of anxiety but I think like mindfulness it's a technique you could employ when not anxious that you would learn to do when anxious".* I was also informed that another member of the LEG attended a meeting stressed and therefore spent the meeting colouring-in, which she found it helpful in allowing her to

relax and engage in the meeting. Despite the limitations of anecdotal evidence, it shows that colouring-in can be helpful in both stressful and non-stressful situations.

Data screening

Data were screened for entry errors via checking the minimum and maximum scores for each participant on each questionnaire. There was no missing data. Outliers were then checked by converting the data into *z*-scores and analysing boxplot graphs. Univariate outliers were found for six participants who had *z*-scores greater than ±3, four in the distraction condition (negative affect at each time point, and SRQ post-manipulation) and two in the colouring condition (both positive affect post-stressor), therefore these outlying scores were converted such that they were 3 SD from the mean. Difference scores from time point 1 to 2 (baseline to post-anagram) and time point 2 to 3 (post-anagram to post-manipulation) were calculated to test for normal distribution in change scores for the SRQ and PANAS across the three conditions. Difference scores were found to be normally distributed in the three conditions, such that parametric tests were appropriate.

Randomization Check

Group means for the PANAS positive and negative mood assessments, SRQ and PHQ-8 scores are presented in Table 1. A between-groups ANOVA found no significant differences in baseline PHQ-8 scores across the rumination, distraction and colouring-in conditions, $F(2, 89) = 0.47$, $p = .62$, partial $\eta^2 = .01$. Similarly, no significant differences were found across the three conditions for baseline PANAS positive affect, $F(2, 89) = 1.12$, $p = .33$, partial $\eta^2 = .25$, baseline PANAS negative affect, $F(2, 89) = 2.92$, $p = .059$, partial $\eta^2 = .06$, or post-stressor SRQ scores across

conditions, $F(2, 89) = 2.32$, $p = .10$, partial $\eta^2 = .05$. These findings suggest that the randomization of participants into conditions was successful.

Manipulation Check

To assess the success of the mood induction, a 3 × 2 mixed ANOVA was performed on state positive affect with condition (rumination, distraction, colouring-in) as the between-groups factor and time of measurement (baseline vs. post-stressor) as the repeated-measures factor. A significant within-subjects effect of time was observed, indicating a reduction in positive affect from before to after the stressor, $F(1, 87) = 154.86$, $p < .001$, partial $\eta^2 = .64$. No significant main effect of condition was found, $F(2, 87) = 0.85$, $p = .43$, partial $\eta^2 = .02$. The interaction between condition and time was not significant, $F(2, 87) = 0.29$, $p = .75$, partial $\eta^2 = .007$.

A similar 3 x 2 mixed ANOVA on negative affect scores also found a significant within subjects effect of time, $F(1, 87) = 44.12$, $p < .001$, partial $\eta^2 = .34$. A significant main effect of condition was found, $F(2, 87) = 3.12$, $p = .049$, partial $\eta^2 = .067$. As apparent from descriptive statistics in Table 1, the distraction condition reported lower negative affect than other conditions pre and post stressor. The interaction between condition and time was not significant, $F(2, 87) = 2.38$, $p = .099$, partial $\eta^2 = .05$. All participants demonstrated significantly increased negative affect and decreased positive affect from before to after the anagram stressor, and there were no significant differences in condition regarding the effect of the stressor.

Table 1.

Means and standard deviations (in parentheses) for study variables at all phases in the rumination, distraction, and colouring-in conditions

	Rumination (*n* = 30)	Distraction (*n* = 30)	Colouring-in (*n* = 30)
Positive affect			
PANAS 1 (baseline)	19.93 (4.32)	20.53 (4.69)	21.83 (5.90)
PANAS 2 (post-stressor)	14.43 (4.56)	14.26 (5.79)	15.63 (6.55)
PANAS 3 (post-manipulation)	17.40 (5.19)	19.10 (5.70)	19.86 (7.01)
Negative affect			
PANAS 1 (baseline)	11.66 (3.68)	9.56 (2.40)	10.93 (3.94)
PANAS 2 (post-stressor)	13.73 (4.77)	11.96 (4.49)	15.15 (5.89)
PANAS 3 (post-manipulation)	11.83 (4.82)	8.96 (2.97)	9.63 (3.47)
State rumination			
SRQ 1 (post-stressor)	17.56 (4.25)	16.66 (4.76)	18.96 (3.35)
SRQ 2 (post-manipulation)	9.30 (3.46)	8.13 (3.60)	11.26 (4.94)
PHQ-8 (baseline)	4.60 (2.74)	5.26 (3.51)	4.56 (3.11)

PANAS = Positive and Negative Affect Schedule, SRQ = State Rumination Questionnaire, PHQ-8 = Patient Health Questionnaire 8.

Positive Affect

Due to the imbalance in positive affect and negative affect between conditions at post-stressor, residualised change scores were calculated by regressing post-manipulation affect on post-stressor affect for positive and negative affect separately and saving the residuals. These residualised change scores were then analysed

using *t*-tests for which the conditions being compared reflected the planned Helmert contrasts.

An independent samples *t*-test using residualised change scores for post-manipulation positive affect in the rumination condition (*M* = -1.17*, SD* = 4.43) to the two distraction conditions combined (colouring and object imagination) (*M* = 0.58, *SD* = 4.61) found no significant difference, *t* = -1.72, *p* = .09, Cohen *d* = -0.38*.* A second *t*-test compared participant's residualised change score for post-manipulation positive affect in the distraction condition (*M* = 0.65, *SD* = 4.45) compared to the colouring-in condition (*M* = 0.52, *SD* = 4.84) and found no significant difference, *t* = 0.10, *p* = .92, Cohen *d* = 0.02. This shows that there was no differential pattern with regards to positive affect recovery between rumination and the two distraction conditions, which fails to support the positive affect aspect of Hypothesis 1a. Also, with no significant differences between the colouring-in and object-imagination distraction condition, thus Hypothesis 2a was not supported.

To explore whether depressive symptoms interacted with the manipulation conditions, a hierarchical multiple regression was conducted, in which post-manipulation positive affect (PA3) was the outcome variable and the predictor variables in the first step were pre-manipulation positive affect (PA2), condition (coded as 0 = rumination, 1 = colouring and distraction), mean-centred PHQ-8, and the interaction between condition and mean-centred PHQ-8 was added in the second step. This followed the logic of the previous *t*-tests (i.e., first comparing rumination vs. colouring and distraction, then comparing colouring vs. distraction). In step one, the proportion of variance in the sample as determined by the unadjusted R^2 was 46% (adjusted R^2 = .44). The model fit was significant, $F(3, 86) = 24.62$, $p < .001$. Greater pre-manipulation positive affect was associated with greater post-

manipulation positive affect scores (β = .69, t = 8.00, p < .001). Participants' PHQ-8 scores were not associated with post-manipulation positive affect scores, controlling for pre-manipulation positive affect, (β = -.29, t = -1.95, p = .054). The predictor variable of condition did not significantly contribute to the model after controlling for depressive symptoms (β = 1.85, t = -3.99, p = .69). In step two, the addition of the interaction explained no significant additional variance in positive affect post-manipulation, R^2 change = .001, F(1, 85) = 0.21, p = .65. This shows that the effect of rumination vs. colouring-in and distraction combined on positive affect recovery did not depend on depressive symptoms.

A similar multiple regression was conducted to see whether depressive symptoms are associated with the advantage of colouring-in versus distraction on positive affect recovery. Post-manipulation positive affect (PA3) was the outcome variable and the predictor variables in the first step were pre-manipulation positive affect (PA2), condition (coded as 0 = distraction, 1 = colouring-in), mean centred PHQ-8 scores, and the interaction between condition and mean centred PHQ-8 was added in the second step. In step one, the proportion of variance in the sample as determined by the unadjusted R^2 was 49% (adjusted R^2 = .46). The model significantly predicted the outcome, F(3, 56) = 17.98, p < .001. Greater pre-manipulation positive affect was associated with greater post-manipulation positive affect scores (β = .71, t = 7.04, p < .001). PHQ-8 scores were not associated with post-manipulation positive affect, when controlling for post-stressor positive affect (β = -.25, t = -1.40, p = .16), nor did condition significantly contribute to the model after controlling for depressive symptoms (β = -.32, t = -0.26, p = .79). In step two, the addition of the interaction explained no significant additional variance in positive affect post-manipulation, R^2 change = .023, F(1, 55) = 2.62, p = .11. This shows that

the effect of colouring-in versus distraction condition on positive affect recovery did not depend on level of depressive symptoms.

