

Effects of behavioural activation on emotional cognition and mood



Tereza Ruzickova

Department of Psychiatry
University of Oxford
Oxford, UK

St John's College

Thesis submitted for the degree of DPhil in Psychiatry

Supervisors:

Prof. Catherine Harmer

Dr. Susannah Murphy

Acknowledgements

I am grateful to all the participants who volunteered their time for the two experiments despite struggling with low mood at the time. I would like to thank my supervisors Professor Catherine Harmer and Dr Susannah Murphy, who were the best supervisors I could ask for and gave me all the support and inspiration that I needed. I'm grateful to all my friends in the Neurosciences building who were there to support me as well as to my dance community who kept me relaxed and entertained outside of the lab.

Funding acknowledgements

I would like to express my thanks to the Medical Research Council, Department of Psychiatry and St John's College for supporting me financially over the course of this DPhil.

Declaration

I declare that this thesis has been composed solely by myself and that I have not submitted it — in whole or in part — in any previous application for an academic degree. I was the primary researcher of both of the experiments presented. With the guidance of my supervisors, I designed each study, obtained ethical approval, carried out daily study management (including recruitment and DSM-V screening), administered the study interventions, processed, and analysed the data and finally wrote this thesis. I also declare that parts of the COVID-19 study presented in Chapters 2 and 3 have previously been published in *Psychological Medicine* (Ruzickova et al., 2021, see References).

Several colleagues provided invaluable help in this process. In the first online experiment, Stirling Argabright assisted with phone screening, James Carson carried out half of the DSM-V screening and half of the interventions, Anna Pearse assisted with data processing and Amy Gillespie provided R scripts for data processing. In the second experiment, several research assistants helped with data collection and data processing, namely Ingrid Martin, Wesley Quadros, Evelyn Watson, Nora Rady, Chloe Robbins and Kristyna Stastna. Cassandra Gould Van Praag and Chamith Halahakoon provided scripts for data processing and Priyanka Panchal provided advice for actigraphy processing.

Table of Contents

Acknowledgements	2
Funding acknowledgements	2
Declaration	3
Abstract	8
Chapter 1: Introduction	10
Depression.....	10
Depression in the context of the COVID-19 pandemic	11
Behavioural activation	12
Treatment rationale.....	12
Treatment efficacy.....	13
Treatment format	14
Behavioural activation mechanism.....	16
Emotional cognition	16
Environmental reinforcement	20
Reward processing.....	21
Physical activity	23
Non-specific factors in psychological interventions.....	24
Thesis objectives.....	25
Chapter 2: Examining the effects of behavioural activation on mood in comparison to a passive control group during the COVID-19 pandemic	26
Introduction.....	26
Methods	28
Participants.....	28
Power calculation	29
Procedure	29
Intervention.....	30
Questionnaire measures	32
Statistical analysis.....	35
Results.....	36
Demographic, clinical and COVID-related baseline information	36
Rating of intervention.....	38
Questionnaire measures	39

Sensitivity analysis: effect of antidepressant treatment	47
Discussion.....	48
Strengths and limitations	49
Conclusion of Chapter 2	50
<i>Chapter 3: Examining the effects of behavioural activation on cognition in comparison to a passive control group during the COVID-19 pandemic</i>	<i>51</i>
Introduction.....	51
Methods	53
Cognitive tasks.....	53
Procedure	57
Power calculation	57
Statistical analysis.....	57
Results.....	58
Facial Emotion Recognition Task (FERT)	58
Emotional categorisation task (ECAT)	66
Correlation findings for early ECAT changes at week 2 and depression symptom changes at week 4 in the BA group.	66
Emotional Recall Task (EREC).....	66
Probabilistic Instrumental Learning Task (PILT)	67
Sensitivity analysis: effect of antidepressant treatment	68
Discussion.....	68
Strengths and limitations	72
Conclusion of Chapter 3	73
<i>Chapter 4: Examining the effects of behavioural activation on mood in comparison to an active and a passive control group</i>	<i>74</i>
Introduction.....	74
Methods	76
Participants.....	76
Power calculation	77
Procedure	78
Interventions	79
Questionnaire measures	80
Statistical analysis.....	81

Results.....	82
Demographic and clinical baseline information	82
Questionnaire measures	84
Discussion.....	96
Strengths and limitations	97
Conclusion of Chapter 4	99
<i>Chapter 5: Examining emotional cognition as a mechanism of behavioural activation in comparison to an active and a passive control group</i>	<i>100</i>
Introduction.....	100
Methods	102
Participants.....	102
Power calculation	102
Cognitive tasks.....	103
Procedure	104
Statistical analysis.....	104
Results	104
Facial Emotion Recognition Task (FERT)	105
Emotional categorisation task (ECAT)	109
Emotional Recall Task (EREC).....	112
Probabilistic Instrumental Learning Task (PILT)	113
Auditory Visual Learning Task (AVLT).....	114
Discussion	116
Strengths and limitations	118
Conclusion of Chapter 5	119
<i>Chapter 6: Examining the mechanism of behavioural activation using actigraphy measurement in comparison to an active and a passive control group</i>	<i>120</i>
Introduction.....	120
Methods	122
Participants.....	122
Procedure	123
Outcome measures	123
Data processing	124
Statistical analysis.....	125

Results	126
M10 activity	126
M10 onset	127
L5 activity	127
L5 onset	128
Relative amplitude (RA)	128
Inter-daily stability (IS)	130
Intra-daily variability (IV)	130
Correlating change in depression with change in actigraphy parameters	130
Discussion	131
Strengths and limitations	132
Conclusion of Chapter 6	133
<i>Chapter 7: General discussion.....</i>	<i>134</i>
Synopsis.....	134
Conclusion	139
<i>Appendix.....</i>	<i>140</i>
Chapter 2.....	140
Chapter 3.....	140
Chapter 4.....	142
Chapter 5.....	143
Chapter 6.....	145
Chapter 6.....	146
<i>References</i>	<i>148</i>

Abstract

Depression is one of the leading causes of disability worldwide and its prevalence has increased further during the COVID-19 pandemic. There is a need for effective treatments that can be easily and cheaply disseminated worldwide. Behavioural activation is known to be an effective intervention for depression which can be disseminated by non-specialist practitioners. However, its exact mechanism of action is currently unknown, which impedes our understanding for how to deliver it most effectively.

The aim of this thesis is to examine the efficacy of a 4-week course of non-specialist behavioural activation and to investigate its possible mechanisms of effect. To this end, two experiments were carried out with participants suffering from low mood. The first study examined the intervention in an online video format during the COVID-19 pandemic and compared it with a passive control group. We examined its effects on self-reported measures of mood and found significant improvements on measures of depression, activation, state anxiety and anhedonia. We also found that the intervention led to a more positive affective bias on accuracy and misclassification measures of the Facial Emotion Recognition Task. Moreover, early changes on emotional cognition correlated with later improvements in depressive symptoms. The results of this study indicate that this format of behavioural activation can not only be effective during a global pandemic, but its effects may be carried out through early changes in emotional cognitive processing, which could help us predict who it will be most suitable for.

The second study compared the intervention to both an active control group, which only carried out activity monitoring, as well as a passive control group. We again found that behavioural activation had significant beneficial effects on self-reported measures of mood and activation in comparison to both of the groups. Activity monitoring on its own also led to significant decreases in depression symptoms, albeit to a lesser extent than behavioural activation. We did not replicate the emotionally cognitive effects from the first study, possibly due to reduced power as a result of technical and COVID-19 related disruptions to the data collection. We found that early increases in environmental reward predicted future decreases in depression symptoms, indicating another mechanistic factor that may be important for the success of behavioural activation. We found no significant effects of the two behavioural interventions on actigraphy parameters indicating participants' physical activity and circadian rhythm. Early

changes on these parameters did not predict future depressive symptoms, suggesting that they may not be mechanistically involved.

In conclusion, both of the experiments clearly showed that a 4-week format of non-specialist behavioural activation can provide significant benefits on self-reported mood measures. Moreover, activity monitoring on its own may be a useful intervention for depression. These findings can inform policy related to mental health provision during societal crisis periods and other times of unusual strain on mental health services. Reward processing and actigraphy factors did not appear to play a mechanistic role in behavioural activation. Further replication is needed to clarify the effects of the intervention on emotional cognition and environmental reinforcement, which may be promising mechanistic factors allowing for prediction of treatment success.

Chapter 1: Introduction

Depression

Depression is a common mental health problem linked to reduced quality of life (Pyne et al., 1997) and functional disability (Spijker et al., 2004). It is associated with poor physical health, such as having an alcohol use disorder (Sullivan et al., 2005) or diabetes (Knol et al., 2006). The total global prevalence is estimated at 5%, although high quality estimates from low- and middle-income countries are missing (Ferrari et al., 2013). It is the fifth cause of global disease burden as measured by years lived with disability (Vos et al., 2017), predicted to be the leading cause of burden in developed countries by 2030 (Mathers and Loncar, 2006).

A major depressive episode is defined by symptoms of depressed mood and/or reduced ability to experience pleasure (anhedonia) experienced consistently for two weeks or more (DSM-5). Additional symptoms must also be present, such as fatigue, feelings of worthlessness as well as dysregulated sleep or appetite, and these symptoms must have a noticeable negative impact on the patient's daily functioning. Major depressive episodes tend to recur for many patients, and each additional episode increases likelihood of further recurrence (Kendler et al., 2001). Nonetheless, even subclinical symptoms of depression that do not meet the diagnostic criteria have been linked to significant impairment in daily life, such as household strain, social irritability, and reduced functioning at work (Judd and Wells, 1996), and can themselves be a precursor to major depression (Brown et al., 1986).

Considering the magnitude of negative impact that depression has on society, ensuring the effectiveness and availability of treatments is of crucial public health importance. Current pharmacological treatments, such as selective serotonin reuptake inhibitors (SSRIs) or tricyclic antidepressants, while more effective than placebo (Cipriani et al., 2018), appear to have limited efficacy for many patients. Estimates from the STAR*D trial indicate that only about 50% of patients respond to their first antidepressant drug, and 30% remain unremitted after trying four different medications (Rush, 2007). According to recent meta-analyses, combining pharmacotherapy with psychotherapy can significantly improve efficacy (Cuijpers et al., 2013), with beneficial effects persisting at six months or longer (Karyotaki et al., 2016). However, as reported by the WHO, evidence-based psychotherapy is expensive to deliver and unavailable in

many parts of the world (Bruckner et al., 2011). As a result, there is still significant room for improvement in the global availability of depression treatment.

Depression in the context of the COVID-19 pandemic

The existing issues in delivering depression treatment have become even more pressing with the emergence of the COVID-19 pandemic on 11th March 2020 as declared by WHO. As foreshadowed in past research on the SARS epidemic (Hawryluck et al., 2004; Reynolds et al., 2008; Mihashi et al., 2009), a public health emergency of this scale is likely to increase many risk factors for mental illness, such as frustration, isolation, uncertainty, income reduction or bereavement. As a result, mental health researchers published a call to action in June 2020, declaring "an urgent need to address how mental health consequences for vulnerable groups can be mitigated under pandemic conditions" (Holmes et al., 2020).

Deterioration in population mental health during the COVID-19 pandemic, including increased rates of depression, has now been widely reported from the United Kingdom (Pierce et al., 2020; Kwong et al., 2021), United States (Ettman et al., 2020), China (Wang et al., 2020), Austria (Pieh et al., 2020), Czech Republic (Winkler et al., 2020) and Cyprus (Solomou & Constantinidou, 2020) among many others. Moreover, these increases in depression may result not only from the unprecedented socio-economic context but also directly from COVID-19 infection (Taquet et al., 2020). Since psychiatric symptoms can exacerbate recovery from physical illness (Shevlin et al., 2020), this could lead to a vicious cycle of mental and physical health burden within the healthcare system. All of this has been taking place at a time of unusual strain on existing services, with most of psychological and psychiatric care reduced in capacity or transformed into a different format of delivery. As a result, there is clear urgency to devise new solutions for disseminating mental health care.

Several cross-sectional studies have investigated what factors might play a protective role within this context. One UK study found that mental wellbeing was associated with levels of physical activity (Jacob et al., 2020), which may be an important factor considering the worldwide decrease in phone-based step-counts estimated during the pandemic (Tison et al., 2020). Furthermore, a Spanish survey reported that several other lifestyle factors, such as following a routine, pursuing hobbies and spending time outdoors, were associated with lower levels of

depression (Fullana, Hidalgo-Mazzei, Vieta, & Radua, 2020). Time spent focusing on the pandemic has also been identified as a risk factor (Fullana et al., 2020). However, due to a lack of randomised controlled studies, it is not possible to infer the direction of causality or investigate whether other variables might better explain these effects.

It is crucial to experimentally examine what interventions may be most helpful during societal crisis periods like the COVID-19 pandemic. It has been argued that the provision of cognitive-behavioural therapy (CBT) should be increased (da Lopes & Jaspal, 2020); however, this intervention is costly and not easily accessible around the world. It has been found that behavioural activation (BA) can be as effective for depression as CBT but more cost-effective because it can be delivered by junior mental health workers with shorter training (Richards et al., 2016). BA helps patients examine their daily behaviour and assists them in finding the right balance of routine, pleasurable and necessary activities in order to overcome lethargy. This could make it uniquely suitable for the pandemic period, which has brought about such a significant lifestyle disruption.

Behavioural activation

Treatment rationale

Behavioural activation (BA) is an established treatment for depression and one of the first-line interventions currently offered on the NHS-funded system for Improving Access to Psychological Therapies (IAPT). It targets a common behavioural pattern in depression, wherein low mood leads to low levels of activity, which then further exacerbates the low mood. The BA model postulates that over time, this leads to a decrease in response-contingent positive reinforcement (Lewinsohn, 1974), such as experiences of connection from social activities or fulfilment from accomplishing necessary tasks. Behaviours end up largely driven by negative reinforcement, the avoidance of or escape from negative experiences. This can lead to a passive lifestyle based around "numbing" activities, such as excessive sleeping, consumption of TV and social media or abusing addictive substances. There may also be an increase in experiences of punishment, due to avoidance of activities that are necessary for maintaining study, work, personal finance, health, or relationships.

BA aims to disrupt this vicious cycle through a combination of psychoeducation on the nature of low mood and activity, activity monitoring, goal setting and problem solving. With the support and accountability of a BA practitioner, clients try to reconnect with meaningful activities despite their low mood and gradually increase their activity levels by completing small and specific goals. This increase in activation then tends to consequently improve mood. Numerous BA manuals exist describing this process, most notably Martell et al. (2010) and Lejuez et al. (2011), but various novel forms of BA have also been explored in recent years. For example, there has been interest in incorporating elements of Acceptance and Commitment Therapy (ACT), such as focus on values, willingness to change as well as using mindfulness to cope with rumination (Kanter et al., 2009).

Treatment efficacy

A body of literature has now amassed in support of the BA model of depression. BA was originally developed based on evidence showing a link between depression and a reduction in pleasant events experienced by the individual (Lewinsohn & Libet, 1972; Lewinsohn & Graf, 1973; Rehm 1978) as well as an increase in unpleasant events (Lewinsohn and Talkington, 1979). During the same time, the term appeared in the extensive work by Grey and McNaughton (1982) who defined the behavioural activation and inhibition systems as important physiological drivers of behaviour. BA subsequently became a part of Beck's cognitive therapy of depression, as it was suggested that behavioural strategies should be used with more severely depressed patients before it is possible to work directly with cognition (Beck et al., 1979).

BA only became a stand-alone treatment after a component analysis study showed that its effects were equivalent to full CBT when treating depression. This was found both at the end of treatment and at 6-month follow-up (Jacobson and Dobson, 1996) as well as in relapse prevention measured at a 2-year follow-up (Gortner et al., 1998). Despite only targeting behaviour, these studies found that BA altered automatic negative thoughts and dysfunctional attributional styles as much as full CBT. Later, BA was found to outperform cognitive therapy and showed equivalent efficacy to antidepressant medication in more severely depressed patients (Dimidjian et al., 2006) with a significantly lower drop-out rate. This was further confirmed in a 12-week clinical trial by Webb et al. (2019), who found that patients' engagement with the BA component of CBT predicted improvement more than their self-rated use of cognitive skills.

Most recently, a network meta-analysis by Ciharova et al. (2021) reported no difference in effectiveness between cognitive restructuring, BA and CBT.

Several meta-analyses reported that BA seems to be effective for depression in general adult population (Cuijpers et al., 2007; Mazzucchelli et al., 2009; Ekers et al., 2014), although most of the randomised controlled trials have only used passive control groups (such as waiting list). Early evidence also suggests that BA is effective in specific subgroups, such as in young people (Tindall et al., 2017), older people (Orgeta et al., 2017) or informal caregivers (Xu et al., 2019; Zabihi et al., 2020). It has also been successfully adapted for different cultural contexts, including low- and middle-income environments, such as in Kenya (Bryant et al., 2017), India (Patel et al., 2017) or Indonesia (Arjadi et al., 2018). Overall, BA has a strong evidence base, which is why it is recommended by the UK guidelines for “persistent sub-threshold depressive symptoms or mild to moderate depression” (National Institute of Health and Care Excellence, [NICE], 2009).