Negative Affect

An independent samples *t*-test using residualised change scores for post-manipulation negative affect in the rumination condition (*M* = 1.64*, SD* = 4.19) to the colouring and object imagination distraction conditions combined (*M* = -0.82, *SD* = 2.33) found a significant difference, *t* = 3.57, *p* = .001, Cohen *d* = 0.72. A second *t*-test compared participant's residualised change score for post-manipulation negative affect in the distraction condition (*M* = -0.45, *SD* = 2.05) compared to the colouring-in condition (*M* = -1.18, *SD* = 2.57) and found no significant difference, *t* = 1.04, *p* = .23, Cohen *d* = 0.31. This shows that there was a differential pattern with regards to negative affect recovery between rumination and the two distraction conditions, which supports Hypothesis 1a. However, no difference between the colouring-in and object-imagination distraction condition was found, thus Hypothesis H2a was not supported.

To explore whether depressive symptoms interacted with the manipulation conditions to predict change in negative affect, a hierarchical multiple regression was conducted, in which post-manipulation negative affect (NA3) was the outcome variable and the predictor variables in the first step were pre-manipulation negative affect (NA2), condition (coded as 0 = rumination, 1 = colouring and distraction), mean-centred PHQ-8, and the interaction between condition and mean-centred PHQ-8 was added in the second step. This set-up followed the logic of the previous *t*-tests (i.e., first comparing rumination vs. colouring and distraction, then comparing colouring vs. distraction). In step one, the proportion of variance in the sample as

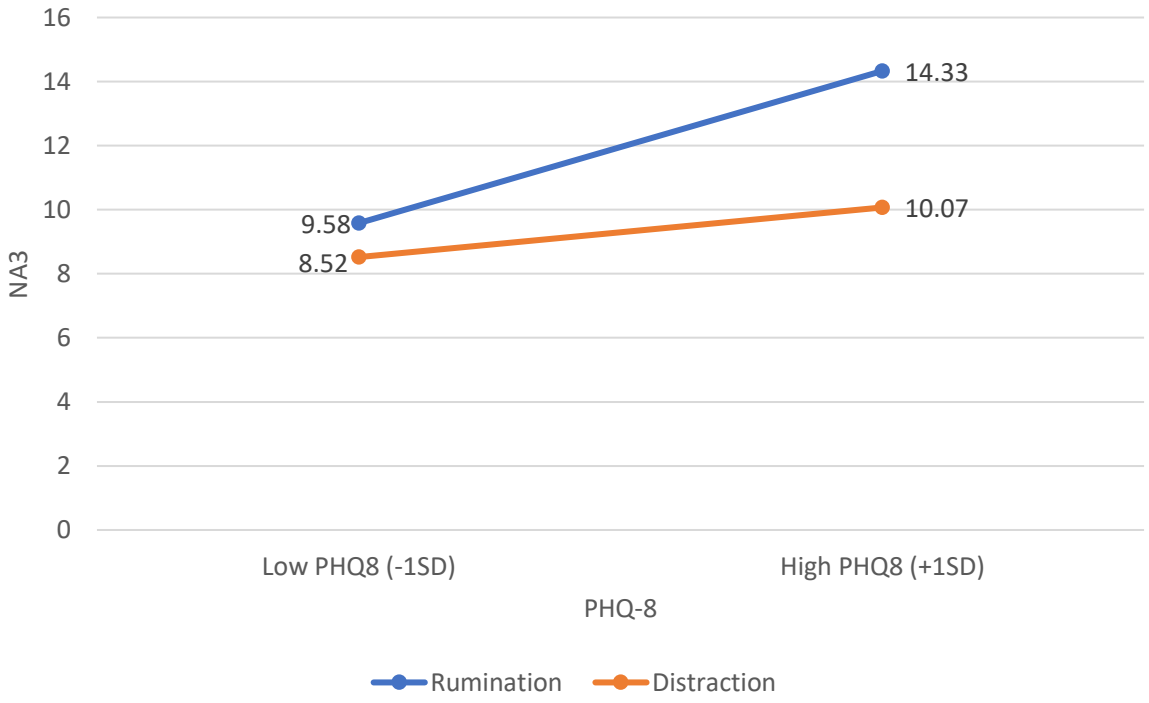
determined by the unadjusted R^2 was 49% (adjusted R^2 = .47). The model significantly predicted the outcome, $F(3, 86) = 27.59$, $p < .001$. Greater pre-manipulation negative affect was associated with greater post-manipulation negative affect scores (β = .47, $t = 5.75$, $p < .001$). Higher PHQ-8 scores were also associated with greater post-manipulation negative affect, controlling for post-stressor negative affect (β = .29, $t = 6.61$, $p = .001$). The predictor variable of condition significantly contributed to the model after controlling for depressive symptoms (β = -0.30, $t = -3.99$, $p < .001$). This effect reflects the lower level of negative affect post-manipulation in the distraction and colouring conditions compared to the rumination condition, whilst controlling for depressive symptoms.

In step two, the addition of the interaction explained significant additional variance in negative affect post-manipulation, R^2 change = .03, $F(1, 85) = 5.39$, $p = .02$. The interaction between condition and PHQ-8 was significantly associated with post-manipulation negative affect (β = -0.51, $t = -2.32$, $p = .02$). This shows that the effect of the rumination condition versus the two distraction conditions depended on level of depressive symptoms (e.g., higher depressive symptoms at baseline were associated with the manipulation being more effective).

Simple slopes analysis was used to break down the significant interaction and assess whether depressive symptoms (baseline PHQ-8 scores) was significantly associated with mood recovery levels within the rumination condition and within the two distraction conditions (e.g., colouring-in and drawing objects). The positive association between PHQ-8 score and NA3 (when controlling for NA2) was significant in both the rumination condition (β = .60, $t = 3.92$, $p < .001$) and in the distraction conditions (β = .19, $t = 2.13$, $p = .035$). As seen clearly in Figure 2, depressive symptoms more strongly predicted negative affect after the manipulation

in the rumination condition than in the distraction conditions. Otherwise stated, the difference between the effects of the rumination versus the distraction condition on negative affect was most pronounced for people reporting higher levels of depressive symptoms.

Figure 2.
Line graph of the simple slopes results for rumination and distraction conditions

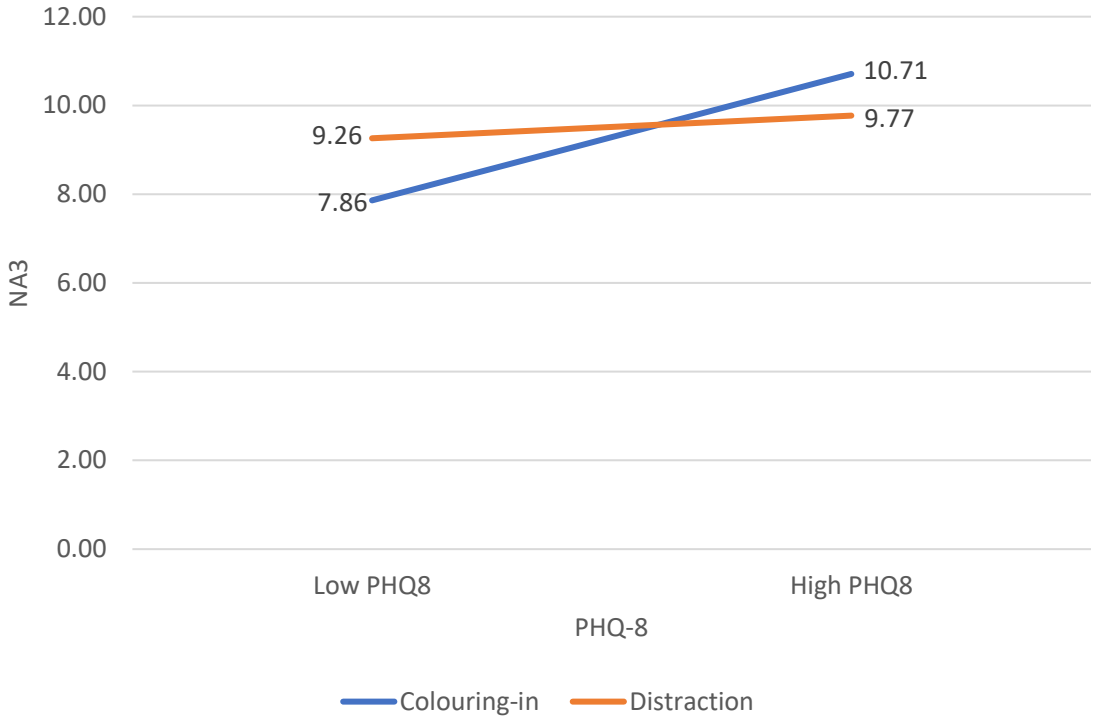


A similar multiple regression was conducted to see whether depressive symptoms are associated with the size of the difference between colouring-in and distraction on negative affect recovery. In step one, the proportion of variance in the sample as determined by the unadjusted R^2 was .53 (adjusted R^2 = .51). The model significantly predicted the outcome, $F(3, 56) = 21.75$, $p < .001$. Greater pre-manipulation negative affect was associated with greater post-manipulation negative affect scores (β = .37, $t = 6.10$, $p < .001$). Higher PHQ-8 scores were also associated

with greater post-manipulation negative affect, controlling for post-stressor negative affect (β = .23, *t* = 2.43, *p* = .02). Condition did not significantly contribute to the model after controlling for depressive symptoms (β = -.32, *t* = -0.53, *p* = .60). In step two, the addition of the interaction explained significant additional variance in negative affect post-manipulation, R^2 change = .04, *F*(1, 55) = 4.55, *p* = .04. The interaction between condition and PHQ-8 was significantly associated with post-manipulation negative affect (β = .38, *t* = 2.13, *p* = .04). Thus, the effect of colouring vs distraction condition on negative affect recovery depended on depressive symptoms.

A second simple slopes analysis was used to assess whether depressive symptoms (PHQ-8 scores) was significantly associated with mood recovery levels within the colouring-in and distraction condition. The association between PHQ-8 score and NA3 (when controlling for NA2) was significant in the colouring-in condition (β = .47, *t* = 3.25, *p* = .002), but not in the distraction condition (β = .09, *t* = 0.70, *p* = .49). As seen in Figure 3, higher depressive symptoms was associated with slower negative affect recovery in the colouring-in condition but no significant association was found in the distraction condition, thus not supporting Hypothesis 3.

Figure 3.
Line graph of the simple slopes results for rumination and distraction conditions



State Rumination

To assess whether state rumination on the SRQ changed differentially across conditions, a 3 x 2 mixed ANOVA was conducted with condition as the between-groups factor and a repeated-measures factor of time (pre-manipulation vs. post-manipulation). A significant within-subjects effect of time was observed indicating a significant decrease in state rumination, $F(1, 87) = 234.21$, $p < .001$, partial $\eta^2 = .73$. There was a significant main effect of condition, $F(2, 87) = 5.36$, $p = .006$, partial $\eta^2 = .11$. The interaction between condition and time was non-significant, $F(2, 87) = 0.21$, $p = .81$, partial $\eta^2 = .005$, showing no significant difference in SRQ reduction between conditions. Nevertheless, hypotheses based on particular conditions were tested using planned comparisons. A planned Helmert contrast comparing the reduction in

state rumination scores from pre-manipulation to post-manipulation in the rumination condition (M = -8.26, SD = 5.20) with the two distraction conditions (colouring and object imagination; M = -8.11, SD = 5.36) showed no significant difference, p = .89, partial η^2 = .002. A second planned contrast also found no significant difference for improvement in state rumination scores from pre-manipulation to post-manipulation when comparing the distraction condition (M = -8.53, SD = 5.53) with the colouring-in condition (M = -7.70, SD = 5.20), p = .52, partial η^2 = .004. This shows that there was no significant difference in reduction in SRQ scores between the distraction conditions and rumination, and between colouring-in and distraction, and therefore Hypotheses 1b and 2b was not supported.

Discussion

The current study explored whether colouring-in was more effective than a neutral distraction activity and rumination in improving affect and state rumination, after participants experienced an induced stressor. The mood induction was successful in significantly increasing levels of negative affect and reducing levels of positive affect, with no significant difference between conditions. This is consistent with previous research which also found the anagram stressor task to induce negative mood (MacLeod et al., 2002; Watkins et al., 2008).

Positive and Negative Affect Findings

The first part of Hypothesis 1 (H1a) predicted that after experiencing an induced stressor, participants in the distraction and colouring-in conditions would experience greater improvements in positive and negative affect, than participants in the rumination condition. This hypothesis was supported by findings for negative affect but not positive affect. As expected for negative affect, participants in the

colouring-in and distraction conditions experienced a significantly greater negative affect recovery than participants in the rumination condition. However, no significant differences were found with regards to improvement in positive affect following manipulation across conditions, thus not supporting the positive affect prediction in H1a. Findings for negative affect were consistent with Nolen-Hoeksema's (1991) research with regards to distraction outperforming rumination.

The second hypothesis (H2a) was specific to colouring-in, in that it was predicted that colouring-in would improve positive affect and negative affect significantly more than distraction (imagining and drawing objects). The reasoning behind this prediction was governed by the emerging empirical evidence for colouring-in in reducing negative affect (Flett et al., 2017), and the potential increased functions of colouring-in, such as flow and creativity (Csikszentmihalyi, 2009; Roston, 2016), that may be underlying its benefits as opposed to traditional short-term distraction interventions used in Nolen-Hoeksema's studies (1991). The findings failed to support this hypothesis in relation to positive and negative affect (H2a). There was no significant difference in affect recovery in the colouring-in condition compared with the distraction condition. Research has shown that colouring-in a complex design reduces negative affect significantly more compared to colouring-in more basic design (Curry & Kasser, 2005). Thus, with participants only colouring for 10 minutes in this study, with a design not seen as overly complex, it may explain why participants did not find it more beneficial than distraction, considering the potential underlying theoretical predictions of colouring-in (e.g., flow, creativity) are unlikely to be achieved in such a short amount of time.

The first part of hypothesis 3 predicted that the contrasts in hypothesis 1, with regards to change in negative affect pre to post manipulation, will be more

pronounced for participants higher in baseline depressive symptoms than participants lower in depressive symptoms. Levels of depressive symptoms more strongly predicted post-manipulation negative affect scores in the rumination condition than in the distraction conditions combined. This suggests that the difference between the effects of the rumination versus the distraction conditions on negative affect was most pronounced for people reporting higher levels of depressive symptoms, thus supporting hypothesis 3. This is the expected pattern that Nolen-Hoeksema (1991) would predict. It is important to acknowledge that these findings resulted from the use of a modified version of the Nolen-Hoeksema (1991) distraction and rumination manipulations that involved greater activity to allow for matching motor activity with colouring, but still showed an advantage of distraction.

The second part of hypothesis 3 predicted that the contrasts in hypothesis 2, with regards to change in negative affect pre to post manipulation when comparing the colouring in against the distraction condition, will be more pronounced for participants higher in baseline depressive symptoms than participants lower in depressive symptoms. Unexpectedly, colouring-in was less effective at reducing negative affect for people with higher depressive symptoms compared to lower depressive symptoms, thus not supporting the predictions of hypothesis 3 in relation to the contrasts between colouring and distraction. There was no significant difference for the distraction object-imagination condition at reducing negative affect for people with high or low depressive symptoms. The overall mean depression score for participants in this study was below the mild range for depression on the PHQ-8, thus inferences from these specific findings only generalise as far as a non-clinical population.

State Rumination Findings

Hypothesis 1b predicted that after experiencing an induced stressor, participants in the distraction and colouring-in conditions would experience greater improvements in state rumination than participants in the rumination condition, although the findings failed to support this hypothesis. All conditions showed a significant reduction in SRQ scores following manipulation, but with no significant difference between levels of reduction between the three conditions. This raises the question as to whether there was an effect of the rumination induction at all. This is inconsistent with the literature around rumination conditions showing that state rumination measures respond to rumination versus distraction manipulations (Nolen-Hoeksema & Morrow, 1991).

Findings from the study did not support the second part of Hypothesis 2 (H2b), which predicted that colouring-in would improve state rumination significantly more than the object-imagination distraction condition. This could be due to the use of a non-clinical student population. It may be that state rumination was not high enough for there to be a significant difference between conditions. Also, the rumination manipulation may not have actually targeted state rumination effectively, perhaps as it was not relevant to participants' goals at the time.

Although the use of colouring-in as an intervention technique is still in its infancy with regards to it being an evidence-based intervention, findings from this study contribute to the emerging literature with regards to colouring-in having the potential to alleviate mood following induced stress in a laboratory setting (Curry & Kasser, 2005; Vennet & Serice, 2012). These findings could be used to help promote colouring-in within student populations and generally as a public health

recommendation as a simple, easy to use self-care strategy to help alleviate negative mood. At a time of austerity within the National Health Service (Baumol, Ferranti, Malach, Mendez, Tabish, & Gomory, 2012), an intervention such as colouring-in would be useful considering its cost-effectiveness and accessibility.

This study was the first to include colouring-in as an alternative distraction task, and the first to use a distraction control condition which explicitly attempted to match some of the processes of colouring-in (e.g., motor action, use of pen, creativity, and colour). Matching the distraction control on such processes may have been one of the reasons as to why there was no significant differences found between the colouring-in and distraction condition. Therefore, it would be useful to carry out further studies with a matched active-control away from the laboratory, to see whether the benefits of colouring-in are consistent in a more ecologically valid setting to help inform their usage in clinical practice as an adjunct to therapeutic models and self-help programs.