Treatment format

BA has been investigated in a variety of formats, suggesting that it can be flexibly adapted to the individual circumstances of each treatment context. For example, it has been administered for various lengths of time, ranging from 8-10 week (Carlbring et al., 2013; Patel et al., 2017; Arjadi et al., 2018) to 12-24-week paradigms (Jacobson et al., 1996; Dimidjian et al., 2006; Richards et al., 2017). The standard provision of low-intensity treatments in the UK IAPT system involves 8 sessions or less. This format is associated with about 50% relapse rate in the subsequent year, with higher likelihood for those who leave the intervention with residual symptoms (Ali et al., 2017).

However, even much shorter courses of BA have been explored. For example, Hopko et al. (2003) proposed a brief behavioural activation paradigm lasting just 2 weeks, delivered in 20min sessions three times a week. They reported that this format was more feasible for shorter inpatient stays and still produced large effects during the short-term in comparison to an active control group. More recently, Gawrysiak et al. (2009) tested the effects of just a single BA session, finding strong effects after a 2-week period, and similarly Nasrin et al. (2017) found that a single session led to a significant decrease in depression over just one week. Nonetheless,

it remains unclear what is the lowest “dose” of BA that can lead to sustained effects and reliable prevention of relapse.

Another widely researched aspect of BA delivery concerns the training necessary for practitioners. It has been argued that a significant advantage of BA over other forms of psychological treatment is its relative simplicity, as the treatment rationale is not difficult to teach to either practitioners or clients. This implies that BA can be delivered by a wide range of roles without the need for long and complex training. Together with other so-called low-intensity CBT therapies, which can be delivered in more cost-effective formats (Bennett-Levy et al., 2010), this has created an unprecedented opportunity for extensive dissemination of treatment through lay-administration.

This possibility was first examined with BA by Ekers et al. (2013), who provided mental health nurses with just 5-days’ worth of training to deliver the treatment and found their outcomes comparable to those of specialist therapists. More robust evidence later came from the large COBRA trial, where Richards et al. (2016) trained junior mental health workers (graduates without formal psychotherapy training) to deliver a guided BA intervention under supervision. In this non-inferiority trial of almost 450 patients, BA administered in this way was found to be as effective as full CBT delivered by accredited therapists - but more cost-effective. The potential ease of BA training and dissemination has also made it appealing for use in developing countries, where specialist mental health care provision is scarce. For example, a few recent studies have shown it to be helpful for depressed women experiencing partner violence in India (Patel et al., 2019) or for substance use in South African HIV care settings (Magidson et al., 2020). This evidence suggests that lay-delivery of simple psychological interventions can be effective and could revolutionise availability to mental health care around the world.

A third important aspect of BA which is currently an active topic of investigation relates to the recent boom in digitalised psychotherapy. Traditionally, BA was delivered through face-to-face contact with a practitioner, but the increasing popularity and evidence base of online mental health programmes, together with the social distancing measures of the pandemic, have stimulated interest in remote forms of BA delivery. Preliminary evidence suggests that BA can be effective in the format of a standalone app (Ly et al., 2014) or as a blended form of treatment combining face-to-face therapy with the use of an app, decreasing therapist time by 47% (Ly et

al., 2015). However, standalone digital programmes face the issue of high dropout rates and at least minimal level of human support may be necessary to maintain treatment adherence (Hilvert-Bruce et al., 2012). Nonetheless, remote formats of BA treatment show potential and may be another important avenue through which it can be widely disseminated.

Together, this evidence suggests that BA can be an effective yet very flexible treatment that can be quickly taught and adapted to different contexts in which depression occurs. Considering the complex global impact on mental health during the COVID-19 pandemic, the rising incidence of depression worldwide and the urgent need for accessible treatments, BA appears to be an extremely relevant intervention that could be usefully deployed to alleviate the current burden of depression.

Behavioural activation mechanism

Despite the amount of evidence supporting BA efficacy, its mechanism of action has not been clearly established. A recent systematic review by Janssen et al. (2021) was not able to find any conclusive evidence for any mediators of BA, only reporting weak support for the importance of activation and environmental reward. Kazdin (2007) and Holmes et al. (2018) have argued for the importance of mechanistic research in psychotherapy to improve our understanding of the variability in response and to aid in predicting successful treatment outcome. It could also be helpful from a precision-medicine standpoint as a more detailed understanding of treatment may allow for a closer match with the individual symptom profile of a patient.

It is therefore important to explore what other factors could be mediating the effects of BA and whether they could predict treatment success. Promising candidates for BA mechanism include emotional cognition, environmental reinforcement, physical activity, reward processing as well as the common factors of psychotherapy, such as the therapeutic relationship and setting.

Emotional cognition

One possible mechanism of BA that has not yet been investigated concerns changes in affective bias on measures of emotional ("hot") cognition. This domain is usually probed by adapting traditional "cold" cognition tasks so that they include emotional stimuli. Common endpoints

include how well participants remember positive and negative emotional words, how quickly they respond to positive or negative emotional images and how they interpret ambiguous facial expressions. Neurocognitive paradigms are also commonly used; for example, measuring the BOLD response in emotion-related neural regions, such as the amygdala, while participants perceive positive and negative affective stimuli.

Depression is known to be associated with negative affective biases in several domains of cognition, including facial perception, attention, working memory as well as long-term memory (Disner et al., 2011). Moreover, it has been linked to a reduction in the predominantly positive bias that is normally found in healthy participants when viewing ambiguous stimuli (Milders et al., 2010). These biases appear to be a risk factor for the onset of a major depressive episode (Joormann, Talbot, & Gotlib, 2007) as well as a predictor of relapse after treatment (Bouhuys, Geerts & Gordijn, 1999). This suggests they may play a causal role in depression rather than just being an epiphenomenon of its symptoms. As a result, it has been argued that cognitive variables may be useful endophenotypes for researching the molecular genetics of this condition (Glahn et al., 2012).

One explanation for how these biases lead to depression is that they constitute a “top-down” negative prediction of what may happen in the environment. For example, an ambiguous face may be interpreted as angry because the individual expects negative outcomes to occur in their social interactions, possibly due to a high frequency of such outcomes in the past, such as during childhood abuse or bullying. This is consistent with Beck's cognitive model of depression (Beck et al., 1979) wherein core “negative schemata” are thought to drive automatic negative cognitions and emotions. This rationale forms the basis of cognitive behavioural therapy with techniques such as cognitive restructuring, wherein patients learn to correct their automatic interpretations of the world (“cognitive errors”). This could be seen as a top-down modulation of one's own negative predictions.

However, an alternative explanation has been put forward by Roiser and Sahakian (2013). They argue that these negative biases may first arise at the “bottom-up” stage of sensory processing as a result of altered monoamine transmission. These neurobiological factors could arise from a number of genetic and environmental influences as well as their possible interaction, although the exact mechanisms remain to be elucidated. The resulting automatic negative perception

would then cause the development of negative schemata, which may bias “top-down” processes later, see Figure 1 below.

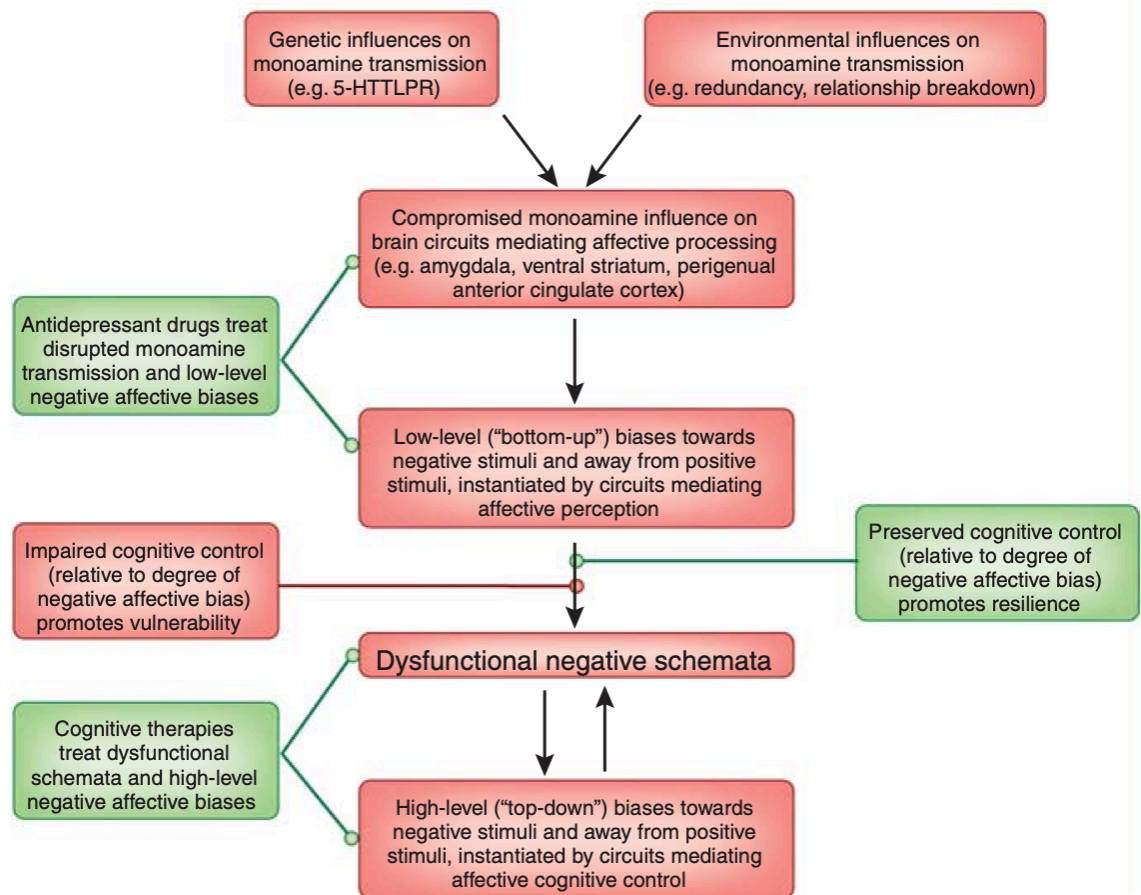


Figure 1. A cognitive neuropsychological model of depression by Roiser and Sahakian (2013). Red boxes illustrate risk factors for developing depression, green boxes represent possible therapeutic mechanisms.

This hypothesis relates to important findings on how emotional cognition responds to treatment with antidepressant drugs. As reviewed by Harmer, Goodwin and Cowen (2009), antidepressant medications appear to change emotional cognition early on in treatment by shifting cognitive biases to a more positive valence, and this is thought to precede mood change. For example, Harmer et al. (2004) reported that just a week-long administration of the antidepressants citalopram or reboxetine in healthy volunteers led to reduced recognition of negative facial expressions when compared to placebo. These findings have led to the proposal that more positively biased interpretation of external stimuli induced by antidepressants may eventually accumulate and improve mood as a result. In line with the “bottom-up” model described above,

the drugs may shift monoamine neurotransmitters in a way that produces positive sensory processing, which could eventually perhaps override negative schemata. This has come to be known as the cognitive neuropsychological model of antidepressant drug action and a significant body of evidence has supported it (see Harmer et al., 2017 for review).

These findings have raised the possibility that emotional cognition may play an important role in other kinds of treatments for depression as well, possibly being a common neural pathway among many, if not all effective treatment approaches. This is supported by similar evidence found with other interventions, such as negative ion treatment for seasonal affective disorder (Harmer et al., 2012) as well as with St John's wort (Warren et al., 2019) and transcranial direct current stimulation (Ironsides et al., 2016) administered to healthy participants. However, this has not been found for bright light treatment (Kaltenboeck et al., 2022), possibly because only a single dose was examined or because the intervention works through a different mechanism.

When it comes to cognitive and behavioural treatments for depression, changes in automatic emotional information processing have not been widely studied, and the few findings that exist have been inconsistent. An early study by Segal and Gemar (1997) deployed the emotional stroop task before and after a course of CBT for depression, reporting that automatic negative information about the self was less accessible after the treatment. A more recent study by Porter et al. (2016) found no changes in facial emotion processing after 16 weeks of CBT or schema therapy in comparison to healthy controls, despite significant decreases in mood symptoms. On the other hand, Vazquez et al. (2018) found a significant decrease in time spent fixating on negative faces and increased attention to positive faces in a pre-post comparison after both 10 weeks of CBT and a positive psychology intervention.

These contradictory findings may result from the fact that CBT contains many different treatment elements, both cognitive and behavioural, that are chosen to suit the individual needs of each patient. These elements could have different effects on emotional cognition, thus confounding the findings. It could therefore be useful to examine emotional cognition just in BA treatment, which only uses a narrow set of behavioural techniques. It is currently unclear whether such changes in emotional processing could occur during this intervention or what role they could play within it.

One possibility is that the increased exposure to positive environmental reinforcement through meaningful activities could provide enough positive “bottom-up” stimuli to shift the “top-down” cognitive model towards expecting more positive outcomes. This would be in line with the computational model of mood by Eldar et al. (2016), according to which mood is determined by the total ratio of positive and negative prediction errors recently experienced in the environment - and the same could apply to affective bias. Nonetheless, it could also be argued that BA includes an important “top-down” component as patients learn to notice and resist their automatic negative instincts towards avoidance and passivity (although the need for cognitive effort should be reduced by the focus on small and easy goals).

All in all, there are plausible hypotheses as to why BA could be another treatment that works through changes in affective bias and how this cognitive shift could unfold. This kind of mechanism could also help explain why the new positive experiences are rewarding for patients and aren’t overshadowed by rumination or negative interpretation as they are being acquired. Testing this hypothesis could help explain the inconsistency in affective bias research on CBT as well as elucidating why some patients respond to BA better than others.

Environmental reinforcement

Another factor that may be important in BA is the nature of environmental reinforcement in the patient’s life. As outlined earlier, the onset of depression is often followed by a loss of positive reinforcement and an increase in negative reinforcement and punishment. However, this can be the cause of depression as well as its consequence, since depression is often preceded by stressful life events (Kendler, Karkowski & Prescott, 1999) that are likely to affect the reward and punishment available in a person's environment (e.g. the loss of a job could lead to a loss of income, routine and social contact). BA aims to rebuild frequency of positive reinforcement, but patients may have varied opportunities for doing so as a result of their individual sociocultural context.

One such source of variability could be the patient’s socioeconomic status (SES). It has previously been reported that lower SES is associated with worse outcomes in depression treatment across several modalities (Falconnier, 2010) and this could plausibly occur in BA as well. For example, BA encourages patients to reconnect with their hobbies, but the feasibility of this could be

affected by the patient's material resources as well as time. Other factors related to SES, such as living in rural areas or having a disability, could also affect the accessibility of patient's preferred activities and thus their ability to carry out the BA protocol. It is therefore important to examine whether this could play a role in BA success.

Another factor constraining positive reinforcement in one's environment is the amount of social support that is available. BA encourages activities that reconnect the patient with friends and family members, but patients may differ widely in their level of isolation or the quality of their relationships, leading to varied opportunities for social connection. Moreover, support from loved ones can help the patient complete other types of activities, such as exercise, by having others accompany them or encourage them throughout the process. While BA can help the patient create new relationships and strengthen their social support network, its baseline state could affect how quickly an improvement in mood is achieved. Examining this factor could further elucidate why patients differ in their treatment response.

The availability of environmental reinforcement is clearly relevant when it comes to treating depression with BA during the COVID-19 pandemic. Social distancing measures have imposed unprecedented limits on everyday activities and significantly constrained the types of goals patients can set in BA. Testing the efficacy of BA in this context would show how important the quality of the environment is and whether BA can help patients find new opportunities for reinforcement or whether it reduces in efficacy in these unusual circumstances.

Reward processing

Another factor that may be a relevant mechanism of BA relates to how patients process rewarding stimuli. One of the core symptoms of depression is anhedonia, the decreased ability to experience pleasure. It is associated with poor outcome in standard antidepressant treatment (Uher et al., 2011), possibly due to unresolved deficits in several reward-related neural circuits (Admon and Pizzagalli, 2015). Anhedonia is increasingly recognised as an important treatment target as it is linked to suicidal ideation independently of depression (Ducasse et al., 2018) and predicts impaired psychosocial functioning (Vinckier et al., 2020).

A recent systematic review and meta-analysis by Halahakoon et al. (2020) examined the association between depression and several cognitive subcomponents of reward processing. The largest impairment was found in reward-related decision-making, indicating impaired cost-benefit evaluations when contemplating different actions. A small to medium deficit was also found in reinforcement learning, a process through which individuals learn from feedback to adjust their future behaviour in order to maximise reward. Similar findings were subsequently reported in a review by Kieslich et al. (2021) who recommended the examination of different reward processing components in the context of anhedonia treatment.