Gilbert's (2009) compassionate mind theory could help explain the beneficial findings of colouring-in in terms of depressive and low positive affect. Gilbert (2009) proposes that there are three main emotion regulation systems: drive, soothing, and threat. It could be hypothesised that colouring-in activates the 'soothing system', which helps to regulate feelings of safety and calmness. Furthermore, if colouring-in leads to the person being more mindful, they are more likely to be focussed on the present moment rather than ruminating about the past or worrying about the future (Teasdale et al., 2006), thus helping to alleviate negative feelings.

Limitations

There is a heavy reliance on self-report measures within the colouring-in, distraction and rumination research, which could have implications for the quality of the literature in this area (Adams, Soumerai, Lomas, & Ross-Degnan, 1999). For example, social desirability and demand characteristics have been known to influence participant responses to self-report questions. (Barker, Horton, Kent, & Tennant, 2013). Also, expectancy effects may have occurred in that participants expected their mood to improve following colouring-in, which then influenced how they completed outcome measures. This could have been overcome by asking participants how much they expected colouring-in to affect their mood, to assess the extent to which expectancy effects influenced results. In addition, participants may have not wanted to be perceived as stressed following the stressor, and therefore altered their responses accordingly. It could be beneficial in future research to consider measuring psychophysiological responses such as cortisol (reflecting the activation of the threat system) and heart-rate variability (soothing system; Rockcliff, Gilbert, Mcewan, Stafford, & Glover, 2008), alongside self-report measures such as the SRQ and PANAS, to help increase the validity and reliability of studies.

The current sample consisted of primarily female undergraduate students, thus making generalisations to the wider population more difficult. It is unclear whether acceptability and effects of colouring differ between genders. A qualitative component should be used to explore this in future research. In relation to this, the sample was non-clinical with participants not seeking treatment for depression nor displaying self-reported depressive symptoms. The decision to use a non-clinical undergraduate population in this study was partly due to the requirement to provide

sufficient statistical power with limited time to recruit participants. Also, the study aimed to obtain proof of colouring-in being beneficial in a non-clinical sample first, before future studies could use clinical samples. Thus, it is possible that the lack of significant difference between conditions (e.g., on state rumination and negative affect) was due to the non-clinical nature of the sample. Replication of this study using a clinical sample would be useful to determine whether the findings are significant in this population, which would lead to further clinical implications being made.

A further potential limitation lies with how the Nolen-Hoeksema (1991) distraction and rumination paradigms were modified. There is a risk that by changing these it weakened their effects and reduced power to detect differences. For example, incorporating the use of writing down ruminative thoughts may have prevented the progressive build-up of internalised ruminative thoughts. Also, perhaps participants found the rumination condition dispiriting but did not ruminate because the anagram task did not map on to a personal goal (Martin & Tesser, 1996), or perhaps participants did not fully believe the false feedback from the anagram stressor task enough for them to ruminate about it. Although modifying the paradigms may have led to a more conservative comparison and thus reducing the possibility of detecting differences, the modifications did lead to a tighter experimental comparison with respect to colouring-in because all conditions involved pen movement.

Future Directions

The focus of future research should prioritise testing of colouring-in against a do-nothing control. This will explore whether colouring-in is better than doing nothing,

which this study and previous studies cannot conclude due to the lack of do-nothing control conditions. The investigation of possible moderators and mediators of its effects could then be conducted. Mediators would allow a deeper understanding of the mechanisms involved in colouring-in. For example, state mindfulness could be measured to see whether this predicts effectiveness of colouring-in by participants answering state mindfulness related questions whilst in the process of colouring. Moderators which could be explored are aspects of the colouring itself, for example; design in use (e.g., complex mandala, simple shapes), use of specific colours, and length of time colouring. A complex mandala design was utilised in this study because it has been used in previous research and found positive results in lowering negative affect (Curry & Kaser, 2005; Vennet and Serice, 2012). It is unclear whether the design that is being coloured-in has any impact on outcome. Heussenstamm (2014) claimed that colouring-in mandalas can bring upon a state of peace and mindfulness due to its links to Buddhism's meditative principles. However, such religious links inherent in the manipulation task could have influenced how useful the task was for participants, depending on their religious and spiritual beliefs. Therefore, research which helps to understand these processes could identify ways in which colouring-in as an intervention could be tailored to personal beliefs to increase its ability at reducing depressive symptoms. Examining moderators could also focus on examining whether the effects of colouring-in at reducing negative affect are stronger or less strong for specific populations, which could help indicate suitability for clinical populations. It is difficult to state clinical implications given the sample used in this study, but colouring-in could provide a better alternative than distraction for clinical samples as it is arguably a simpler task with fewer instructions.

Furthermore, research around the use of colouring-in as a therapeutic intervention may be more useful when conducted via a qualitative stance, to understand more about the phenomenological experience of the person doing the colouring-in in order to learn about the components of colouring-in which lead to positive outcomes. Overall, it is hoped that in adding to the body of growing literature in this area, this study could help guide future research towards more empirical examination of the underlying mechanisms of colouring-in and their impact on mood processes.

References

Abbott, K. A., Shanahan, M. J., & Neufeld, R. W. (2013). Artistic tasks outperform non-artistic tasks for stress reduction. *Art Therapy, 30*, 71-78. doi:10.1080/07421656.2013.787214

Abramson, L. Y., Seligman, M. E., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87*, 49-74. doi:10.1037/0021-843X.87.1.49

Adams, A. S., Soumerai, S. B., Lomas, J., & Ross-Degnan, D. (1999). Evidence of self-report bias in assessing adherence to guidelines. *Journal of International Society for Quality in Healthcare, 11,* 187-192.

Alloy, L. B., Abramson, L. Y., Hogan, M. E., Whitehouse, W. G., Rose, D. T., Robinson, M. S., & Lapkin, J. B. (2000). The Temple-Wisconsin cognitive vulnerability to depression project: Lifetime history of axis I psychopathology in individuals at high and low cognitive risk for depression. *Journal of Abnormal Psychology, 109*, 403-418. doi:10.1037//0021-843x.109.3.403

Barker, S. M., Horton, R., Kent, M., & Tennant, A. (2013). Development of a self-report scale of spasticity. *Topics in Stroke Rehabilitation, 20*, 485-492. doi: 10.1310/tsr2006-485

Baumol, W. J., Ferranti, D., Malach, M., Mendez, A., Tabish, H., & Gomory, W. L. (2012). *The cost disease: Why computers get cheaper and health care doesn't.* New York: Yale University Press.

Bell, C. E., & Robbins, S. J. (2007). Effect of art production on negative mood: a randomized, controlled trial. *Art Therapy, 24*, 71-75. doi:10.1080/07421656.2007.10129589

Boothby, D. M., & Robbins, S. J. (2011). The effects of music listening and art production on negative mood: A randomized, controlled trial. *The Arts in Psychotherapy, 38*, 204-208. doi:10.1016/j.aip.2011.06.002

Burman, J. T., Green, C. D., & Shanker, S. (2015). On the meanings of self-regulation: digital humanities in service of conceptual clarity. *Child Development, 86*, 1507-1521. doi:10.1111/cdev.12395

Carolan, R., & Betts, D. (2015). The adult coloring book phenomenon. Retrieved from http://3blmedia.com/News/.

Crawford, J. R., & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology, 43*, 245-265. doi:10.1348/0144665031752934

Csikszentmihalyi, M. (2009). *Flow: The psychology of optimal experience*. New York: Harper Row.

Curry, N. A., & Kasser, T. (2005). Can coloring mandalas reduce anxiety? *Art Therapy, 22*, 81-85. doi:10.1080/07421656.2005.10129441

Davis, C. G., & Nolen-Hoeksema, S. (2009). Making sense of loss, perceiving benefits, and posttraumatic growth. In C. R. Snyder & S. J. Lopez (Eds.), Oxford *handbook of positive psychology* (pp. 641-649). New York, NY: Oxford University Press.

Flett, J. A., Lie, C., Riordan, B. C., Thompson, L. M., Conner, T. S., & Hayne, H. (2017). Sharpen your pencils: Preliminary evidence that adult coloring reduces depressive symptoms and anxiety. *Creativity Research Journal, 29*, 409-416. doi:10.1080/10400419.2017.1376505

Forkosh, J., & Drake, J. E. (2017). Coloring versus drawing: Effects of cognitive demand on mood repair, flow, and enjoyment. *Art Therapy, 34*, 75-82. doi:10.1080/07421656.2017.1327272

Gilbert, P. (2009). *The compassionate mind.* London: Constable & Robinson Ltd.

Heussenstamm, P. (2014). *Colour yourself calm: A mindfulness colouring book*. London: Quadrille Publishing.

Just, N., & Alloy, L. B. (1997). The response styles theory of depression: Tests and an extension of the theory. *Journal of Abnormal Psychology, 106*, 221-229. doi:10.1037//0021-843x.106.2.221

Kleiman, E. M., & Riskind, J. H. (2013). Negative cognitive style and looming cognitive style synergistically predict stress generation. *Anxiety, Stress, and Coping, 27*, 347-359. doi:10.1080/10615806.2013.853048

Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: A new depression diagnostic and severity measure. *Psychiatric Annals, 32*, 509-515. doi: 10.3928/0048-5713-20020901-06

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9. *Journal of General Internal Medicine, 16*, 606-613. doi:10.1046/j.1525-1497.2001.016009606.x

Lemoult, J., Arditte, K., D'avanzato, C., & Joormann, J. (2013). State rumination: Associations with emotional stress reactivity and attention biases. *Journal of Experimental Psychopathology, 4,* 471-484. doi:10.5127/jep.029112

Lyubomirsky, S., & Nolen-Hoeksema, S. (1995). Effects of self-focused rumination on negative thinking and interpersonal problem solving. *Journal of Personality and Social Psychology, 69*, 176-190. doi:10.1037//0022-3514.69.1.176

Mackinnon, A., Jorm, A. F., Christensen, H., Korten, A. E., Jacomb, P. A., & Rodgers, B. (1999). A short form of the Positive and Negative Affect Schedule: Evaluation of

factorial validity and invariance across demographic variables in a community sample. *Personality and Individual Differences, 27*, 405-416. doi:10.1016/s0191-8869(98)00251-7

Macleod, C., Rutherford, E., Campbell, L., Ebsworthy, G., & Holker, L. (2002). Selective attention and emotional vulnerability: Assessing the causal basis of their association through the experimental manipulation of attentional bias. *Journal of Abnormal Psychology, 111*, 107-123. doi:10.1037//0021-843x.111.1.107

Martin, L. L., & Tesser, A. (1996). Some ruminative thoughts. In R. S. Wyer, Jr. (Ed.), Advances in social cognition, Vol. 9. Ruminative thoughts (pp. 1-47). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Merritt, R., Wilde, C., Chapman, L., & Schrey, S. (2015). *Color therapy: An anti-stress coloring book*. Philadelphia: Running Press.

Morrow, J., & Nolen-Hoeksema, S. (1990). Effects of responses to depression on the remediation of depressive affect. *Journal of Personality and Social Psychology, 58*, 519-527. doi:10.1037//0022-3514.58.3.519

Morrow, J., & Nolen-Hoeksema, S. (1991). A prospective study of depression and posttraumatic stress symptoms after a natural disaster: The 1989 Loma Prieta earthquake. *Journal of Personality and Social Psychology, 61,* 115-121.

Muthard, C., & Gilbertson, R. (2016). Stress management in young adults: Implications of mandala coloring on self-reported negative affect and psychophysiological response. *Journal of Psychological Research, 21*, 16-28. doi:10.24839/2164-8204.jn21.1.16

Niven, K., Totterdell, P., & Holman, D. (2009). A classification of controlled interpersonal affect regulation strategies. *Emotion, 9*, 498-509. doi:10.1037/a0015962

Nolen-Hoeksema, S. (1991). Responses to depression and their effects on the duration of depressive episodes. *Journal of Abnormal Psychology, 100,* 569–582.

Nolen-Hoeksema, S. (2000). Further evidence for the role of psychosocial factors in depression chronicity. *Clinical Psychology: Science and Practice, 7*, 224-227. doi:10.1093/clipsy/7.2.224

Nolen-Hoeksema, S., & Morrow, J. (1991). A prospective study of depression and distress following a natural disaster: The 1989 Loma Prieta earthquake. *Journal of Personality and Social Psychology, 61*, 105-121.

Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking rumination. *Perspectives on Psychological Science, 3*, 400-424. doi:10.1111/j.1745-6924.2008.00088.x

Petrillo, L. D., & Winner, E. (2005). Does art improve mood? A test of a key assumption underlying art therapy. *Art Therapy, 22*, 205-212. doi:10.1080/07421656.2005.10129521

Robinson, M. S., & Alloy, L. B. (2003). Negative cognitive styles and stress-reactive rumination interact to predict depression: A prospective study cognitive therapy and research. *Cognitive Therapy and Research, 27*, 275-291. doi: 10.1023/A:1023914416469

Rockcliff, H., Gilbert, P., Mcewan, K., Lightman, S., & Glover, D. (2008). A pilot exploration of heart rate variability and salivary cortisol responses to compassion-focused imagery. *Clinical Neuropsychiatry, 5*, 132-139.

Rood, L., Roelofs, J., Bögels, S. M., & Meesters, C. (2011). Stress-reactive rumination, negative cognitive style, and stressors in relationship to depressive symptoms in non-clinical youth. *Journal of Youth and Adolescence, 41*, 414-425. doi:10.1007/s10964-011-9657-3

Roston, T. (2016). Why grown-ups love coloring books too. Retrieved from http://ideas.ted.com/why-grown-ups-love-coloring-books-too/.

Seligman, M. (1976). Learned helplessness: theory and evidence. *Journal of Experimental Psychology, 105*, 3-46. doi:10.1037/0096-3445.105.1.3

Small, S. (2006). Anxiety reduction: Expanding previous research on mandala coloring. *The Undergraduate Journal of Psychology*, *19*, 15-21.

Spasojevic, J., & Alloy, L. B. (2001). Rumination as a common mechanism relating depressive risk factors to depression. *Emotion, 1*, 25-37. doi:10.1037//1528-3542.1.1.25

Staricoff, L. R. (2003). Arts in health: a review of the medical literature. *Psychological Reviews, 102*, 4, 671-683.

Thompson, R. A. (1994). Emotion regulation: A theme in search of definition. *Monographs of the Society for Research in Child Development, 59*, 25-52. doi:10.2307/1166137

Thwaites, R., & Freeston, M. H. (2005). Safety-seeking behaviours: Fact or function? How can we clinically differentiate between safety behaviours and adaptive coping strategies across anxiety disorders? *Behavioural and Cognitive Psychotherapy, 33*, 177-188. doi:10.1017/s1352465804001985

Uttl, B., Morin, A., & Hamper, B. (2011). Are inner speech self-report questionnaires reliable and valid? *Social and Behavioural Sciences, 30,* 1719-172.

Vennet, R. V., & Serice, S. (2012). Can coloring mandalas reduce anxiety? A replication study. *Art Therapy, 29*, 87-92. doi:10.1080/07421656.2012.680047

Watkins, E. R. (2008). Constructive and unconstructive repetitive thought. *Psychological Bulletin, 134*, 163-206. doi:10.1037/0033-2909.134.2.163

Watkins, E., Moberly, N. J., & Moulds, M. L. (2008). Processing mode causally influences emotional reactivity: Distinct effects of abstract versus concrete construal on emotional response. *Emotion, 8*, 364-378. doi:10.1037/1528-3542.8.3.364

Watkins, E., & Teasdale, J. D. (2001). Rumination and overgeneral memory in depression: Effects of self-focus and analytic thinking. *Journal of Abnormal Psychology, 110*, 333-357. doi:10.1037//0021-843x.110.2.333

Watkins, E., & Teasdale, J. D. (2004). Adaptive and maladaptive self-focus in depression. *Journal of Affective Disorders, 82*, 1-8. doi:10.1016/j.jad.2003.10.006

Wisco, B. E., & Nolen-Hoeksema, S. (2008). Ruminative response style. In: K. S. Dobson & D. J. A. Dozois (Eds.), Risk factors in depression (pp. 221-236). San Diego, CA: Elsevier Academic Press. doi:10.1016/b978-0-08-045078-0.00010-1

Zung, W. W. (1986). Zung self-rating depression scale and depression status inventory. In: T. A. Ban & N. Sartorius (Eds.), Assessment of depression (pp. 221–231). New York: Springer Verlag. doi: 10.1007/978-3-642-70486-4_21

Appendices

Appendix 1 - Distraction condition information

Imagination Task

For the next 10 minutes, try your best to focus your attention on each of the ideas on the following pages.

Read each item slowly and silently to yourself. As you read the items, use your imagination and concentration to focus your mind on each of the ideas. Spend a few moments visualising and concentrating on each item.

For each item, I would like you to sketch any rough images that express what you are imagining. Do not spend very long on each item, perhaps 20 seconds or so on each. You are not expected to reach the end of the items.

Please continue until the experimenter returns.