Reward learning and anhedonia have not been extensively investigated in relation to BA treatment for depression, even though it could be expected to play an important role. Since the BA model postulates that repeated exposure to pleasurable activities should lead to improved mood and subsequently increased reward-seeking behaviour, patient's capacity for decision-making and reinforcement learning could moderate their engagement with treatment and thus affect its success. Alternatively, the process of BA, particularly the psychoeducation on the relationship between low mood and low activity, could encourage the patient to make more rewarding decisions, and the process of regular reflection on the benefit of different activities could enhance reinforcement learning. In this way, BA could impact these cognitive processes directly and act as a mediator of treatment outcome.

The current body of evidence has not investigated the effects of BA on reward-related decision making and reinforcement learning. Dichter et al. (2009) reported that a course of BA may lead to increased reward processing activity during a gambling paradigm, specifically by increasing dorsal striatum activity during reward anticipation. However, other studies on reward processing have had methodological weaknesses. Walsh et al. (2019) found that a decrease in anhedonia after BA was predicted by decreased connectivity between the left middle frontal gyrus and right temporoparietal regions during exposure to positive stimuli, but this was explored without a control group. Russo et al. (2018) found decreased anhedonia after a course of BA combined with a TMS intervention, but only in a small sample and again without appropriate controls. It would therefore be important to investigate the effect of BA on the aforementioned components of reward processing in a larger sample with adequate methodological design.

In recent years, deficits in decision-making and reinforcement learning have been assessed using the probabilistic instrumental learning task, which tests the ability to adjust one's behaviour based on long-term rates of reward and punishment associated with different stimuli. Unmedicated depressed patients have been found not to develop a bias towards more rewarding stimuli on this task (Pizzagalli et al., 2008) and this level of reward responsiveness has been found to negatively correlate with self-reported anhedonia, as well as predicting it at one-month follow-up (Pizzagalli et al., 2005). However, to our knowledge, this task has not been used in the context of BA treatment before. Examining whether this intervention modifies patients' reward learning or whether baseline variability in this domain affects treatment outcome would help further elucidate which patients may be most suitable for this treatment.

Physical activity

The intensity of physical activity experienced during BA could also be an important variable influencing the treatment outcome. Depression is known to be linked with increased sedentary behaviour (Teychenne et al., 2010), which may result from symptoms of fatigue as well as anhedonia and lack of motivation. Increased levels of physical activity have been extensively linked to improved affect (Mata et al., 2012; Wichers et al., 2012), therefore including physical activity goals as part of BA could be particularly effective for symptom relief. Moreover, increased motor activity has been linked to improved incentive drive one hour after exercising (Bewernick et al., 2017), which could further help the patient motivate themselves to do other activity goals. Baseline levels of physical activity could also influence patients' ability to complete goals and the subsequent success of the intervention.

To our knowledge, no studies have examined physical activity levels in the context of BA treatment using objective actigraphic measures. A systematic review of 19 studies using actigraphic devices in depression (Burton et al., 2013) concluded that these measures can provide valuable information about patients' symptom variability. They also reported that effective depression treatments can lead to objective increases in patients' daytime activity, possibly providing an objective marker of treatment success when it comes to alleviate the fatigue-related symptoms of the disorder. Using actigraphy devices in the context of BA treatment could help us evaluate whether increases in physical activity are important in treatment success and whether specific emphasis should be placed on setting such goals.

Non-specific factors in psychological interventions

Lastly, an important mechanism of BA may also concern the so called “non-specific factors” of psychotherapy (also known as the common factors or the psychological placebo effect). These include factors such as the therapeutic alliance, the therapeutic context and the patient’s belief in the treatment rationale (Ahn and Wampold, 2001). Some authors have argued that these factors may be more important than specific treatment techniques (Messer and Wampold, 2006), such as the act of activity monitoring and goal setting in the case of BA. However, these arguments have largely resulted from studies showing that different types of psychotherapy can have equivalent effects. But this could be because many types of specific techniques can be similarly effective, rather than not being effective at all.

The use of active control groups, which provide patients with the common factors but not with any specific techniques, could help resolve this debate. However, these control groups aren’t commonly used in psychotherapy research. For example, three recent meta-analyses on BA efficacy showed that the majority of studies use waitlist or treatment-as-usual control groups (Cuijpers et al., 2007; Mazzucchelli et al., 2009; Ekers et al., 2014). This all-or-nothing approach does not allow for a component analysis differentiating the efficacy of common and specific factors.

Together with measures that may serve as objective markers of efficacy, the comparison of both an active and a passive control group with BA could also help control for the demand effects that are likely to occur when using self-report instruments. Presuming that demand effects would be equivalent in BA and its active control group, finding that BA shows stronger effects on objective markers, such as measures of emotional cognition or actigraphy, would strengthen the evidence for the usefulness of specific techniques of BA in depression treatment. This would help us better understand the mechanism of this intervention and subsequently design for what participants and in what contexts it can be most effective.

Thesis objectives

The first experiment of this thesis aimed to examine the mood effects as well as potential mechanisms of BA as a treatment for depression during the COVID-19 pandemic when delivered remotely by non-specialists after a short training. Chapter 2 compares the results of a BA group in comparison to a passive control group on several self-report measures of efficacy as well as examining possible mechanisms relating to levels of environmental reinforcement. Chapter 3 shows the analysis of emotional cognition and reward processing data collected during the COVID-19 experiment to investigate whether they may be plausible mechanisms of BA action and whether they could serve as early objective markers of efficacy.

Chapter 4 introduces a second experiment which compares the effects of the BA treatment on self-report measure of mood to both an active and a passive control group in order to control for the effect of common factors and to investigate the role of specific BA components. Chapter 5 compares the three groups using objective measures of emotional cognition and reward processing, while Chapter 6 uses actigraphy to investigate physical activity as a possible treatment mechanism. Chapter 7 finally gathers all findings to discuss their wider implications and to suggest further areas of research.

It is worth noting that the original plans of this DPhil were significantly disrupted by the COVID-19 pandemic, as in-person data collection was not possible between March 2020 and May 2021. Instead of the originally planned experiment testing a combination of BA with antidepressant drugs, we devised the online experiment instead to examine BA and emotional cognition in the context of the pandemic. Due to the disruptive conditions and the short timeframe for running this online experiment, it was not pre-registered. The second experiment testing BA in comparison to an active and a passive control group remained as intended, albeit the intended sample size had to be reduced due to the effects of the pandemic on recruitment (see Methods for details of impact). This experiment was pre-registered on Clinical Trials (<https://clinicaltrials.gov/ct2/show/NCT03995186>).

Chapter 2: Examining the effects of behavioural activation on mood in comparison to a passive control group during the COVID-19 pandemic

Introduction

The first experiment of this thesis examined the effects of a short course of online behavioural activation delivered by non-specialists in comparison to a passive control group during the COVID-19 pandemic. As detailed in Chapter 1, the pandemic has had a detrimental effect on public mental health, increasing rates of depression, anxiety, burnout, substance use, post-traumatic stress disorder and eating disorders worldwide (Ettman et al., 2020; Winkler et al., 2020; Clemente-Suarez et al., 2021) as well as often decreasing access to treatment (Aragona et al., 2020; Purrington & Beail, 2021). It is therefore of crucial importance to examine what interventions could help prevent and treat mental health problems during societal crisis periods like a pandemic.

Behavioural activation (BA) is a simple yet effective treatment for depression that helps patients examine how they spend their time, what activities they avoid because of their low mood and how they could reconnect with those conducive to their emotional wellbeing (such as productive work, hobbies, or socialising). In addition to treating depression, evidence suggests that BA may also have beneficial effects on substance use (Martinez-Vispo et al., 2018; Magidson et al., 2020) and anxiety (Hopko et al., 2016). It can be effectively delivered in a variety of formats, such as through an online self-help course with therapist support (Carlbring et al., 2013) or as a standalone smartphone application (Ly et al., 2014). This flexibility could make it particularly suitable for a pandemic context where social distancing policies necessitate remote forms of treatment.

Importantly, evidence suggests that BA does not necessarily have to be administered by extensively trained specialists. Studies show that it can be effectively delivered through lay counsellors without any mental health qualifications; and what's more, the length of training required has ranged from just a few days (Richards et al., 2017) to a few weeks (Patel et al., 2017). This suggests that the simplicity of BA makes it uniquely amenable to wide dissemination through lay practitioners that could be trained much faster than traditional mental health

specialists. This could help to supplement treatment provision in places where mental health services are oversubscribed, or - as in many parts of the world - barely existent at all (WHO, 2008).

While the number of people with mental health problems has increased during the COVID-19 pandemic, available healthcare resources have rapidly decreased at the same time. As a result, it is also important to examine what could be the smallest "dose" of BA that would provide lasting relief from mental health problems. The current body of research has not provided a clear answer to this question. While the standard BA protocol takes place over at least 8 weeks or more (Jacobson et al., 1996; Dimidjian et al., 2006; Richards et al., 2016), other studies have found effective short-term help just from a single BA session (Gawrysiak et al., 2009; Nasrin et al., 2017). However, since an 8-session protocol is associated with a 50% relapse rate within the following year (Ali et al., 2017), a single session may not be likely to cause robust long-term improvement. Nonetheless, considering the low availability of mental health resources around the world, a compromise protocol of 4 weeks would be worth examining, both in terms of immediate as well as sustained follow-up effects.

While all the above factors suggest BA could be suitable for dissemination during a pandemic, there are also reasons to be apprehensive about its efficacy in this context. Importantly, it is unclear whether BA can still be effective when options for activities are significantly constrained due to lockdowns and other social distancing measures. The COVID-19 pandemic has led to an unprecedented reduction of access to professional as well as recreational opportunities that many find to be a source of meaning and enjoyment. Considering the inability to take part in most travel, sport, cultural and hospitality activities, patients may not be able to practice the behaviours that would be most in line with their interests and values as the BA model recommends. Having to find alternative activity options, which may not be as intrinsically fulfilling, could cause the treatment to take longer than usual or to be generally less effective. In this context, it is essential to experimentally examine whether BA can remain helpful in this unique set of circumstances.

In this randomised controlled study, we investigated whether a 4-week course of remote non-specialist administered BA can reduce mild to moderate depression and increase activation during a period of social distancing restrictions. We hypothesised that, in comparison to a

passive control group, BA would lead to a significant decrease in depression symptoms by the end of the 4 weeks. To examine sustained effects, we also collected data at one-month follow-up, hypothesising that the effect would persist.

As secondary outcomes for exploratory analysis, we also monitored the effects on state and trait anxiety, anhedonia, automatic negative thoughts, social support, and COVID-related lifestyle disruption. The main aims of the study were to test the efficacy of the intervention as well as examining its possible mechanisms of effect. The mechanistic analysis, which examined participants' emotional cognition using an established battery of tasks, will be presented in Chapter 3.

The study was published in *Psychological Medicine* (Ruzickova et al., 2021), see References.

Methods

Participants

Seventy participants were recruited online via advertising on social media, University of Oxford's Department of Psychiatry website, the MQ charity website and Oxford's Daily Info platform. Two participants dropped out after screening and the final sample included 57 women and 11 men (see Figure 2 below for details of numbers at enrolment, allocation and follow-up). All participants scored between 10 and 28 on the BDI-2 at baseline and were in the age range of 18 to 65. Exclusion criteria included not being from the UK, not being able to use Microsoft Teams, past participation in other studies using the Emotional Testing Battery, frequent use of recreational drugs (once a month or more), undergoing any other psychological treatment, reporting suicidal thoughts, having a past or present diagnosis of psychosis or bipolar disorder, or having a current diagnosis of an eating disorder, borderline personality disorder, substance abuse, OCD or PTSD (as assessed by a structured clinical interview for the DSM-V). Participants could be on antidepressant medication, but any other medication was evaluated on a case-by-case basis by a study medic to assess the likelihood of interfering with the study measures. Recruitment was carried out between May and October 2020.

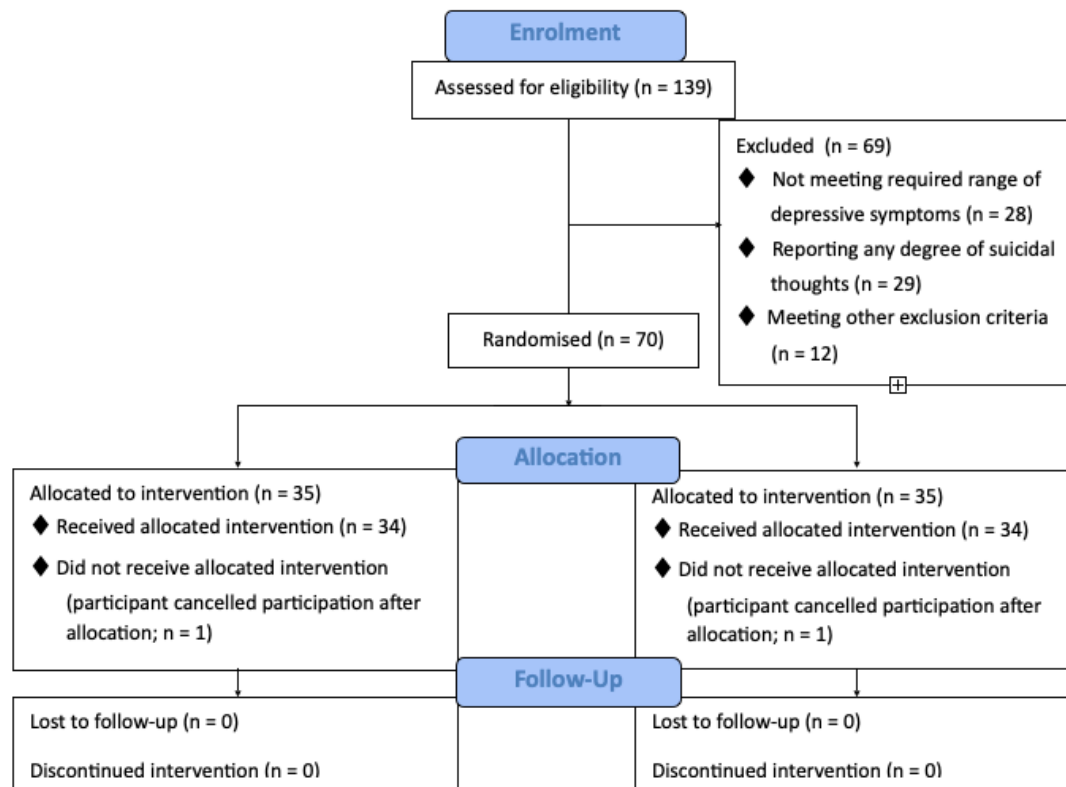


Figure 2. A flowchart showing the numbers of participants assessed for eligibility through the SCID interview as well as the numbers of participants excluded, randomised and followed up.

Power calculation

An a priori power calculation showed that we needed 36 participants per group in order to detect a moderate-to-large effect size (Cohen's d of 0.6) for a difference between two means with 80% power. This should allow us to detect a between-group difference in depression symptoms, our primary outcome measure, as a previous meta-analysis by Ekers et al. (2014) found BA to have a large effect size in reducing depressive symptoms when compared to passive control groups (Hedge's g of 0.75).

Procedure

All applicants to the study first filled in a pre-screening questionnaire, where they were asked to complete an electronic consent form and to read the participant information sheet. They were asked about their level of depressive symptoms using the Beck Depression Inventory (BDI-2; excluding the suicidality question for ethical reasons) and they were assessed on the main inclusion and exclusion criteria of the study. Potentially eligible participants then proceeded onto phone screening with a research assistant, who asked them about their symptoms in more

detail as well as assessing any current or past mental health treatments, physical health status and family history of mental illness. Finally, participants who appeared eligible were invited for a video call with one of the study researchers who carried out the Structured Clinical Interview (SCID) for DSM-5 with them to confirm they weren't suffering from any of the excluded conditions.

Fig. 3 below shows the study schedule for eligible participants, who were randomly allocated to either the BA intervention or to a control group, stratified by sex. Questionnaire measures (BDI-2, BADS, STAI, SHAPS, ATQ, MSPSS as well as all COVID-related questionnaires) were administered remotely at baseline (week 0), halfway through the intervention (week 2) and at the end (week 4) using the Qualtrics and the Gorilla testing platforms. Before each of these data collection sessions, participants were reminded that they should pay full attention and should not be distracted by their surroundings. A link to the daily Mood Zoom questionnaire was texted to participants every evening using the Text Marketer platform and participants were asked to reflect on the whole day, rather than just the present moment, when filling it in. The BDI-2, BADS, STAI and all COVID-related questionnaires were also administered at one-month follow-up (week 8).