Think about:

and imagine a boat slowly crossing the Atlantic

Think about:

the layout of a typical classroom

Think about:

the shape of a large black umbrella

Think about:

the movement of an electric fan on a warm day

Think about:

raindrops sliding down a window pane

Think about:

a double-decker bus driving down a street

Think about:

and picture a full moon on a clear night

Think about:

clouds forming in the sky

Think about:

the layout of the local shopping centre

Think about:

and imagine a plane flying overhead

Think about:

fire darting round a log in a fire-place

Think about:

the car park at a large supermarket

Think about:

two birds sitting on a tree branch

Think about:

the shadow of a stop sign

Think about:

the layout of the local post office

Think about:

the structure of a high-rise office building

Think about:

and picture the Eiffel Tower

Think about:

the shape of the continent of Africa

Think about:

a group of polar bears fishing in a stream

Think about:

the shape of Great Britain

Think about:

a train stopped at a station

Think about:

a lone cactus in the desert

Think about:

the shape of the country Italy

Think about:

a row of shampoo bottles on display

Think about:

a petrol station on a major road

Think about:

the fuzz on the shell of a coconut

Think about:

the queens' head on a stamp

Think about:

a band playing the National Anthem

Think about:

the shape of the United States of America

Think about:

the baggage claim area at the airport

Think about:

the size of the Statue of Liberty

Think about:

the shape of a cricket bat

Think about:

a freshly painted door

Think about:

the shiny surface of a trumpet

Think about:

a kettle coming to the boil

Appendix 2 - Rumination Condition Participants Information Sheet

Self-reflection task

For the next 10 minutes, try your best to focus your attention on each of the ideas on the following pages.

Read each item slowly and silently to yourself. As you read the items, use your imagination and concentration to focus your mind on each of the ideas. Spend a few moments visualising and concentrating on each item.

For each Idea, I would like you to write down any prominent words that describe your thinking. You do not have to write sentences, just a few words for each topic. Do not spend very long on each item, perhaps 20 seconds or so on each.

Please continue through the items until the experimenter prompts you to finish. You are not expected to reach the end of the items.

Think about:

the physical sensations you feel in your body

Think about:

your character and who you strive to be

Think about:

the degree of clarity in your thinking right now

Think about:

why you react the way you do

Think about:

the way you feel inside

Think about:

the possible consequences of your current mental state

Think about:

how similar or different you are relative to other people

Think about:

what it would be like if your present feelings lasted

Think about:

why things turn out the way they do

Think about:

trying to understand your feelings

Think about:

how awake or tired you feel now

Think about:

the amount of tension in your muscles

Think about:

whether you are fulfilled

Think about:

your physical appearance

Think about:

whether you feel stressed right now

Think about:

the long-term goals you have set

Think about:

the amount of certainty you feel

Think about:

your present feelings of fatigue or energy

Think about:

possible explanations for your physical sensations

Think about:

how hopeful or hopeless you are feeling

Think about:

the level of motivation you feel right now

Think about:

the degree of helplessness you feel

Think about:

the degree of calmness or restlessness you feel

Think about:

the possible consequences of the way you feel

Think about:

what your feelings might mean

Think about:

how sad or happy you are feeling

Think about:

the expectations your family has for you

Think about:

why your body feels this way

Think about:

why you get this way sometimes

Think about:

how passive or active you feel

Think about:

what people notice about your personality

Think about:

how optimistic or pessimistic you feel about the future

Think about:

how weak or strong your body feels right now

Think about:

the degree of relaxation or agitation you feel

Think about:

the kind of person you think you should be

Think about:

the degree of control you feel right now

Think about:

what would happen if your current physical state lasted

Think about:

sitting down and analysing your personality

Think about:

why you turned out this way

Think about:

the things that are most important in your life

Think about:

how quick or slow your thinking is right now

Think about:

the degree of decisiveness you feel

Think about:

trying to understand who you are

Think about:

how you feel about your friendship

Think about:

whether you have accomplished a lot so far

-

Appendix 3 - Ethical Approval Application

Your application for ethical approval (2017/1591) has been accepted

A apache@exeter.ac.uk on behalf of Ethics Approval System <D.M.Salway@exe

Reply all

Mon 13/03/2017, 15:58

Drew, Neil

Ethical Approval system

Your application (2017/1591) entitled Colouring-in, a distraction technique? A study looking at the effects of colouring-in and distraction on rumination and negative affect. has been accepted

Please visit http://www.exeter.ac.uk/staff/ethicalapproval/

Please click on the link above and select the relevant application from the list.

Ethical Approval

Title of project: Colouring-in, a distraction technique? A study looking at the effects of colouring-in and distraction on rumination and negative affect.

Name of researchers : Neil Drew

Nick Moderly

Lay summary of project: Aim

Our aim is to examine whether colouring-in is helpful for self-reported ruminators, in regards to reducing ruminating thoughts and aiding mood recovery, compared to other kinds of active distraction and rumination.

Primary research question

1. Is the use of adult colouring books more effective than a neutral distraction activity and rumination a condition in reducing negative mood and state rumination in self-reported ruminators after experiencing a laboratory stressor?

Dependent Variables: State rumination and negative affect

I want to examine whether colouring-in is as good as or better than another distraction activity and rumination. This study will look to build on the current distraction based interventions for rumination and aim to discover whether colouring-in can improve mood and reduce state rumination after a stressor. If positive benefits are discovered when compared to other distraction techniques, this would serve as an easy to use, simple intervention for those suffering from rumination, and would provide an initial evidence base for this approach.

Comparing colouring-in with an active control tells you something more specific about whether colouring-in has specific effects over and above an alternative distraction condition. The active control will involve participants completing the Nolen-Hoeksema (1991) distraction activity. This involves imagining an object (e.g. Empire State Building). To help match this condition with colouring, participants will also write down what they imagine. In order to support findings and increase their implications, a rumination control will also be included. A written version of Nolen-Hoeksema's (2000) rumination condition task will be used. This will allow the exploration of whether colouring-in is more effective than simply no intervention.

Communication and consent:

Will you describe the main experimental procedures to participants in advance, so that they are informed in advance about what to expect?	Yes ◉ No ○ n/a ○
Will you tell participants that their participation is voluntary?	Yes ◉ No ○ n/a ○
Will you obtain written consent for participation?	Yes ◉ No ○ n/a ○
If the research is observational, will you ask participants for their consent to being observed?	Yes ○ No ○ n/a ◉
Will you tell participants that they may withdraw from the research at any time and for any reason?	Yes ◉ No ○ n/a ○
With questionnaires, will you give participants the option of omitting questions they do not want to answer?	Yes ◉ No ○ n/a ○
Will you tell participants that their data will be treated with full confidentiality and that, if published, it will not be identifiable as theirs?	Yes ◉ No ○ n/a ○
Will you debrief participants at the end of their participation (ie. give them a brief explanation of the study)?	Yes ◉ No ○ n/a ○

Psychology: Details on misleading participants or inducing stress.
A stressor will be used with the aim to induce negative mood. Participants will be asked to complete a computerized version of an anagram stressor task. The anagrams have been used many times in previous studies with this population and the effects on mood have been shown to wear off by the end of the study, thus the distress is limited.

Stressor instructions:

This task tests your cognitive ability in solving anagrams, which has been shown to predict academic and career success. Most university students complete 60-70% of the anagrams successfully. During the task, each of the anagrams will be presented for 20 seconds, after which you will be given 10 seconds to type in the solution. The computer will then display “correct” or “incorrect” feedback, together with the correct solution and the proportion of anagrams solved so far. Please type carefully as there will not be a chance to correct errors. You should not begin typing until the computer prompts you to do so. For example, if you see the anagram “NHCLU”, the correct answer would be “LUNCH”.

The above description of the stressor to participants is misleading as most students don't get 60-70% of anagrams correct, they get 50% correct. Therefore, when the participants score below the "average", we hope this elicits some stress. This minor level of misleading participants will be fully explained to them during the oral and paper debrief (see attachment).

I will not recruit participants who are severely depressed as they may react badly to the stressor. Participants will be shown a humorous video at the end of the study. The aim of this is to avoid the participants leaving the study with heightened stress levels. All participants, including those excluded due to severe low mood, will be given details of where to seek psychological support (see attachment for services)

Vunerable Groups

Do participants fall into the following categories?	**Are your participants under the age of 18?**	Yes ○ No ◉ n/a ○
	If your participants are under the age of 18, will you be recruiting from schools/colleges?	Yes ○ No ○ n/a ◉
	People with learning or communication difficulties	Yes ○ No ◉ n/a ○
	Those at risk of psychological distress or otherwise vulnerable	Yes ◉ No ○ n/a ○
	People in custody	Yes ○ No ◉ n/a ○
	People engaged in illegal activities (e.g. drug taking)	Yes ○ No ◉ n/a ○

Purpose of project and academic rationale:
Doctorate Clinical Psychology Thesis

The purpose of this project is to examine whether colouring-in is helpful for self-reported ruminators, in regards to reducing ruminating thoughts and aiding mood recovery, compared to other kinds of active distraction and rumination. This will contribute to the current research on distraction and rumination.