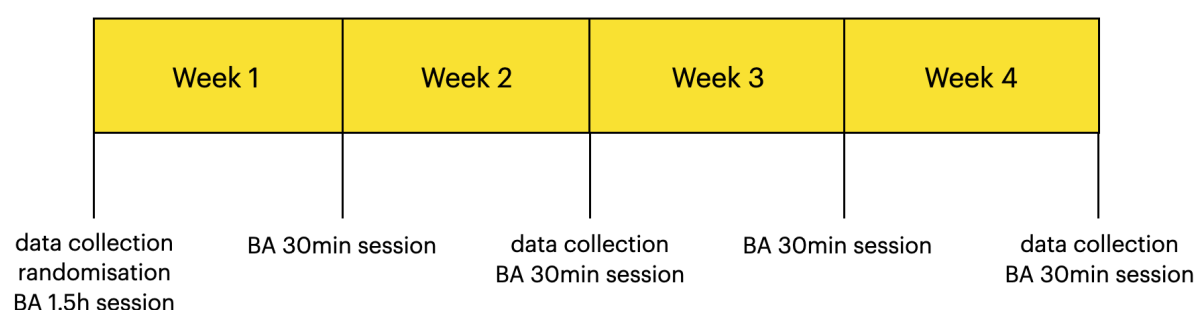


Figure 3. Study schedule showing the timing of data collections (baseline, midway and final) as well as intervention sessions applying to the BA group only.

Intervention

We followed an established BA programme 'Get Active, Feel Good' (Farrand, Taylor, Greaves, & Pentecost, 2013) routinely used in the UK Improving Access to Psychological Therapies (IAPT) services. Our intervention lasted 4 weeks and it was administered by two psychology researchers who were trained and supervised by a Psychological Wellbeing Practitioner (PWP) from a local IAPT service. Each researcher received 15 hours of training prior to the study with special

emphasis on adapting the programme for remote administration. The training included skills related to psycho-education on the BA model of depression, activity monitoring, planning and problem-solving, as well as general interpersonal skills such as empathy and encouragement. After the training, each researcher submitted recordings of their practice sessions for evaluation and approval. Once the study started, ongoing support from the PWP was available.

Participants had an initial introductory session lasting 1.5 hours following the official BA programme booklet. The session started with psychoeducation on the possible causes of low mood, the relationship between low mood and low activity and the general CBT model of how thoughts, behaviours, and physical sensations bidirectionally affect each other. Participants were then asked to set three main behavioural goals that they would like to work on during the study and together with the practitioner these goals were formulated to be specific, measurable, and achievable for the participant. Goals were chosen based on participants' own interests and priority. Afterwards, participants carried out retrospective activity monitoring of the week that had just gone by to help the practitioner estimate their baseline level of activation. The activities in their baseline week were divided into three categories for routine, pleasurable and necessary behaviours and participants were instructed to aim for a balance of all three types in their life. Participants were asked about other activities they would like to introduce beyond the three main goals that were set out previously. All the activities were then ordered according to subjective difficulty and emphasis was placed on starting with the easiest tasks.

Finally, small and measurable activity goals were set out for the first week, aiming for only a small increase in activation above the participants' baseline levels. Participants were asked to follow the "action first" principle (also called "outside in"), which reminded them to take actions before they feel motivated and to try to do the behaviours despite negative thoughts and emotions. When struggling with motivation, they were told to use the "5-minute rule" instructing them to try just five minutes of any given activity. They were also recommended not to increase their activation beyond the plan made in the session to avoid a "boom and bust" scenario, where exhaustion or disappointment could worsen their depression. Participants were given an electronic template in which they could monitor their activities while also having the option of keeping their notes on paper. They were also informed that full participation in every session of the intervention and weekly submission of activity and goal monitoring was required to receive the full study reimbursement.

After the first introductory session, participants had weekly 30min check-in meetings with the practitioner (4 check-ins in total) to examine their activity monitoring log, to discuss which activities they accomplished and to problem solve any issues. They were reminded of the BA principles and a new activity plan was prepared for the following week. If previous week's plan was carried out successfully, a slight increase in difficulty or frequency of activities was encouraged. At the final meeting, participants were asked to reflect on all the progress they made and on the strategies that they found most helpful. A relapse-prevention plan was written up to help remind them of these techniques so they could continue using them on their own beyond the context of the study. In total, participants spent 3.5 hours on average in BA meetings with the practitioner. All meetings were carried out remotely through video calls via Microsoft Teams.

The control group did not receive any intervention but received materials about BA at the end of the study, which included the intervention booklet, a video about a patient who is undergoing BA as well as resources on where to find BA in their area. All participants were compensated with £80 for their full participation together with any amount won in the Probabilistic and Instrumental Learning Task (see Chapter 3).

Questionnaire measures

We measured depression using the full version of the Beck Depression Inventory 2 (BDI-2; Beck, Steer, & Brown, 1996), which is one of the most widely used instruments for evaluating depressive symptoms (McDowell, 2006). It uses a 4-point Likert scale to assess 21 components of depression, such as feelings of sadness or frequency of crying. Participants are asked to assess their experience over the past two weeks. As opposed to the original version (Beck et al. 1961), BDI-2 contains new items on difficulty concentrating, agitation, worthlessness, and energy loss, as well as adding that both an increase and a decrease in appetite and sleep may be possible symptoms of depression. The ranges of possible scores on this questionnaire have been classified into minimal (0-13), mild (14-19), moderate (20-28) and severe depression (29-63). A large naturalistic study in psychiatric hospitals reported that a score of 12 or lower best corresponded to clinician-evaluated remission based on global impression of patients' functioning (Riedel et al., 2010). BDI-2 has been reviewed to have good psychometric properties

across a variety of populations (Wang and Gorenstein, 2013). Considering the extensive body of research that has validated this questionnaire, we chose this scale as our primary outcome measure of mood.

Anxiety was measured using the State-Trait Anxiety Inventory (STAI; Spielberger et al., 1970), which uses a 4-point Likert scale to evaluate 40 statements assessing both momentary (state) anxiety as well as a more stable personality tendency towards experiencing anxiety (trait). It has been found to be reliable and valid in both non-clinical (Metzger, 1976) and clinical populations (Oei et al., 1990).

Depression-related activation was assessed using the long form of the Behavioural Activation for Depression Scale (BADSD; Kanter et al., 2007). It asks participants to rate 25 statements such as “I stayed in bed for too long even though I had things to do” on a 6-point Likert scale to indicate how often these were true for them over the past week. It uses 4 subscales to assess work and school impairment, social impairment, avoidance and rumination and goal-directed behaviour. This scale has been found to have good psychometric properties when tested on an undergraduate sample (Kanter et al., 2007) as well as a community sample with symptoms of depression (Kanter et al., 2009).

Social support was measured using the Multidimensional Scale of Perceived Social Support (MSPSS; Zimet et al., 1988). This instrument is designed to measure level of support on three subscales corresponding to participants’ friends, family and a significant other. Twelve statements such as “I can count on my friends when things go wrong” are rated on a 7-point Likert scale. It has been found to have good psychometric properties on several measures (Zimet et al., 1988; Zimet et al., 1990).

Anhedonia was assessed using the Snaith–Hamilton Pleasure Scale (SHAPS; Snaith et al., 1995), which is a self-report measure of 16 items assessing hedonic capacity on a 4-point Likert scale by asking participants to imagine and evaluate various scenarios such as “I would enjoy my favourite meal”. It has previously been found to have good psychometric properties in a large sample of adult outpatients suffering from major depressive disorder (Nakonezny et al., 2010).

Automatic negative thoughts were measured using the short form of the Automatic Thoughts Questionnaire (ATQ; Hollon and Kendall, 1980). On a 5-point Likert scale, participants are instructed to indicate how often a particular negative thought, such as "I'm so disappointed in myself" occurred to them over the last week, evaluating 15 thoughts in total. This shortened form of the scale has been confirmed to have acceptable psychometric properties (Netemeyer et al., 2002).

We sent out electronic daily mood questionnaires to participants based on the Mood Zoom format (Tsanas et al., 2016), asking participants to rate on a 7-point Likert scale the extent to which they felt happy, energetic, anxious, angry, tired and irritable that day. Participants' data were averaged across each of the four weeks of their participation of the study; a week had to have at least 3 days of data provided to be included in the analysis. If the BA group participants had any delay between the first data collection and the start of their intervention, only the data since the start of the intervention was included. As this scale has primarily been used in research on bipolar disorder and has not been extensively validated in depressed samples, we included it as a secondary outcome measure only.

We measured COVID-19-related lifestyle disruption using several questionnaires developed in our department (see Appendix). The Current Stressors Questionnaire asked about the extent to which participants have been stressed about 18 life domains (including finances, getting medication or job security) over the past week on a four-point Likert scale. The COVID Anxiety Questionnaire asked participants to evaluate nine statements (such as 'I am worried that I will catch COVID-19') on a five-point Likert scale. The Current Disruption Questionnaire asked participants to evaluate how disrupted six different life areas, such as work, friendship, or leisure time, had been on a four-point Likert scale. Participants were also asked to report if they were in a high-risk COVID-19 category and whether they had previously tested positive for COVID-19.

We also asked participants about several aspects of their living and work situation in relation to the pandemic, including their current isolation status, number of people per household, number

of rooms per house, level of access to outside space, whether they are key workers¹, their work status, how their work hours have changed due to the pandemic, work location and whether they feel able to balance work and other responsibilities. We also asked them to estimate how much exercise they have done and how much time they spent outside over the past week, whether they are a morning or an evening chronotype, how well they slept over the past week and how their sleep has changed because of the pandemic.

Finally, all BA participants' experience of the intervention was assessed in a final feedback questionnaire, asking them to answer seven questions such as "Did you find the intervention helpful?" using a sliding scale from 0 to 100 points.

Statistical analysis

All questionnaire measures were analysed using two-way mixed ANOVAs with group as a between-subject factor (BA and control) and time as a within-subject factor (baseline, week 2, week 4 and follow-up at week 8 where relevant). Mood Zoom data were analysed in a three-way mixed ANOVA with the additional factor of valence, which had three levels corresponding to positive (happy, energetic), negative (sad, anxious) and irritable (irritable, angry) emotions. Significant interactions were followed up with simple main effect analyses; independent t tests were used for simple main effects of group and pairwise comparisons, separated by group, were used for simple main effects of time.

Normality was assessed using the Shapiro–Wilk test. Outliers were assessed by the inspection of box plots and studentised residuals; they were removed if they were 3 box-lengths away from the edge of the box or if the studentised residual was greater than ± 3 standard deviations. Mauchly's test was used to assess the assumption of sphericity; if sphericity was violated, Greenhouse–Geisser correction was applied. Box's test was used to assess the equality of covariance.

¹ As defined by Department of Education (2020). *Children of critical workers and vulnerable children who can access schools or educational settings*. (n.d.). [GOV.UK](https://www.gov.uk/government/publications/coronavirus-COVID-19-maintaining-educational-provision/guidance-for-schools-colleges-and-local-authorities-on-maintaining-educational-provision). Retrieved 25 March 2021, from <https://www.gov.uk/government/publications/coronavirus-COVID-19-maintaining-educational-provision/guidance-for-schools-colleges-and-local-authorities-on-maintaining-educational-provision>

All data were processed and analysed using R and SPSS software. For a sensitivity analysis, all analyses were compared with and without the 13 participants who were taking antidepressants to investigate whether any effects may have been driven by the concurrent pharmacological treatment.

Results

Demographic, clinical and COVID-related baseline information

The two groups were well matched in terms of the main demographic and clinical characteristics at baseline (see Table 1 below).

Table 1. Demographic, clinical and COVID-related baseline characteristics

Variable	BA group (n = 34)	Control group (n = 34)
Age: years (mean, SD)	32.38 (10.92)	30.79 (11.27)
Full time education: years (mean, SD)	16.29 (3.23)	15.88 (2.29)
Highest education level attained	26.5% A-level/GCSE 38.2% Undergraduate or professional qualification 35.3% Postgraduate	23.5% A-level/GCSE 50% Undergraduate or professional qualification 26.5% Postgraduate
Race	76.5% white 23.5% non-white	96.9% white 3.1% non-white
Current antidepressant treatment	14.7% yes 85.3% no	23.5% yes 76.5% no
Current MDE	48.5% yes 51.5% no	41.2% yes 58.8% no
Current PDD	6.1% yes 93.9% no	5.9% yes 94.1% no
Current GAD	6.1% yes 93.9% no	8.8% yes 91.2% no
Current panic disorder	6.1% yes 93.9% no	2.9% yes 97.1% no

Current social anxiety disorder	0% yes 100% no	2.9% yes 97.1% no
Work status in the past week	25.7% full time 25.7% part time 48.6% unable to work	37% full time 20% part time 43% unable to work
Percentage key workers as defined by the UK government	14% yes 86% no	17% yes 83% no
Baseline COVID-19 risk	0% yes 100% no	6% yes 94% no
Baseline COVID-19 symptoms	0% yes 100% no	3% yes 97% no
Baseline COVID-19 diagnosis	3% suspected 97% no	20% suspected 80% no
Current isolation status	8.8% leaving the house as normal 41.2% social distancing where possible 44.1% only leaving the house for essential responsibilities 5.9% shielding with access to outside space 0% shielding with no access	2.9% leaving the house as normal 52.9% social distancing where possible 41.2% only leaving the house for essential responsibilities 0% shielding with access to outside space 2.9% shielding with no access to
Current number of people per household (mean, SD)	2.79 (1.18)	2.44 (1.12)
Current number of rooms in household (mean, SD)	5.00 (2.62)	4.59 (2.12)
Outdoor access	88.2% yes 11.8 % no	85.3% yes 14.7% no
Time spent exercising in the past week (self-report)	51% less than 30min 49 % more than 30min	77% less than 30min 23% more than 30min
Time spent outside in the past week (self-report)	80% less than 2h 20% more than 2h	77% less than 2h 23% more than 2h

Currently categorised as high- risk of COVID-19 (due to existing medical conditions or age)	0% yes 100% no	6% yes 94% no
Family members, friends or housemates currently categorised as risk of COVID-19	23% yes 77% no	20% yes 80% no
Experienced symptoms of COVID-19 in the past 2 weeks	0% yes 100% no	3% yes 97% no
Baseline COVID-related stress (score out of 72; mean, SD)	32.40 (6.44)	31.57 (7.86)
Baseline COVID-related anxiety (score out of 45; mean, SD)	32.31 (4.36)	30.63 (4.66)
Baseline COVID-related lifestyle disruption (score out of 24; mean, SD)	17.40 (3.35)	16.91 (3.82)

Diagnostic categories were assessed using the Structured Clinical Interview for DSM-5. MDE = major depressive episodes, PDD = persistent depressive disorder, GAD = generalised anxiety disorder

Rating of intervention

The majority of BA participants evaluated the intervention positively on all the questions that were asked in the feedback survey (see Table 2 below). Notably, when asked to evaluate how helpful the intervention was, the average participant rating was 82 out of 100. This indicates that the intervention was appropriately delivered by the two non-specialists.

Table 2. Participants' self-reported rating of the intervention (mean rating on 0-100 scale)

Question	Mean, SD
Overall, did you find the intervention helpful?	81.58 (17.61)
Did the intervention match your ideas of what helps people with low mood?	72.97 (21.29)
Was the intervention well explained?	95.97 (6.90)
Did you get on well with your BA practitioner?	97.35 (6.14)
Was your BA practitioner supportive and empathetic?	97.68 (5.49)

Do you think you'll continue using the strategies learnt in this intervention?	86.39 (19.77)
Would you recommend this intervention to other people experiencing low mood during the lockdown?	92.23 (12.75)

Questionnaire measures

Baseline descriptive statistics for all questionnaires can be found in the Appendix (Table 51).

Depression

There was a significant time by group interaction for depression scores [$F(3,189) = 7.75$, $p < 0.001$, $\eta^2 = 0.11$], see Fig.4 below.

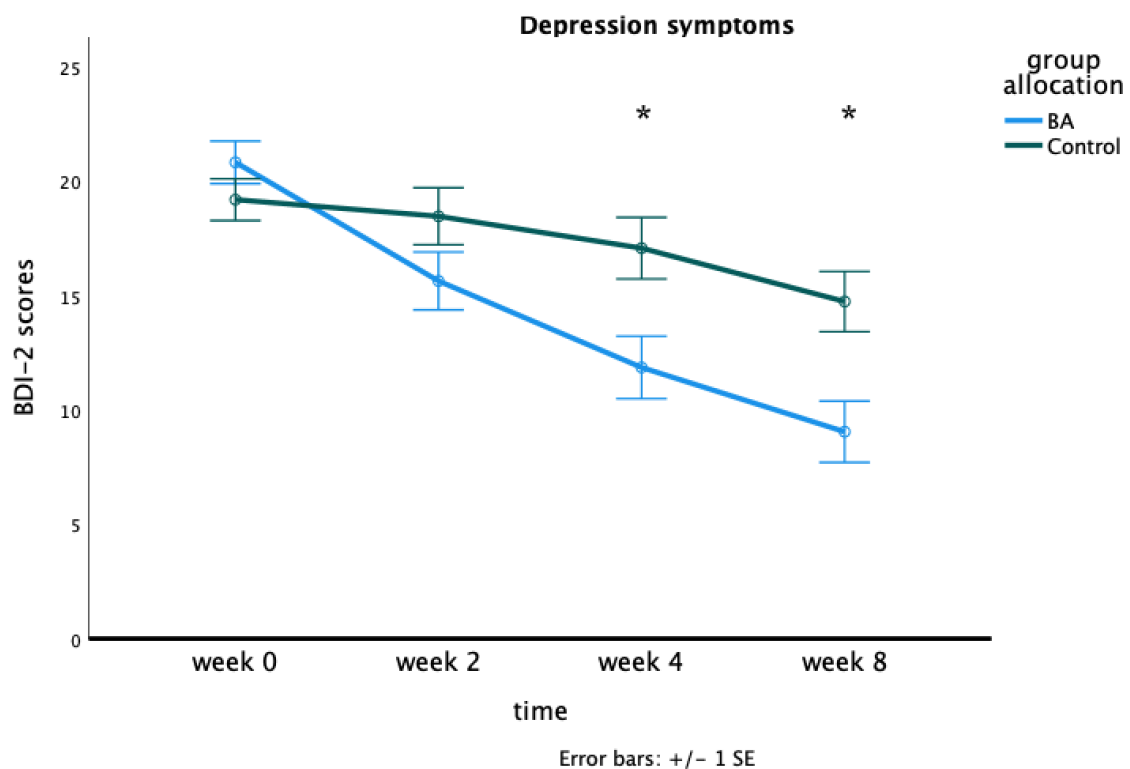


Figure 4. Depression scores as measured by the BDI-2 from baseline (week 0), to midway (week 2), end of study (week 4) and one-month follow-up (week 8). Asterisks indicate $p > .05$.

Follow-up analyses revealed that this interaction was driven by a significant difference between groups at the end of the intervention, with the BA group showing lower depression scores than the control group (see Table 3 for detailed results). The significant difference persisted at 1-

month follow-up, again with depression symptoms significantly lower in the BA group than the control group.

Table 3. Detailed statistics for post-hoc t-tests comparing the behavioural activation (BA) and control group (C) in depression scores as measured by the BDI. One asterisk indicates $p < .05$.

BA vs control	T-test statistics	Group M and SD
Week 0	$t(66) = 0.75, p = .46, d = 0.18$	BA: $M = 20.38, SD = 5.54$ C: $M = 19.41, SD = 5.19$
Week 2	$t(64) = -1.63, p = .11, d = 0.40$	BA: $M = 15.64, SD = 7.12$ C: $M = 18.48, SD = 7.08$
Week 4 *	$t(65) = -2.68, p = .009, d = 0.66$	BA: $M = 12.18, SD = 7.61$ C: $M = 17.24, SD = 7.83$
Week 8 *	$t(65) = -3.00, p = .004, d = 0.74$	BA: $M = 9.45, SD = 6.92$ C: $M = 15.21, SD = 8.64$

Activation

There was also a significant time \times group interaction for self-reported activation scores, as measured by the BADS [$F(3,183) = 4.75, p = .003, \eta^2 = 0.07$], see Fig.5 below.

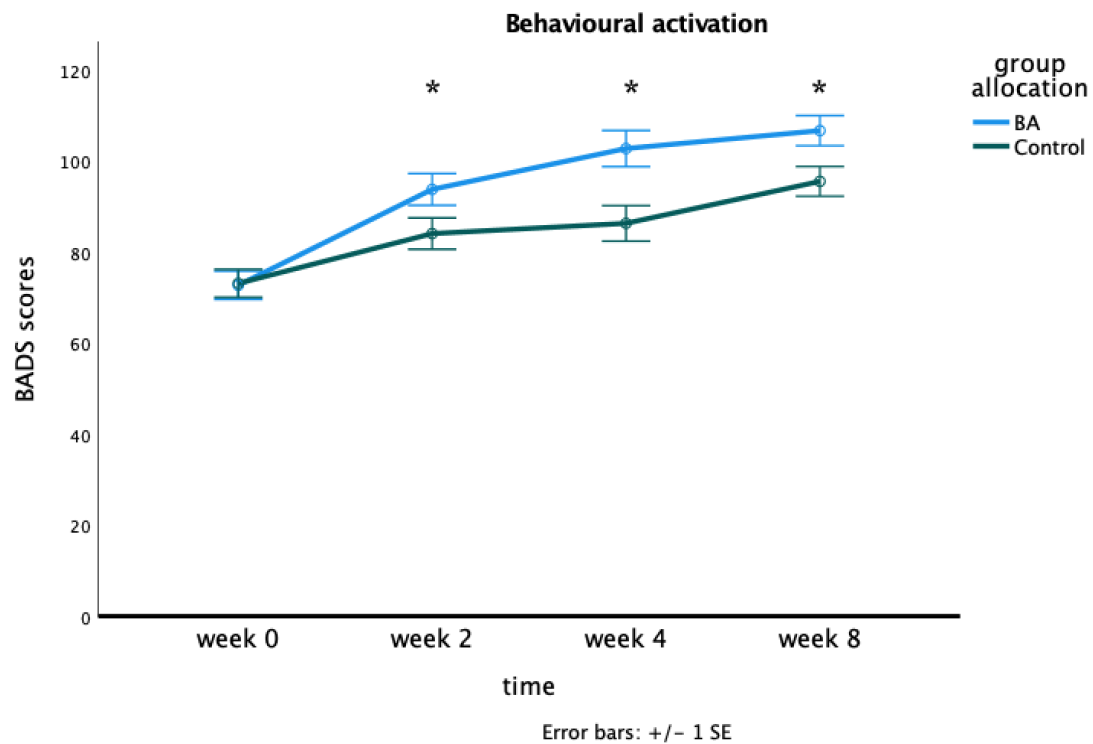


Figure 5. Behavioural activation scores as measured by the BADS from baseline (week 0), to midway (week 2), end of study (week 4) and one-month follow-up (week 8). Error bars show ± 1 standard error. Asterisks indicate $p < .05$.

Follow-up analyses revealed significant differences between the groups midway through the intervention at week 2, with the BA group showing higher activation than the control group (see Table 4 for detailed results). Further difference was found at the end of intervention at week 4, with the same trend for the BA group being higher in activation than the control group. Finally, differences persisted at 1-month follow-up with the BA group again showing higher activation scores than control group.

Table 4. Detailed statistics for post-hoc t-tests comparing the behavioural activation (BA) and control group (C) in activation scores as measured by the BADS. One asterisk indicates $p < .05$.

BA vs control	T-test statistics	Group M and SD
Week 0	$t(64) = -0.14, p = .89, d = 0.04$	BA: $M = 72.36, SD = 16.24$ C: $M = 72.97, SD = 18.57$
Week 2 *	$t(64) = 2.12, p = .04, d = 0.52$	BA: $M = 93.97, SD = 19.20$ C: $M = 84.06, SD = 18.85$
Week 4 *	$t(65) = 2.99, p = .004, d = 0.73$	BA: $M = 101.42, SD = 21.59$ C: $M = 85.35, SD = 22.36$

Week 8 *	$t(65) = 2.16, p = .04, d = 0.53$	BA: $M = 104.61, SD = 20.40$ C: $M = 93.62, SD = 21.27$
----------	-----------------------------------	--

Anhedonia

Our data on anhedonia were moderately positively skewed and square root transformation was applied to prevent violations of ANOVA assumptions. The measure showed a significant time \times group interaction [$F(1.61, 102.75) = 4.07, p = 0.03, \eta^2 = 0.06$], see Figure 6.

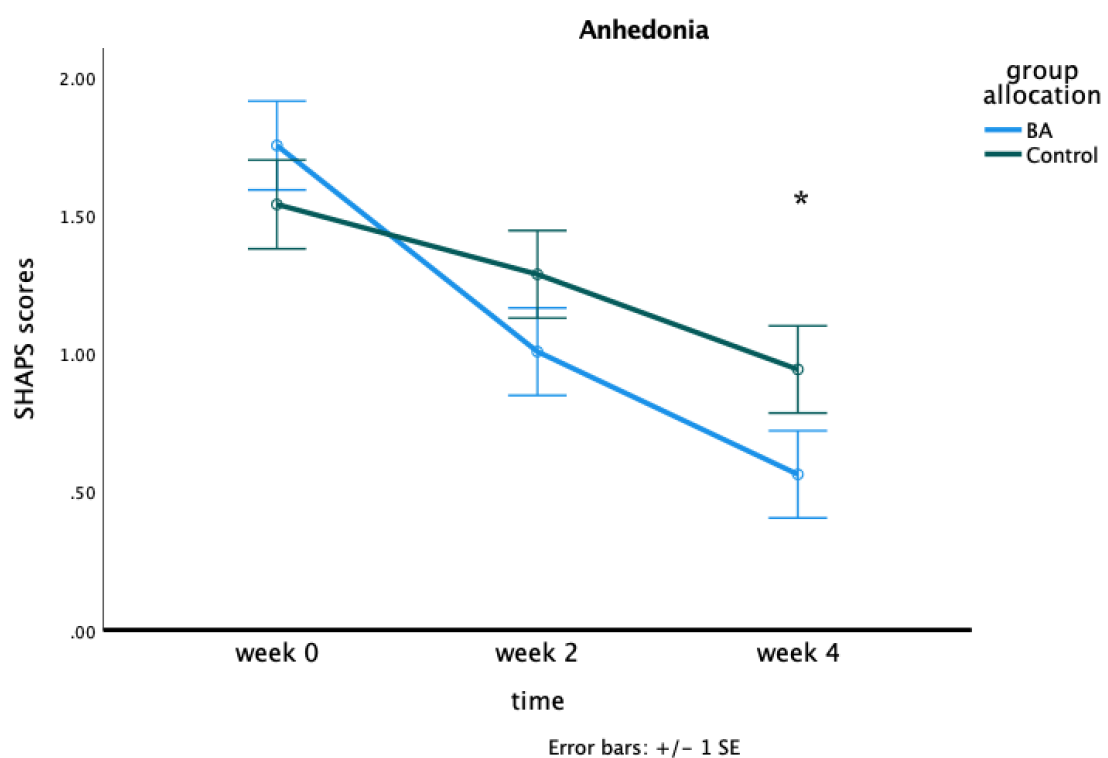


Figure 6. Anhedonia scores (after square root transformation) as measured by the SHAPS from baseline (week 0), to midway (week 2) and end of study (week 4). Error bars show ± 1 standard error. Asterisks indicate $p < .05$.

This was driven by a significant difference between groups at final time point, with the BA group showing lower anhedonia than the control group, see Table 5 below for detailed results.

Table 5. Detailed statistics for post-hoc t-tests comparing the behavioural activation (BA) and control group (C) in anhedonia scores as measured by the SHAPS. One asterisk indicates $p < .05$.

BA vs control	T-test statistics	Group M and SD
Week 0	$t(66) = 0.94, p = .35, d = 0.23$	BA: $M = 1.75, SD = 1.05$ C: $M = 1.54, SD = 0.80$
Week 2	$t(66) = -1.25, p = .22, d = 0.30$	BA: $M = 1.01, SD = 0.95$ C: $M = 1.29, SD = 0.90$
Week 4 *	$t(64) = -2.32, p = 0.02, d = 0.58$	BA: $M = 0.45, SD = 0.71$ C: $M = 0.94, SD = 1.0$

Anxiety

There was also a significant time \times group interaction for state anxiety [$F(3,183) = 2.75, p = 0.04, \eta^2 = 0.04$], see Fig.7 below.

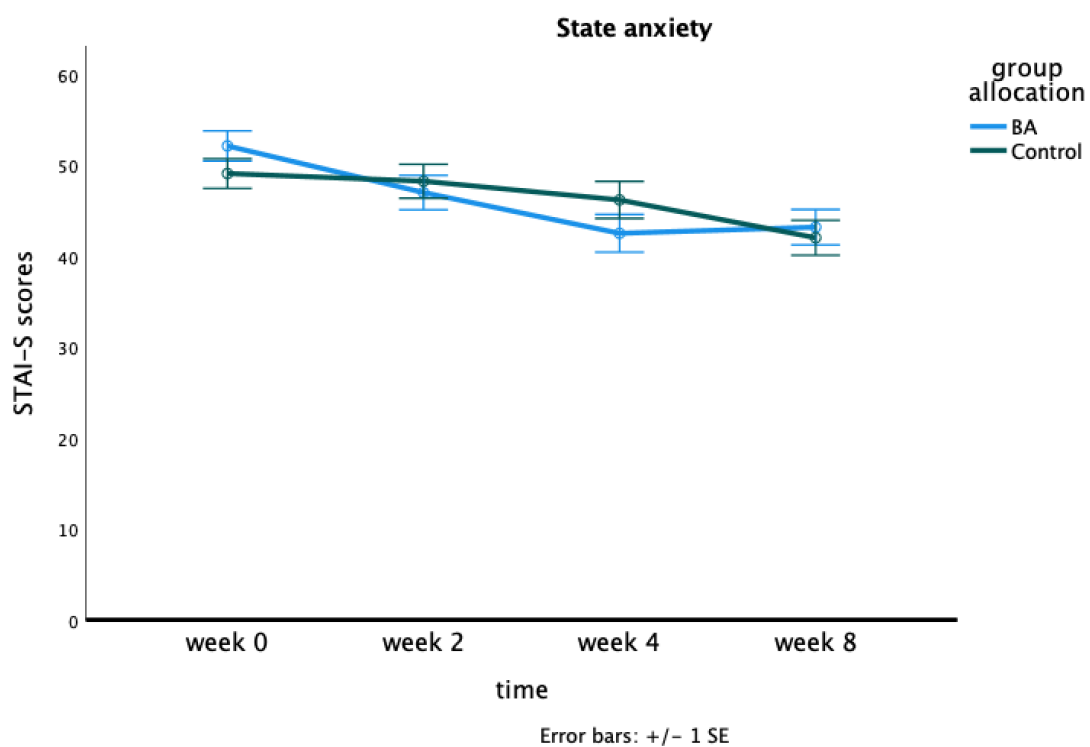


Figure 7. State anxiety symptoms as measured by the STAI-S from baseline (week 0), to midway (week 2), end of study (week 4) and one-month follow-up. Error bars show ± 1 standard error.

Simple main effects of group did not show any differences between groups at any time point ($p > 0.05$), see Table 6 for detailed results.

Table 6. Detailed statistics for post-hoc t-tests comparing the behavioural activation (BA) and control group (C) in state anxiety scores as measured by the STAI-S.

BA vs control	T-test statistics	Group M and SD
Week 0	$t(64) = 0.98, p = .33, d = 0.24$	BA: $M = 51.91, SD = 9.26$ C: $M = 49.64, SD = 9.59$
Week 2	$t(64) = -0.57, p = .57, d = 0.14$	BA: $M = 47.09, SD = 10.16$ C: $M = 48.55, SD = 10.55$
Week 4	$t(65) = -1.27, p = .21, d = 0.31$	BA: $M = 43.33, SD = 11.34$ C: $M = 46.94, SD = 11.84$
Week 8	$t(65) = 0.09, p = .93, d = 0.02$	BA: $M = 43.52, SD = 11.33$ C: $M = 43.26, SD = 11.15$

Simple main effects of time showed that the BA group showed significant differences between baseline and midway, baseline and end of intervention, baseline and follow-up, and midway and end of intervention (see Table 7 below for detailed results).

Table 7. Detailed statistics for post-hoc t-tests comparing the behavioural activation (BA) group across time in state anxiety scores (STAI-S). One asterisk indicates $p < .05$, two asterisks indicate $p < .001$.

BA	T-test statistics	Group M and SD
Week 0 vs Week 2 *	$t(31) = 2.93, p = 0.006, d = 0.48$	Week 0: $M = 52.16, SD = 9.15$ Week 2: $M = 47.03, SD = 10.46$
Week 0 vs Week 4 **	$t(31) = 4.10, p < 0.001, d = 0.83$	Week 0: $M = 52.16, SD = 9.15$ Week 4: $M = 42.55, SD = 11.24$
Week 0 vs Week 8 **	$t(31) = 4.10, p < 0.001, d = 0.78$	Week 0: $M = 52.16, SD = 9.15$ Week 8: $M = 43.23, SD = 11.49$
Week 2 vs Week 4 *	$t(32) = 2.88, p = .007, d = 0.35$	Week 2: $M = 47.09, SD = 10.46$ Week 4: $M = 42.55, SD = 11.24$
Week 2 vs Week 8 *	$t(31) = 2.21, p = .04, d = 0.37$	Week 2: $M = 47.16, SD = 10.32$ Week 8: $M = 43.16, SD = 11.31$
Week 4 vs Week 8	$t(31) = -0.18, p = .86, d = 0.02$	Week 4: $M = 42.88, SD = 11.21$ Week 8: $M = 43.16, SD = 11.31$

The control group only showed a significant difference between baseline and follow-up, as well as midway and follow-up, see Table 8 below.