Primary research question

1. Is the use of adult colouring books more effective than a neutral distraction activity and rumination a condition in reducing negative mood and state rumination in self-reported ruminators after experiencing a laboratory stressor?

Dependent Variables: State rumination and negative affect

I want to examine whether colouring-in is as good as or better than another distraction activity and rumination. This study will look to build on the current distraction based interventions for rumination and aim to discover whether colouring-in can improve mood and reduce state rumination after a stressor. If positive benefits are discovered when compared to other distraction techniques, this would serve as an easy to use, simple intervention for those suffering from rumination, and would provide an initial evidence base for this approach.

Comparing colouring-in with an active control tells you something more specific about whether colouring-in has specific effects over and above an alternative distraction condition. The active control will involve participants completing the Nolen-Hoeksema (1991) distraction activity. This involves imagining an object (e.g. Empire State Building). To help match this condition with colouring, participants will also write down what they imagine. In order to support findings and increase their implications, a rumination control will also be included. A written version of Nolen-Hoeksema's (2000) rumination condition task will be used. This will allow the exploration of whether colouring-in is more effective than simply no intervention.

Brief desscription of methods and measurements:

The study is a mixed (3 x 3) experimental design, with 'time' as a within factor (3 levels) and condition as the between subjects factor (3 levels). Participants will be randomly assigned to one of three conditions: (1) colouring-in; (2) written Nolen-Hoeksema distraction and; (3) written Nolen-Hoeksema rumination. Participants' state rumination and negative affect (dependent variables) will be measured after completing the anagram stressor, and then again after completing the distraction task (e.g. colouring-in, Nolen-Hoeksema activity).

Measures:

The State Rumination Questionnaire (SRQ).

The Positive and Negative Affect Schedule (PANAS).

The Patient Health Questionnaire 8 (PH8) - to help screen out participants who score within the depressed or anxiety range.

Participants

a. Human research:

I aim to recruit 90 participants (30 per condition), all of which will be 18 years or over, mixed sex, and current students at the University of Exeter. The participants will be recruited via: (1) the University of Exeter Research Participation System; (2) a general email to all students within the School of Psychology; (3) posters around the University campus; (4) an advert on the University website; (5) a general email to academic schools other than psychology. If required, participants will be encouraged to sign up to the study to have a chance to win Amazon gift vouchers via a prize draw. The exception will be participants who are recruited via the online Research Participation System, of whom will be awarded course credits.

exclusion/inclusion:

Participants do not have to have prominent symptoms of depression or a diagnosis to sign up to the study. When assigning to the study, participants will complete a pre-screening mood measure. Participants who score within the severe range of depression, based on the screening measures, will also be excluded from taking part

in the study for ethical reasons. I will not exclude participants based on their state rumination scores, however when advertising the study I will seek to recruit people who "think a lot" and have "repetitive thought" which affects their daily living. Thus it is realistic that participants will experience some mild distress if they are self-reported ruminators.

Clear consise statement of ethical considerations

A stressor will be used with the aim to induce negative mood. Participants will be asked to complete a computerised version of an anagram stressor task. Participants will be falsely informed that their scores fall below average. The anagrams have been used many times in previous studies with this population and the effects on mood wear off by the end of the study, so the distress will be limited.

I will not recruit participants who are severely depressed as they may react badly to the stressor (see exclusion criteria). Participants in all conditions will be fully orally debriefed (see attachment) at the end of their participation. This will explain the minor deception that occurred in the instructions of the stressor. At the end of the study, participants will also take part in a de-stressing activity (e.g. watching a humorous video).The aim of this is to avoid the participants leaving the study with heightened stress levels. All participants, including those excluded due to severe low mood, will be given details of where to seek psychological support if required, details of which are in an attachment I've provided (e.g. University counselling, GP, Increasing Access to Psychological Therapies).

Appendix 4 - Patient Health Questionnaire

Over the **last 2 weeks**, how often have you been bothered by any of the following problems?
*(circle **one** number on each line)*

How often during the past 2 weeks were you bothered by...	Not at all	Several days	More than half the days	Nearly every day
1. Little interest or pleasure in doing things	0	1	2	3
2. Feeling down, depressed, or hopeless	0	1	2	3
3. Trouble falling or staying asleep, or sleeping too much	0	1	2	3
4. Feeling tired or having little energy	0	1	2	3
5. Poor appetite or overeating	0	1	2	3
6. Feeling bad about yourself, or that you are a failure, or have let yourself or your family down	0	1	2	3
7. Trouble concentrating on things, such as reading the newspaper or watching television	0	1	2	3
8. Moving or speaking so slowly that other people could have noticed. Or the opposite – being so fidgety or restless that you have been moving around a lot more than usual	0	1	2	3

Appendix 5 - State Rumination Questionnaire (SRQ)

Post-anagram SRQ - 5-point Likert scale that ranged from 1 (*Not at all*) to 5 (*A lot*)

1. **To what extent did you keep thinking about the anagram task after it finished?**

2. **When you were performing the anagram task, to what extent did thoughts about your performance interfere with trying to solve the puzzles?**

3. **To what extent did you find it hard to stop thinking about the anagram task, since you started it?**

4. **To what extent did you have self-critical thoughts about your performance on the anagram task?**

5. **How much did you think about your past performance on academic or intelligence tasks?**

Post-manipulation SRQ: SRQ to be completed following manipulation (e.g. post colouring-in, distraction.). The SRQ questions have been tweaked slightly to reflect the amount participants continued to ruminate during the manipulation.

5-point Likert scale that ranged from 1 (*Not at all*) to 5 (*A lot*)

1. **When you were performing the previous task, to what extent did you keep thinking about the anagram task?**

2. **When you were performing the previous task, to what extent did thoughts about your performance on the anagram task interfere?**

3. **When you were performing the previous task, to what extent did you find it hard to stop thinking about the anagram task?**

4. **When you were performing the previous task, did you have self-critical thoughts about your performance on the anagram task?**

5. **When you were performing the previous task, how much did you think about your past performance on academic or intelligence tasks?**

Appendix 6 - Positive and Negative Affect Schedule

PANAS—SHORT FORM (Mackinnon et al., 1999)

This scale consists of a number of words that describe different feelings and emotions. Read each item and then type the appropriate answer to indicate your response. Indicate to what extent you feel this way **right now**, that is, at the present moment.

Not at all = 0
Very slightly = 1
A little = 2
Moderately = 3
Quite a bit = 4
Extremely = 5

Distressed

Excited

Upset

Scared

Enthusiastic

Alert

Inspired

Nervous

Determined

Afraid

Appendix 7 - Induced Stressor Anagram Task

Anagram task instructions

This task tests your cognitive ability in solving anagrams, which has been shown to predict academic and career success. Most university students complete 60-70% of the anagrams successfully. During the task, each of the anagrams will be presented for 20 seconds, after which you will be given 10 seconds to type in the solution. The computer will then display “correct” or “incorrect” feedback, together with the correct solution and the proportion of anagrams solved so far. Please type carefully as there will not be a chance to correct errors. You should not begin typing until the computer prompts you to do so. For example, if you see the anagram “NHCLU”, the correct answer would be “LUNCH”.