Table 8. Detailed statistics for post-hoc t-tests comparing the control group (C) across time in state anxiety scores (STAI-S). One asterisk indicates $p < .05$, two asterisks indicate $p < .001$.

Control	T-test statistics	Group M and SD
Week 0 vs Week 2	$t(31) = 0.59, p = .56, d = 0.09$	Week 0: $M = 49.13, SD = 9.27$ Week 2: $M = 48.28, SD = 10.61$
Week 0 vs Week 4	$t(32) = 2.09, p = .05, d = 0.27$	Week 0: $M = 49.64, SD = 9.59$ Week 4: $M = 46.70, SD = 11.94$
Week 0 vs Week 8 **	$t(32) = 3.82, p < .001, d = 0.66$	Week 0: $M = 49.64, SD = 9.59$ Week 8: $M = 42.85, SD = 11.05$
Week 2 vs Week 4	$t(32) = 1.30, p = .21, d = 0.19$	Week 2: $M = 48.55, SD = 10.55$ Week 4: $M = 46.48, SD = 11.72$
Week 2 vs Week 8 *	$t(32) = 3.30, p = .002, d = 0.58$	Week 2: $M = 48.55, SD = 10.55$ Week 8: $M = 42.52, SD = 10.41$
Week 4 vs Week 8	$t(33) = 1.81, p = .08, d = 0.32$	Week 4: $M = 46.94, SD = 11.84$ Week 8: $M = 43.26, SD = 11.15$

For trait anxiety, there was no significant interaction [$F(1.78, 87.08) = 2.86, p = .07, \eta^2 = .04$].

Social support

Levels of social support were measured by the MSPSS and showed significant time \times group interaction [$F(2,120) = 5.21, p = .007, \eta^2 = 0.08$]. Simple main effects of group did not show any differences between groups at any time point (see Table 9 for detailed results).

Table 9. Detailed statistics for post-hoc t-tests comparing the behavioural activation (BA) and control group (C) in social support scores as measured by the MSPSS.

BA vs control	T-test statistics	Group M and SD
Week 0	$t(63) = -1.0, p = .32, d = 0.25$	BA: $M = 58.73, SD = 12.76$ C: $M = 61.94, SD = 13.14$
Week 2	$t(63) = 0.28, p = .78, d = 0.07$	BA: $M = 62.58, SD = 11.49$ C: $M = 61.72, SD = 12.93$

Week 4	$t(64) = 0.96, p = .34, d = 0.24$	BA: $M = 62.66, SD = 11.64$ C: $M = 59.26, SD = 16.61$
--------	-----------------------------------	---

Simple main effect of time analysis showed that the BA group had a significant increase in social support scores between baseline and midway and baseline and end of intervention (see Table 10 for detailed results).

Table 10. Detailed statistics for post-hoc t-tests comparing the behavioural activation (BA) group across time in social support scores (MSPSS). One asterisk indicates $p < .05$.

BA	T-test statistics	Group M and SD
Week 0 vs Week 2 *	$t(31) = -2.47, p = .02, d = 0.27$	Week 0: $M = 59.09, SD = 12.79$ Week 2: $M = 62.34, SD = 11.59$
Week 0 vs Week 4 *	$t(30) = -2.75, p = .01, d = 0.28$	Week 0: $M = 58.42, SD = 12.41$ Week 4: $M = 62.52, SD = 11.80$
Week 2 vs Week 4	$t(31) = -0.57, p = .58, d = 0.06$	Week 2: $M = 61.97, SD = 11.12$ Week 4: $M = 62.66, SD = 11.64$

The control group did not show any significant differences between any time points ($p > 0.05$), see Table 11 below.

Table 11. Detailed statistics for post-hoc t-tests comparing the control group across time in social support scores (MSPSS).

Control	T-test statistics	Group M and SD
Week 0 vs Week 2	$t(30) = -0.03, p = .98, d = 0.003$	Week 0: $M = 61.35, SD = 12.93$ Week 2: $M = 61.39, SD = 13.00$
Week 0 vs Week 4	$t(31) = 1.50, p = .15, d = 0.12$	Week 0: $M = 61.94, SD = 13.14$ Week 4: $M = 60.25, SD = 15.21$
Week 2 vs Week 4	$t(31) = 1.50, p = .15, d = 0.12$	Week 2: $M = 61.72, SD = 12.93$ Week 4: $M = 60.09, SD = 15.08$

Automatic negative thoughts

Automatic negative thoughts, as measured by the ATQ, showed no significant time \times group interaction [$F(2,122) = 0.42, p = 0.66, \eta^2 = 0.007$] but showed a significant main effect of time [$F(2,122) = 15.46, p < 0.001, \eta^2 = 0.20$]. This was driven by a significant decrease in negative

thoughts from baseline ($M = 34.97$, $SD = 9.77$) to midway time point ($M = 32.33$, $SD = 10.43$, $p = 0.02$), as well as from baseline to final time point ($M = 29.05$, $SD = 10.43$, $p < 0.001$) for all participants.

Daily mood

Due to a large amount of missing data in week 4 of the Mood Zoom dataset, we only analysed the first three weeks of intervention. There was no time \times group \times valence interaction [$F(4,260) = 1.25$, $p = .29$, $\eta^2 = 0.02$].

Covid-related lifestyle factors

There was a significant main effect of time for COVID-related stress [$F(2.48,160.93) = 3.25$, $p = .03$, $\eta^2 = 0.05$]. A significant decrease in stress ($p = .02$) was found for all participants between baseline ($M = 32.18$, $SD = 0.93$) and end of study ($M = 29.83$, $SD = 0.95$).

There was also a significant main effect of COVID-related life-style disruption [$F(2.36,153.67) = 33.37$, $p < .001$, $\eta^2 = 0.34$], which significantly decreased for all participants between baseline ($M = 17.30$, $SD = 0.46$) and all time points, including the one-month follow-up ($M = 12.28$, $SD = 0.53$, $p > .001$).

COVID-related anxiety also showed a main effect of time [$F(2.09,131.37) = 11.75$, $p < .001$, $\eta^2 = 0.16$] with significant decreases between baseline ($M = 31.59$, $SD = 0.58$) and final time point ($M = 29.85$, $SD = 0.62$, $p = .001$) as well as follow-up ($M = 28.74$, $SD = 0.68$, $p > .001$) for all participants.

Sensitivity analysis: effect of antidepressant treatment

Exclusion of participants on antidepressants from the analyses did not affect the interaction effects for any of the questionnaire measures.

Discussion

This chapter investigated whether a short (4 week) course of non-specialist BA administered over video calls can be effective for treating low mood in the context of the COVID-19 pandemic, when options for activities are significantly constrained. We found that the intervention had a significant positive effect on our primary measures with medium to large effect sizes, decreasing participants' ratings of depression and increasing activation. Crucially, beneficial effects remained at one-month follow-up, suggesting the intervention had persistent effects beyond the treatment window, consistent with previous reports (Carlbring et al., 2013).

Our study adds further evidence of the flexibility with which BA can be administered online through video calls, making it safe during a period of social distancing. Since it was delivered by two researchers after just 15 hours of training and under the supervision of a PWP, this suggests that new non-specialist practitioners could be trained very quickly during a period of heightened mental health risk in the population. The intervention was rated as acceptable by the participants, and it also increased their rating of social support and decreased state anxiety and anhedonia. This demonstrates that BA may be a particularly helpful intervention during a public health crisis, as it can improve multiple facets of mental health quite quickly and it is more cost-effective than full CBT (Richards et al., 2016).

To our knowledge, this is the first study to show that BA can be effective in a shorter, 4-week format, as opposed to the standard protocol of 8 to 24 weeks (Jacobson et al., 1996; Dimidjian et al., 2006; Richards et al., 2016). Nonetheless, other converging evidence also supports this possibility. For example, Fu et al. (2021) recently published a mediation analysis of a large 8-week BA trial, reporting that a significant change in activation had already occurred at 4 weeks, which then led to a decrease in depression at 6 weeks. The reason why it took 2 weeks longer to show a symptomatic effect than the current study may be due to the self-help format of BA that was used, which may have slowed down the efficacy, as well as due to their comparison with an active control group, where it could take longer for a between-group difference to emerge. It is possible that the effects of the 4-week paradigm would not last as long as effects of the longer programmes. Further research would be warranted to examine more long-term effects beyond just one-month follow-up.

Several of our secondary measures improved for all participants across time. Across both groups, we found significant decreases in automatic negative thoughts as well as COVID-related stress, anxiety and lifestyle disruption. This indicates a degree of spontaneous recovery, possibly as participants adapted to the circumstances of the pandemic. One longitudinal study found that the highest level of mental health problems occurred early in lockdown and then declined rapidly (Fancourt et al., 2021), suggesting that some of the emotional impact of the pandemic may be temporary and people eventually adapt to it. Considering the period of recruitment (May–October 2020), most participants experienced the UK lockdown measures as either stable or gradually improving, rather than becoming stricter, which may have impacted these measures.

Strengths and limitations

The strengths of this study include its randomised controlled design, the use of well-validated measurement tools as well as new tools designed to measure COVID-specific factors. The intervention was based on an established BA protocol routinely used in UK NHS services, developed in collaboration with a Psychological Wellbeing Practitioner and well accepted by participants.

From the baseline descriptive measures it appears that our sample was fairly similar to other studies on mild-to-moderate low mood. Previous research using the behavioural activation scale (BADs) found a very similar average activation score and standard deviation in a moderately depressed community sample (Kanter et al., 2009) and our state anxiety scores were also similar to other research on depressed participants (Godlewska et al., 2012). Trait anxiety was found to be somewhat lower than the average reported in a past review (Fisher and Durham, 1999), which is likely because we did not focus our recruitment on participants with anxiety. Our study showed slightly higher social support levels than another depressed sample (Gladstone et al., 2007), which may affect the generalisability of our findings.

While online recruitment allowed us to include a wider range of participants from all of UK, there were limitations to the representativeness of our sample. Firstly, we excluded participants with a number of physical and psychiatric comorbidities, higher levels of depression severity as well as suicidal thoughts. This was to ensure the safety of all participants, since the remote

format of the study did not allow us to intervene in the event of a mental health crisis. Moreover, our sample had a much higher proportion of females than males. Research has shown that women may be particularly at risk of depression during COVID-19 (Daly, Sutin, & Robinson, 2020), which may explain their greater interest in our study. While these limitations should be considered, another research study in India found a similar BA protocol to be effective in a representative primary care sample with virtually no exclusion criteria (Patel et al., 2017). This suggests that lay-delivered BA can remain effective across different populations and contexts.

It is important to note that the BA practitioners of this study were researchers holding degrees in psychology and their general knowledge from the field as well as interpersonal skills may have made the success of BA more likely. Nonetheless, neither of them had any previous formal training in therapeutic skills. While other studies recruited BA counsellors without any mental health related education (Patel et al., 2017), assessment of communication and interpersonal skills was still an integral part of the selection process (Singla et al., 2014). It is therefore likely that BA can be successfully carried out through lay delivery, as long as practitioners have the right social skills to establish a therapeutic relationship.

Conclusion of Chapter 2

This study presents compelling evidence that online BA can be a helpful intervention to disseminate during a public health crisis when the need for effective mental health treatments increases. Crucially, non-specialists may be trained quickly to supplement over-subscribed mental health services. A shorter treatment protocol can have more robust effects on depression symptoms than previously thought. While our effects persisted at one-month follow-up, future studies should examine more long-term effects of this treatment paradigm.

Chapter 3: Examining the effects of behavioural activation on cognition in comparison to a passive control group during the COVID-19 pandemic

Introduction

Changes in emotional cognitive processing, particularly the increase in negatively biased perception and memory, are believed to play an important role in the risk and maintenance of depressive symptoms (Disner, Beevers, Haigh, & Beck, 2011). As a result, negative affective bias appears to be a useful objective marker of psychiatric vulnerability. For example, it can be detected in remitted patients before relapse (Bouhuys et al. 1999), as well as in never-depressed individuals with high neuroticism (Chan et al., 2007) or family history of depression (Joormann, Talbot, & Gotlib, 2007). This chapter investigates whether behavioural activation treatment for depression during the COVID-19 pandemic affects these cognitive markers and whether early cognitive changes could be a possible mechanism through which the intervention provides benefits.

Evidence suggests that the negative mental health effects of the COVID-19 pandemic can also be detected using measures of emotional cognition. For example, Bland et al. (2021) reported that pandemic-related disruption of social contact was significantly associated with reduced recognition of happy faces in a sample of participants with no previous mental health problems. Similarly, Melendez et al. (2020) compared emotional cognition in young adults in and out of pandemic-related confinement and found that confinement was associated with decreased facial recognition of happiness and increased perception of sadness. A cross-sectional study by Gillespie et al. (2020) also found that vulnerability to depression during COVID-19 was significantly associated with negative bias in the Facial Expression Recognition Task (FERT). This further supports the evidence that psychological vulnerability has increased during COVID-19, particularly in the context of social isolation, and that this can be detected using objective cognitive markers.

Measures of emotional cognition may also be a useful tool to detect treatment efficacy in this unique context. Previous research suggests that affective bias can show therapeutic effects early on in treatment, often before the effects on mood appear, and may therefore serve as an early

marker of response (Tranter et al., 2009), possibly allowing for better selection of the right treatment type for specific patients. However, these markers have mainly been explored in pharmacological treatment for depression (see Harmer, Duman, & Cowen, 2017 for review). While some promising findings suggest there could be similar early cognitive effects in CBT for panic (Reinecke et al., 2013), research on cognitive markers in psychological treatments for depression has been scarce and its results mixed (Groves et al., 2021). To our knowledge, behavioural activation specifically has not been investigated using these measures and it is unknown whether a largely behavioural intervention would be able to influence these kinds of cognitive processes.

Interestingly, preliminary evidence suggests that emotional cognition could play an important mediatory role in BA treatment. For example, previous research reported that cognitive bias plays an important role in mediating the relationship between behavioural activation and inhibition systems (see Chapter 1) and symptoms of depression (Farashbandi et al., 2020; Farashbandi et al., 2021); this suggests that a change in cognitive bias could be a plausible mechanism of BA efficacy. However, this research only included self-report measures of cognition, which could be biased by demand effects and may not tap into the same automatic processes measured by objective behavioural tasks. Gillespie et al. (2020) reported that behavioural activation appears to protect against the development of negative affective bias during the COVID-19 pandemic, but this evidence is only cross-sectional. The only available experimental study using an objective task measured participants' approach and avoidance tendencies before and after BA treatment, but authors did not find a significant effect (Nasrin et al., 2017). Since this study only used a single-session BA paradigm, the therapeutic effects may not have been strong enough to elicit a significant cognitive change. It would therefore be informative to investigate a longer course of BA using behavioural measures of affective bias.

The present study investigated whether a 4-week BA treatment for depression administered online during the COVID-19 pandemic influences objective markers of emotional cognition. We used a battery of emotion-related cognitive tasks, which is known to be sensitive to antidepressant drug effects, and computed the same outcome measures routinely used in previous research (see review by Harmer et al., 2017). Our primary outcome measures were accuracy and misclassification rates during the Facial Emotion Recognition Task as these were the outcomes previously found to be affected by the COVID-19 pandemic (Gillespie et al., 2020;

| Melendez et al., 2020; Bland et al., 2021). We hypothesised that, in comparison to a passive control group, the BA group would show an increase in positive bias or a decrease in negative bias from baseline early on in treatment (week 2 of the intervention). We predicted that this would occur before conscious effects in mood could be detected at the end of intervention (week 4) or at one-month follow-up (week 8).

As secondary measures, we also examined reaction times in the Facial Emotion Recognition Task, accuracy and reaction times in the Emotion Categorisation task, accuracy and false alarm recall in the Emotion Recall Task as well as choice behaviour and amount of money won and lost in the Probabilistic Instrumental Learning Task. This would allow us to investigate other domains of emotional cognition that have been linked to antidepressant drug treatment. Similarly, we hypothesised that significant changes in these measures would occur at week 2, preceding conscious mood effects at week 4 and week 8.

Overall, we hypothesised that changes in emotional cognition would precede and predict improvement in depression symptoms, indicating that affective bias plays a mechanistic role in BA treatment and can serve as an early objective marker of efficacy.

Methods

| Detailed descriptions of the study participants and the BA intervention have previously been presented in Chapter 2.

Cognitive tasks

All tasks were delivered online via the Gorilla platform.

In the **Facial Emotion Recognition Task** (FERT; see Harmer et al., 2003), participants were presented with faces of four actors (two male and two female) showing the six basic emotions (anger, disgust, fear, happiness, sadness, and surprise) at different levels of intensity (from 10 to 100%, increasing in 10% steps) alongside neutral faces. Face images were selected from the Karolinska Directed Emotional Faces (KDEF) set and morphed from neutral to each emotional

expression to create the steps of 10% intensity. Sixty emotional images are presented from each actor (six emotions at ten stages of intensity), together with 10 neutral images (2 from two actors and 3 from the two other actors), leading to 250 images in total. Each image was presented only once. A different set of stimuli (same emotions expressed by four different actors) was presented in each of the three versions of the task (used at week 0, week 2 and week 4) to allow for repeated testing without the confounding effects of practice.

Participants were asked to identify the emotion in each face as quickly and accurately as possible. This was done by selecting the correct emotion label from a circular arrangement using a cursor and then moving the cursor to a fixation cross in the centre. They were instructed to press just one option for every face and the next face would not appear until they responded to the current one. They were also informed that some expressions may look ambiguous, and the decision may be difficult, but they should make their best guess. All participants carried out an engagement check before the start of the task, which asked them four questions about the task instructions, such as “What should you do when presented with a face on screen?”. Correct completion of all engagement checks was necessary to proceed to the task. Six practice faces were presented at the beginning to further allow participants to check their understanding of the task, but no feedback was given.

The task consisted of 4 blocks, with two blocks containing 63 trials and two blocks containing 62 trials (250 trials in total). Faces were presented in a fixed randomised order, ensuring that there would be a balance of emotions, actors, and level of difficulty (i.e., emotion intensity) across all blocks, that there would not be more than two consecutive trials showing the same emotion or the same actor, as well as no more than six consecutive trials of all positive or all negative expressions.

Each trial started with a fixation cross presented in the centre of the screen for 1s, followed by a face presented for 500ms. Once the face has been presented, participants were given a further 9.5s to identify the emotion by clicking on the correct label with their cursor (10s in total). The following trial began once the participant responded, or after 10s if no response was made. The total duration of each trial was therefore determined by the participant’s reaction time. Participants were given a rest period for as long as necessary when the task was 25%, 50% and 75% complete.

We measured percentage accuracy and reaction times, comparing responses to neutral faces, faces of positive valence (happiness and surprise) and negative valence (anger, disgust, fear, and sadness), as well as comparing all emotions individually. We also measured misclassification counts, both in terms of specific emotions misclassified as something else (e.g., how many times an angry face was misclassified as something other than “angry”), as well as how often faces were incorrectly misclassified as a specific emotion (e.g., how many times a non-angry face was misclassified as “angry”), comparing individual emotions as well as groups of emotions (positive, negative, and neutral).

In the **Emotion Categorisation Task** (ECAT; see Chan, Harmer, Goodwin, & Norbury, 2008), participants were presented with 40 personality adjectives (such as ‘cheerful’ or ‘dishonest’), 20 positives and 20 negatives, based on previously established likableness ratings (Anderson, 1968). Words were also matched for length, frequency, and meaningfulness. Different sets of words were presented at each of the three time points (week 0, week 2 and week 4) to decrease practice effects from repeated testing.

Participants were first shown the instructions for the task, stating that they will be shown various personality characteristic words and that they should imagine overhearing someone describing them in this way. They should then indicate whether they would like or dislike to be described that way by pressing either the “L” or “D” keys on the keyboard with their index fingers as quickly and accurately as possible. They were also told that later in the testing session they would be asked to recall as many words from this task as they can. Participants had to complete an engagement check consisting of two questions testing their understanding of the instructions and successful completion allowed them to proceed to the task.

Participants were first given 3 practice trials, followed by 40 test trials. In each of trial, a fixation cross was first presented for 500ms, followed by a word shown for 500ms. Participants were given 2500ms to indicate whether they would like or dislike to be described with the word. After they made a response, there was a 750ms pause before the next trial. If no response was made, this pause started after 3000ms. Participants were not given any breaks in this task.

We measured percentage accuracy as well as reaction times, comparing the groups of positive and negative words together.

In the **Emotional Recall Task** (EREC; see Harmer, Shelley, Cowen, & Goodwin, 2004), which took place 15min after the ECAT, participants were asked to remember as many of the ECAT personality adjective words as possible. They were given 3 minutes to type these words into a text box on the screen. Halfway through the task, a timer appeared on the screen to count down the remaining 90 seconds. After 3 minutes, the screen automatically moved on to the next task. We measured the accuracy of positive and negative recall, as well as the number of positive and negative intrusions that the participants wrote down.

In the **Probabilistic Instrumental Learning Task** (PILT; see Walsh, Browning, Drevets, Furey, & Harmer, 2018), participants were tested on probabilistic reward learning. In this “2-arm bandit” task, participants were simultaneously shown two abstract symbols (taken from the Agathodaimon font), which had reciprocal probabilities (0.7 v. 0.3) of either winning v. not winning money (in ‘gain’ trials) or losing v. not losing money (in ‘loss’ trials).

Participants were instructed that they would start the task with 100 pence and should aim to win as much money as possible, because the final amount would be added to their study reimbursement. They were specifically informed of the probabilistic nature of the task with the instruction: “The best symbol won’t always result in winning (or not losing) points - it will just result in winning (or not losing) points more frequently than the other symbol.” Before the task started, participants had to successfully complete an engagement check consisting of two questions about the task instructions. Afterwards, they were given three practice trials with unique symbols that would not appear again in any of the task blocks, and they were not shown a running total.

The task itself consisted of three blocks of 60 trials each, with 30 ‘gain’ trials and 30 ‘loss’ trials (180 trials in total). The trials were randomised within each block and new randomisation was generated with each run of the task. Only the left/right position of the symbol was pre-randomised to ensure symbols appeared equally on the left and right sides. In each trial, the fixation cross was first presented for 250ms, followed by the pair of symbols presented for 2000ms or until one symbol is selected with a mouse click. The total duration of each trial was

thus determined by individual reaction times of participants. After each choice, participants were given feedback about their monetary outcome (20 pence won, 20 pence lost or no change) and their current total. Participants were given three rest periods during which they were shown their progress through the task (25%, 50% and 75%).

A different set of stimuli was used at each run of the task to prevent learning effects. The main outcome measures were total amount won and lost, percentage accuracy for choosing high probability symbols across the whole task and in the last 20 trials.

Procedure

All cognitive tasks (FERT, ECAT, EREC and PILT) were administered remotely via the Gorilla platform at baseline (week 0), halfway through the intervention (week 2) and at the end (week 4). Multiple choice questions designed to test understanding of the task instructions had to be completed before participants could proceed with each task.

Power calculation

As described in Chapter 2, the study was well-powered to detect a moderate-to-large effect size (Cohen's d of 0.6) for a difference between two means. It was difficult to estimate the expected effect size for our primary measure of emotional cognition (FERT accuracy), as most of the past research using this task did so in the context of pharmacological interventions with either placebo pill or healthy participant controls. Nonetheless, one study with a passive control group testing facial recognition after CBT in panic disorder patients found a large effect size (Cohen's d of 1.5; Reinecke et al., 2013), indicating that there may be at least a moderate effect size that we could detect.

Statistical analysis

Measures of emotional cognition were analysed as change-from-baseline scores (calculated as "week 4 minus week 0" and "week 2 minus week 0") using a two-way ANCOVA controlling for baseline, with group as a between-subject factor and time (week 2 v. 4) and task condition as within-subject factors (ECAT/EREC/FERT: valence - positive, negative; FERT: emotion - anger, disgust, fear, happiness, sadness, surprise or neutral; PILT: trial type - win, loss). Change-from-baseline scores were correlated with emotional cognitive measures at week 2 and depression scores at week 4 and 8 while controlling for baseline.

The rest of the analysis procedure remained the same as described in Chapter 2.

Results

Baseline descriptive statistics for all cognitive tasks can be found in the Appendix (Table 52).

Facial Emotion Recognition Task (FERT)

Accuracy for recognising positive and negative emotions

FERT accuracy showed a significant time \times group \times valence interaction [$F(1,60) = 4.36, p = .04, \eta^2 = 0.07$]. To further understand this interaction, separate analyses were conducted for week 2 and week 4. There was no significant group \times valence interaction at week 2 [$F(1,63) = 0.51, p = .48, \eta^2 = 0.008$]. However, there was a significant group \times valence interaction at week 4 [$F(1,63) = 12.58, p = .001, \eta^2 = 0.17$].

Simple main effect analysis revealed no significant difference in accuracy for positive faces at week 4 [$t(64) = 1.17, p = .79, d = 0.23$]. However, there was a significant difference in accuracy for negative faces at week 4 [$t(63) = -3.36, p = .001, d = 0.83$]; with the BA group showing a greater decrease in accuracy for negative faces compared with baseline ($M = -5.37, SD = 6.63$) than the control group ($M = -0.57, SD = 4.77$), see Fig. 8 below.

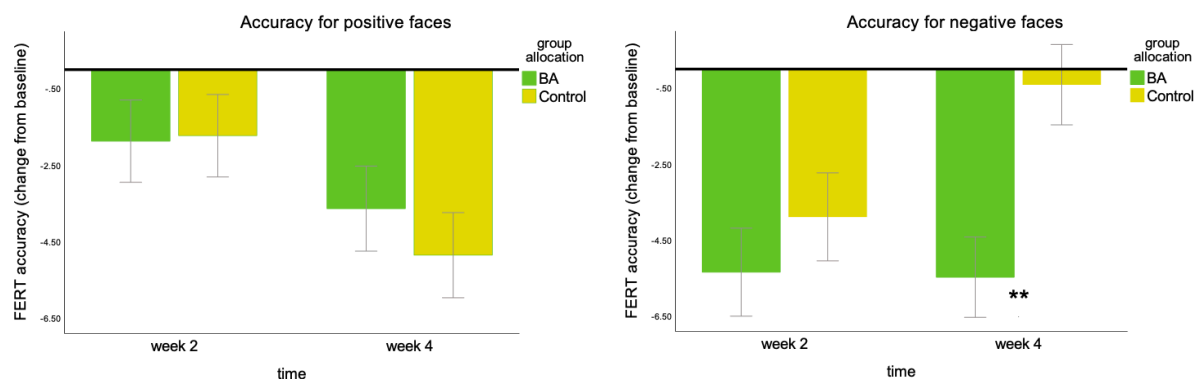


Figure 8. Change from baseline in accuracy towards positive and negative faces in the Facial Emotion Recognition Task at midway (week 2) and at the end of the study (week 4). Error bars show ± 1 standard error. Double asterisks indicate $p = .001$.

Accuracy for recognising individual emotions

When comparing accuracy to individual emotions, there was a significant time \times group \times emotion interaction [$F(6,324) = 2.60, p = .02, \eta^2 = .05$], but the subsequent group \times emotion interaction was not significant at midway time point [$F(6,324) = 0.33, p = .92, \eta^2 = .006$] or at the final time point [$F(6,330) = 1.81, p = .10, \eta^2 = .03$].

Exploratory analysis of simple main effects of group showed only a significant difference for recognition of fear at the end of the intervention (see Figure 9 below). For these fearful faces, the control group significantly increased in accuracy compared with baseline, while the BA group decreased (see Table 12 for detailed results).

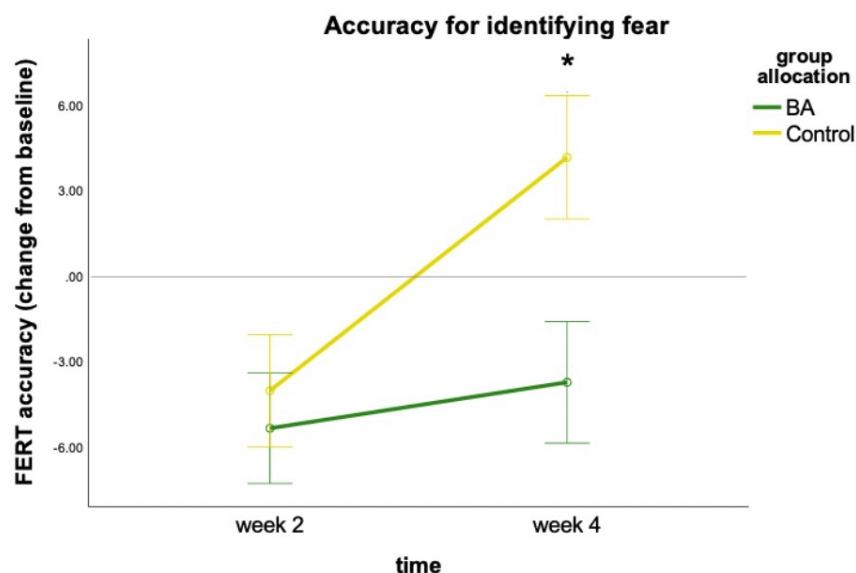


Figure 9. Exploratory analysis comparing change from baseline in accuracy for fearful faces in the Facial Emotion Recognition Task at midway (week 2) and at the end of the study (week 4). Error bars show ± 1 standard error. Asterisk indicates $p < .05$.

Furthermore, there was a trend towards a significant difference in recognition of sadness at the end of the intervention, with the BA group decreasing in accuracy compared to baseline more than the control group. There was also a trend towards a significant difference in recognition of surprise, with the control group decreasing in accuracy compared to baseline significantly more than the BA group - see detailed results in Table 12 below.

Table 12. Detailed statistics for post-hoc t-tests comparing the behavioural activation (BA) and control (C) groups in FERT accuracy for individual emotions at week 2 and week 4 (change from baseline scores). One asterisk indicates $p < .05$.

FERT accuracy parameter	T-test statistics	Group M and SD
Anger at Week 2	$t(63) = -0.69, p = .50, d = 0.17$	BA: $M = -4.77, SD = 11.12$ C: $M = -2.65, SD = 13.55$
Disgust at Week 2	$t(57.66) = -0.36, p = .72, d = 0.09$	BA: $M = 0.55, SD = 14.02$ C: $M = 1.67, SD = 10.58$
Fear at Week 2	$t(62) = -1.15, p = .25, d = 0.29$	BA: $M = -6.25, SD = 9.31$ C: $M = -3.13, SD = 12.18$
Happy at Week 2	$t(63) = 1.18, p = .24, d = 0.29$	BA: $M = 3.28, SD = 8.17$ C: $M = 0.91, SD = 8.00$
Sad at Week 2	$t(63) = -0.11, p = .91, d = 0.03$	BA: $M = -11.33, SD = 12.59$ C: $M = -10.99, SD = 11.71$
Surprise at Week 2	$t(63) = -0.70, p = .49, d = 0.17$	BA: $M = -6.56, SD = 9.33$ C: $M = -4.92, SD = 9.47$
Neutral at Week 2	$t(62) = -0.55, p = .58, d = 0.14$	BA: $M = 5.00, SD = 14.37$ C: $M = 6.88, SD = 12.81$
Anger at Week 4	$t(64) = -0.60, p = .55, d = 0.15$	BA: $M = 5.31, SD = 8.79$ C: $M = 6.77, SD = 10.74$
Disgust at Week 4	$t(64) = -0.45, p = .65, d = 0.11$	BA: $M = 4.14, SD = 13.81$ C: $M = 5.59, SD = 12.25$
Fear at Week 4 *	$t(63) = -3.07, p = .003, d = .76$	BA: $M = -4.45, SD = 11.96$ C: $M = 4.39, SD = 11.27$
Happy at Week 4	$t(64) = -0.35, p = .73, d = 0.09$	BA: $M = -3.28, SD = 10.23$ C: $M = -2.50, SD = 7.79$
Sad at Week 4	$t(64) = -1.89, p = .06, d = 0.46$	BA: $M = -26.48, SD = 16.32$ C: $M = -19.56, SD = 13.34$
Surprise at Week 4	$t(64) = 2.01, p = .05, d = 0.50$	BA: $M = -3.20, SD = 8.24$ C: $M = -7.65, SD = 9.61$
Neutral at Week 4	$t(63) = -0.66, p = .51, d = 0.16$	BA: $M = -4.38, SD = 21.99$ C: $M = -1.21, SD = 16.35$

Misclassification of positive and negative emotions

When analysing the number of misclassifications of positive faces as either negative or neutral (negative bias), there was no significant time \times group interaction [$F(1,61) = 1.15, p = 0.30, \eta^2 = 0.02$], see Fig. 10 (left). For misclassifications of negative faces as positive or neutral (positive bias), there was a significant time \times group interaction [$F(1,61) = 6.65, p = 0.01, \eta^2 = 0.10$], see Fig.10 (right).