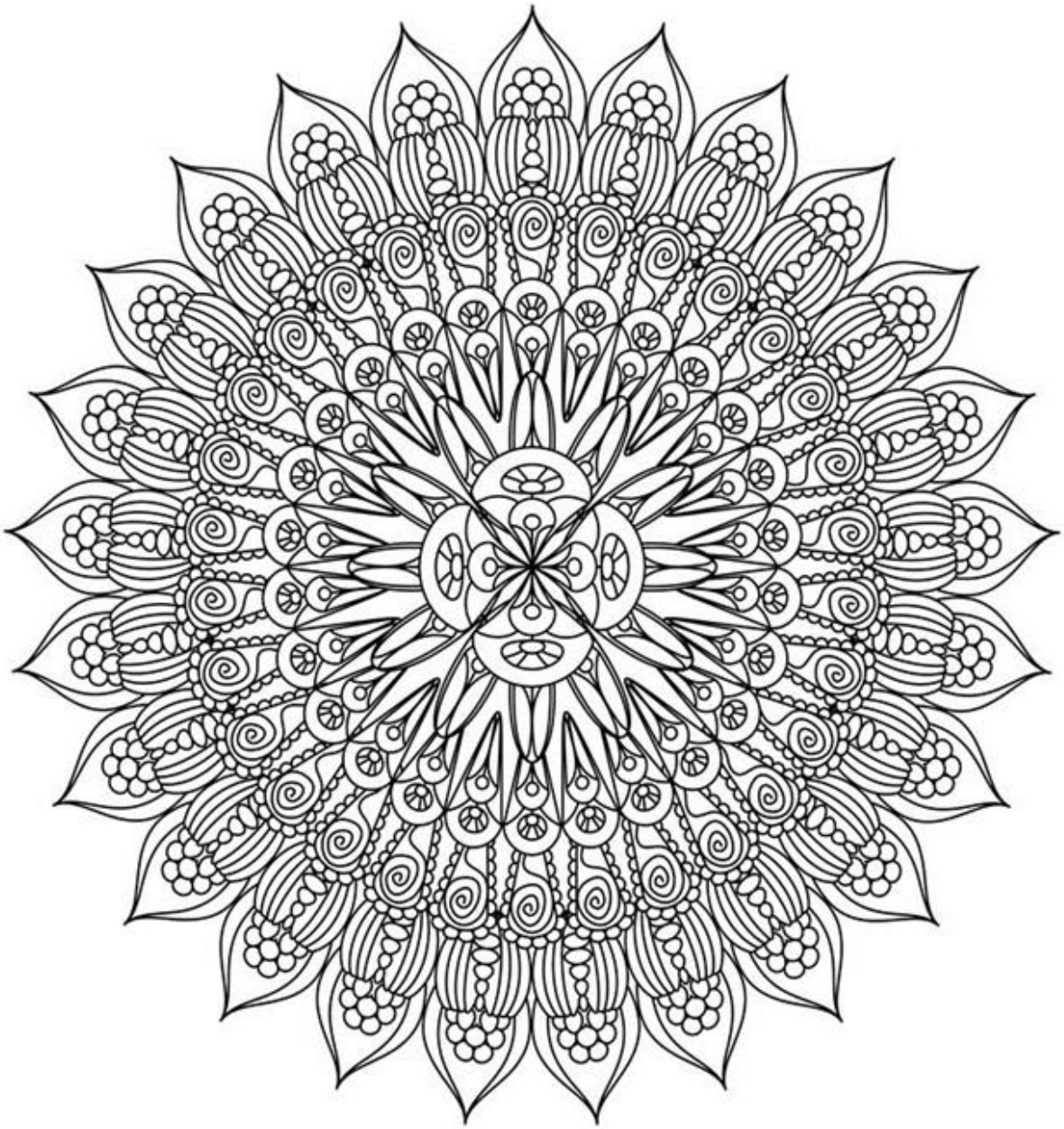
Anagrams used:

OLDME – MODEL
HROAB – ABHOR
DATIR – TRIAD
TINGA – GIANT
AEWTK – TWEAK
DGROU – GOURD
UCRSO – SCOUR
TANBO – BATON
AITOP – PATIO
MGEON – GNOME
RIGON – GROIN
MALBY – BALMY
RCOHA – ROACH
ARFYO – FORAY
GLAEI - AGILE

Appendix 8 - Colouring-in Condition

Colouring task

I would like you to spend 10 minutes colouring-in the pattern below. You can use any of the colours provided. You are not expected to finish colouring the entire pattern, so please take your time and do not rush. There is no right or wrong way to colour in the pattern. And it is up to you which colours you use and how to use them.



Appendix 9 - Information Sheet

Participant Information Sheet

UNIVERSITY OF EXETER

Aim of the study

This study aims to examine the relationship between our imagination and mood.

What do I have to do if I take part?

Taking part in the study will first involve completing some demographic details and mood measures online. You will then meet with me at the University to complete the study, which should last approximately 45 minutes. The study will involve completing questionnaires that ask you about your mood and thinking style. You will then complete a computer-based cognitive task that involves solving puzzles. Following this, you will complete another simple task.

All your personal details will remain confidential and secure, and will be stored securely and separately from other research data. The reported results of the research will only include information about the range of participants who took part (e.g., average age, results of questionnaires) and will not identify individual participants.

Psychology undergraduate will be given two study credits for taking part in this research. If you are not a Year 1 undergraduate, you will be entered into a prize draw. There will be approximately a 1 in 6 chance of winning a £10 Amazon voucher. Participants can only win one voucher.

What are the possible benefits of taking part?

The information I get from this study should help deepen the understanding of the way in which people regulate their emotions and will inform possible interventions for those who experience sad mood.

What will happen to the results of the study?

I aim to publish the work in an academic journal. Upon request, I will provide you with an information sheet about the results of the research. Your identity will not be revealed in any report or publication.

Possible risks:
This research has been approved by University of Exeter Psychology Ethics Committee. However, some measures/questionnaires will ask about difficult feelings, which some people may find emotive. The puzzle task has a time limit and some people may find completing it mildly stressful.

Contact for further information

If you would like further information about participating in this study, please email any questions to

nd320@exeter.ac.uk.

Appendix 10 - Consent Form

Consent form

UNIVERSITY OF EXETER

Participant Consent Form

Name of researcher: Neil Drew

1. I confirm that I have read and understood the information sheet for the above study	☐
2. I understand that my personal details will be kept secure and no identifiable details will be used as part of the research results.	☐
3. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason.	☐
4. I agree to take part in the study.	☐

Name:
Date of birth:

Signature:

Date:

Appendix 11 - Debrief Form

Debrief Form

Thank you for taking part in this research. The study aims to look at whether different distraction interventions can help reduce rumination (repetitive thinking about the self) and negative affect. You were placed in one of three conditions colouring-in, distraction via imagination, or rumination. Based on the recent popularity of colouring-in books, we wanted to see whether colouring-in can help relieve rumination and negative affect. We hypothesised that participants in the colouring-in condition would show a greater reduction in rumination and negative affect than those in the imaginative distraction and rumination conditions. Previous research has suggested that distraction is more effective than rumination in alleviating negative mood, but it is unclear whether colouring-in is more effective than distraction.

The anagram task was set up with the purpose of inducing mild stress, rumination and negative mood. Therefore, we told you some false information before the task. First, your performance in this task is not known to predict your intelligence, academic or career success. Second, most undergraduates solve less than half the anagrams. I ask you to please not break confidentiality by sharing details of this study to others, as it may hinder the value of this study.

This study included measures of negative affect, rumination and depression. If completion of these measures led to concerns about your mental health, it may be worth making an appointment with your GP to seek support. I can also provide you with a list of contact details for additional services to seek for support. Please email me if you would like a copy.

If you have any complaints regarding your participation in this research, you can email your concerns to Dr Lisa Leaver (Research Ethics Committee Chair) at L.A.leaver@exeter.ac.uk

Thank you again for your participation.
Name of researcher: Neil Drew
Contact details: nd320@exeter.ac.uk

Appendix 12 - Contact details of mental health support

CONTACT DETAILS OF ORGANISATIONS IF YOU REQUIRE HELP

YOUR GP

If you have been experiencing low mood most of the day for several days or weeks, you should consider consulting your GP, who can provide professional guidance and help. If you are a student at the University of Exeter, you should be able to access a GP via registration with the Student Health Centre:
http://www.exeterstudenthealthcentre.co.uk/

UNIVERSITY OF EXETER WELLBEING SERVICE

The Wellbeing Service is available free of charge to all students, full-time, part-time, undergraduate and postgraduate. Because student life can be stressful, the Wellbeing Service is there to provide confidential help and support. We aim to help students cope more effectively with any personal problems or emotional difficulties that may arise during their time at University.

Telephone: (01392) 724381

Or book an appointment via the website:
http://www.exeter.ac.uk/wellbeing/appointments/

Student Wellbeing Service
Reed Mews Wellbeing Centre
Streatham Drive
Exeter EX4 4QP

VOICE (University of Exeter)

Voice is a student-run listening and information service, run by students for fellow students at the University of Exeter and is available from 8pm to 8am every night during term time. It is completely confidential, anonymous and prejudice-free, which means you can call with the confidence of knowing you can discuss anything you want without being judged.

Telephone (8pm – 8am): 4000 (internal, free of charge)
External: (01392) 724000
Email: exetervoice@googlemail.com

Website: http://www.exeterguild.org/voice/

EXETER SAMARITANS

Samaritans provides confidential emotional support, 24 hours a day for people who are experiencing feelings of distress or despair. Samaritans are there if you're worried about something, feel upset or confused, or you just want to talk to someone.

10 Richmond Road
Exeter
Devon
EX4 4JA (open daily 10.30am – 9.30pm)

24 hour telephone helpline: 01392 411711 (Exeter branch) / 08457 909090 (national)

Email: jo@samaritans.org

Website: http://www.samaritans.org/branches/samaritans-exeter-mid-east-devon

DEPRESSION ALLIANCE

Depression Alliance are a charity working to relieve and to prevent depression by providing information, support and understanding. Depression Alliance offer a range of publications and self-help groups.

Depression Alliance
20 Great Dover Street
London
SE1 4LX

Telephone: 0845 123 23 20 (for an information pack only)

Email: information@depressionalliance.org

Website: http://www.depressionalliance.org/