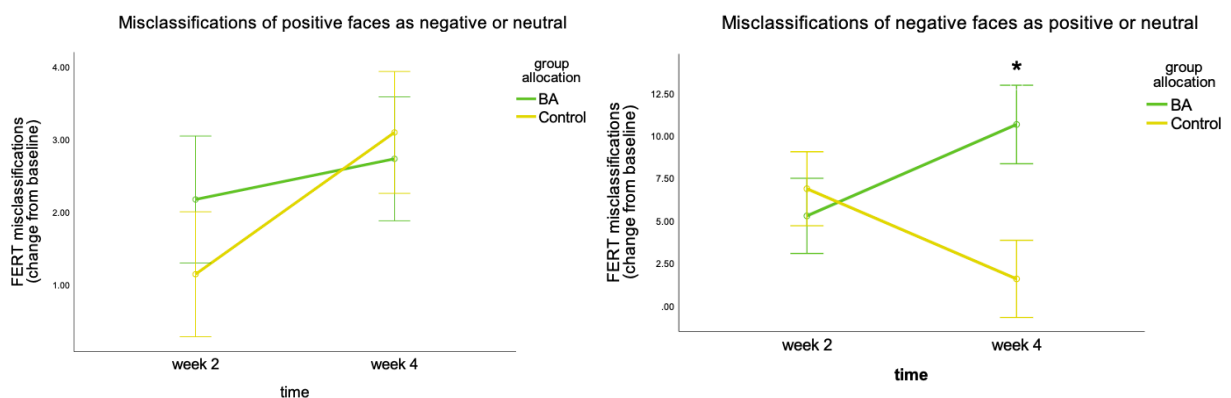


Figure 10. (left) shows no significant difference between the groups in misclassification of positive faces as negative or neutral (negative bias) at either time point; (right) shows a significant difference between the behavioural activation group and the control group in misclassifying negative faces as positive or neutral (positive bias) at the end of the study (week 4). Error bars show ± 1 standard error. Asterisk indicates $p < .05$.

The interaction was driven by a significant difference between groups at the final time point, wherein the BA group showed a significantly greater increase in misclassifications than the control group, see table 13 below.

Table 13. Post-hoc tests comparing the behavioural activation (BA) group and the control group (C) in the number of FERT misclassifications of negative faces as positive or neutral (positive bias). One asterisk indicates $p < .05$.

Between-group comparison	T-test statistics	Group M and SD
Week 2	$t(62) = 0.98, p = .33, d = .25$	BA: $M = 7.38, SD = 11.43$ C: $M = 4.56, SD = 11.54$
Week 4 *	$t(64) = 2.52, p = 0.01, d = 0.62$	BA: $M = 10.44, SD = 13.84$ C: $M = 2.35, SD = 12.19$

There was no significant time \times group interaction for misclassifying negative faces as other negative emotions [$F(1,61) = 0.39, p = 0.54, \eta^2 = 0.006$].

Misclassification of individual emotions

When comparing how individual emotions were misclassified (what emotions were selected instead of the correct option), there was a significant time \times group \times valence interaction [$F(6,306) = 3.69, p = .001, \eta^2 = .07$]. There was no significant group by emotion interaction at the midway time point [$F(4.82,284.34) = 0.23, p = .95, \eta^2 = .004$]. A significant group by emotion interaction persisted only at the final time point [$F(6,324) = 2.86, p = .01, \eta^2 = .05$].

When comparing the misclassification of individual emotions, there was a significant difference between groups in how fear was misclassified at week 4, with the BA group increasing in tendency to misclassify fear as other emotions compared with baseline while the control group decreased, see Fig. 11 (left). There was also a difference in misclassifying surprise, with the control group increasing from baseline significantly more than the BA group, see Fig. 11 (right).

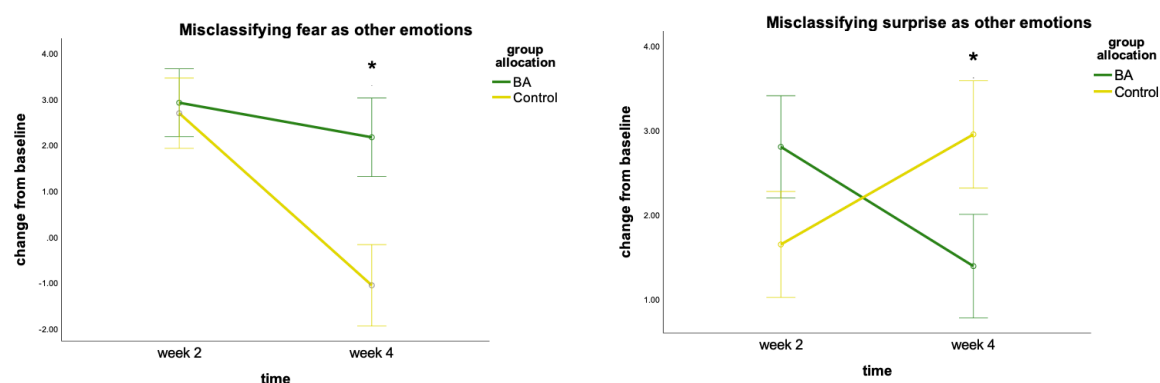


Figure 11 (left) shows that the behavioural activation group significantly increased from baseline in tendency to misclassify fear as other emotions by the end of the intervention (week 4), while the control group decreased; (right) shows that the control group increased in tendency to misclassify surprise as other emotions significantly more than the BA group. Error bars show ± 1 standard error. Asterisk indicates $p < .05$.

The full results of misclassifications of individual emotions is in table 14 below.

Table 14. Detailed statistics for post-hoc t-tests comparing the behavioural activation (BA) and control (C) groups in FERT misclassifications for individual emotions at week 2 and week 4 (change from baseline scores). One asterisk indicates $p < .05$.

Emotion	T-test statistics	Group M and SD
Anger	$t(63) = 0.90, p = .37, d = 0.22$	BA: $M = -2.13, SD = 3.44$ C: $M = -2.97, SD = 4.10$
Disgust	$t(64) = 0.24, p = .81, d = 0.06$	BA: $M = -1.81, SD = 5.27$ C: $M = -2.12, SD = 4.93$
Fear *	$t(63) = 2.42, p = .02, d = .60$	BA: $M = 1.94, SD = 5.05$ C: $M = -0.97, SD = 4.63$
Happy	$t(63) = 0.74, p = .46, d = 0.18$	BA: $M = 1.77, SD = 3.17$ C: $M = 1.21, SD = 3.03$
Sad	$t(64) = 1.91, p = .06, d = 0.47$	BA: $M = 10.66, SD = 6.50$ C: $M = 7.82, SD = 5.56$
Surprise *	$t(64) = -2.22, p = .03, d = 0.55$	BA: $M = 1.31, SD = 2.99$ C: $M = 3.15, SD = 3.67$
Neutral	$t(63) = 0.47, p = .64, d = 0.12$	BA: $M = 0.38, SD = 2.17$ C: $M = 0.15, SD = 1.68$

When comparing which emotions faces were most likely to be falsely selected, there was again a significant three way interaction ($F(6,270) = 3.90, p = .001, \eta^2 = .08$) and a significant group by emotion interaction persisted only at the final time point ($F(6,294) = 4.10, p = .001, \eta^2 = .08$).

As shown in Fig. 12 below, this was driven by the control group falsely selecting fear significantly more than the BA group. Moreover, the BA group was more likely to misclassify emotions as neutral than the control group.

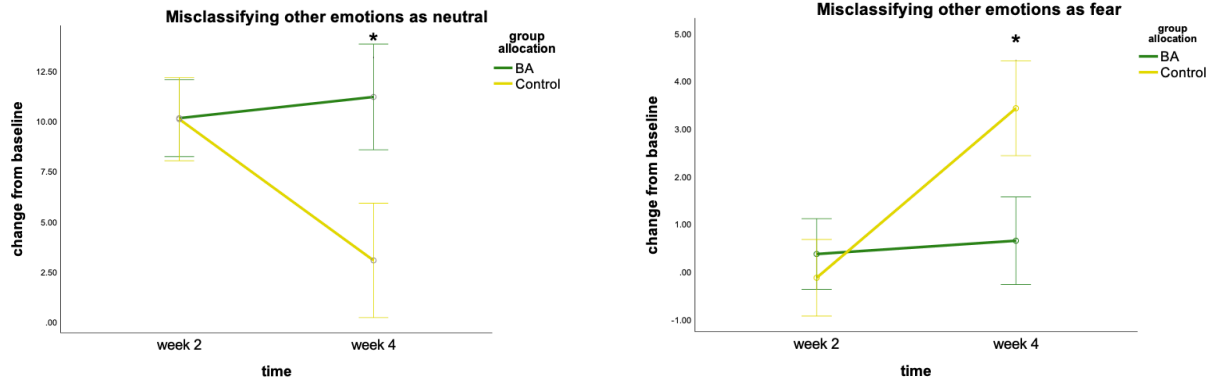


Figure 12 (left) shows that the control group was more likely to incorrectly identify faces as fearful by the end of the study (week 4); (right) shows that the behavioural activation group significantly increased in tendency to misclassify faces as neutral. Error bars show ± 1 standard error. Asterisk indicates $p < .05$.

The full results of misclassifications of individual emotions is in table 15 below.

Table 15. Detailed statistics for post-hoc t-tests comparing the behavioural activation (BA) and control (C) groups in FERT false alarms for individual emotions at week 2 and week 4 (change from baseline scores). One asterisk indicates $p < .05$.

Emotion	T-test statistics	Group M and SD
Anger	$t(63) = -0.91, p = .37, d = 0.23$	BA: $M = 1.48, SD = 6.29$ C: $M = 2.85, SD = 5.81$
Disgust	$t(64) = -0.06, p = .95, d = 0.02$	BA: $M = 1.03, SD = 6.05$ C: $M = 1.12, SD = 5.38$
Fear *	$t(64) = -2.59, p = .01, d = 0.64$	BA: $M = -0.03, SD = 4.62$ C: $M = 3.27, SD = 5.64$
Happy	$t(55.05) = 0.53, p = .60, d = 0.14$	BA: $M = 0.38, SD = 2.09$ C: $M = 0.13, SD = 1.43$
Sad	$t(62) = 0.14, p = .89, d = 0.04$	BA: $M = -5.56, SD = 7.07$ C: $M = -5.81, SD = 6.80$
Surprise	$t(63) = 0.77, p = .44, d = 0.19$	BA: $M = 1.44, SD = 4.84$ C: $M = 0.55, SD = 4.49$
Neutral *	$t(58.51) = 2.67, p = .01, d = .66$	BA: $M = 11.78, SD = 15.58$ C: $M = 2.59, SD = 12.11$

Reaction times to identify faces

There was no significant time \times valence \times group interaction for change from baseline scores for reaction time to classify positive and negative faces at week 2 and week 4 while controlling for baseline [$F(1,58) = 0.08, p = 0.77, \eta^2 = 0.001$].

There was also no significant time \times emotion \times group interaction for change from baseline scores for reaction times to individual emotions at week 2 and week 4 while controlling for baseline [$F(6,294) = 0.32, p = 0.93, \eta^2 = 0.006$].

Correlating changes in facial emotion recognition with changes in depression symptoms

The BA group showed a significant negative correlation between change in accuracy for positive faces at week 2 and change in depression scores at week 4, see Fig. 13 below (left). There also was no significant correlation between change in accuracy for negative faces at week 2 and change in depression at week 4 or week 8, see table 16 for full findings.

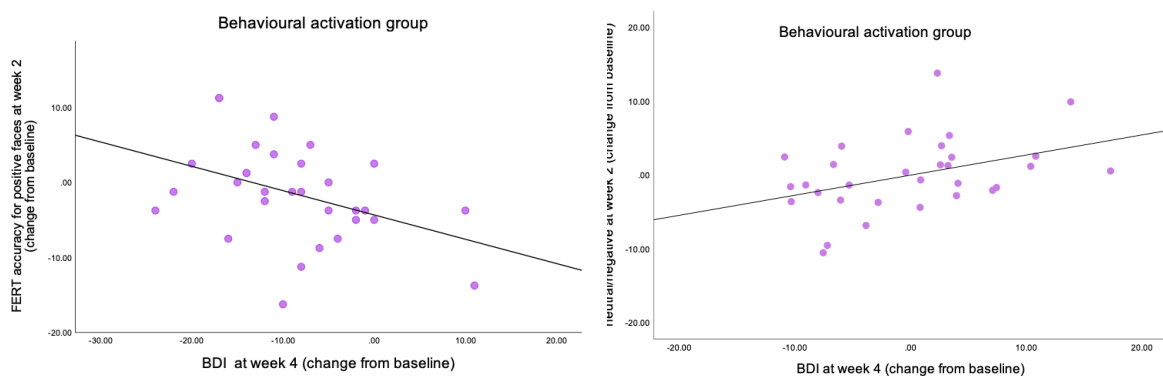


Figure 13 (left). The BA group showing a significant correlation between early change in accuracy for positive faces midway through the study (week 2) and change in depression symptoms at the end of the study (week 4); (right) the BA group showing a significant correlation between early change in misclassification of positive faces as negative/neutral midway through the study (week 2) and change in depression symptoms at the end of the study (week 4).

Moreover, the BA group showed a significant positive correlation between misclassifications of positive faces as negative or neutral (negative bias) at week 2 and depression scores at week 4, see Fig. 13 (right) above. This suggested that early change in the processing of positive faces was associated with later therapeutic gain following treatment. All correlation findings can be found in Table 16 below.

Table 16. Correlation findings for early FERT changes at week 2 and depression symptom changes at week 4 in the BA group.

FERT parameter at Week 2 (BA group only)	Correlation statistics
Accuracy for positive emotions	$r(27) = -0.41, p = .03$
Accuracy for negative emotions	$r(27) = -0.35, p = .07$
Misclassification of positive faces as negative or neutral	$r(27) = 0.41, p = .03$
Misclassification of negative faces as positive or neutral	$r(28) = 0.21, p = .26$
Reaction times to positive emotions	$r(27) = 0.14, p = .48$
Reaction times to negative emotions	$r(27) = 0.29, p = .13$

Emotional categorisation task (ECAT)

There was no significant time \times valence \times group interaction for accuracy in categorising emotional words [$F(1,54) = 1.77, p = 0.19, \eta^2 = 0.03$] or in reaction times [$F(1,52) = 1.26, p = 0.27, \eta^2 = 0.02$].

We did not find any significant partial correlations between change from baseline of any of these variables with change from baseline in depression as measured by the BDI, see table 17.

Table 17. Correlation findings for early ECAT changes at week 2 and depression symptom changes at week 4 in the BA group.

ECAT parameter at Week 2 (BA group only)	Correlation statistics
Accuracy for positive words	$r(29) = 0.17, p = .37$
Accuracy for negative words	$r(29) = 0.26, p = .17$
Reaction times for positive words	$r(29) = 0.05, p = .80$
Reaction times for negative words	$r(29) = -0.19, p = .31$

Emotional Recall Task (EREC)

There was no significant time \times valence \times group interaction for either accurate emotional recall [$F(1,61) = 0.04, p = 0.83, \eta^2 = 0.001$] or the number of intrusions in emotional recall [$F(1,61) = 1.92, p = 0.17, \eta^2 = 0.03$].

However, only the BA group showed a significant positive correlation between the change in negative intrusions at week 2 and the change in depression at week 4, suggesting that reduction in negative intrusions may be related to later therapeutic gain. See Table 18 below for full correlation findings.

Table 18. Correlation findings for early EREC changes at week 2 and depression symptom changes at week 4 in the BA group.

EREC parameter at Week 2 (BA group only)	Correlation statistics
Accuracy for positive words	$r(29) = -0.10, p = .60$
Accuracy for negative words	$r(29) = -0.10, p = .59$
False alarms for positive words	$r(29) = 0.10, p = .59$
False alarms for negative words	$r(27) = 0.38, p = .04$

Probabilistic Instrumental Learning Task (PILT)

There was no significant time \times group \times trial type interaction effect on choice behaviour [full task $F(1,60) = 0.10, p = 0.76, \eta^2 = 0.002$; last 20 trials $F(1,63) = 0.01, p = 0.94, \eta^2 < 0.001$]. There was also significant time \times group \times trial type interaction for the total amount won and lost [$F(1,65) = 1.57, p = 0.21, \eta^2 = 0.02$].

There were no significant correlations between early changes in PILT parameters and later changes in depression scores (see Table 19).

Table 19. Correlation findings for early PILT changes at week 2 and depression symptom changes at week 4 in the BA group.

PILT parameter at Week 2	Correlation statistics
Win trials	$r(30) = -0.23, p = .21$
Loss trials	$r(28) = -0.19, p = .73$
Win trials (last 20)	$r(30) = -0.25, p = .18$
Loss trials (last 20)	$r(30) = 0.11, p = .56$

Sensitivity analysis: effect of antidepressant treatment

After the exclusion of participants on antidepressants ($n = 13$), the interaction effects for FERT accuracy and FERT misclassifications were no longer significant. All other results remained the same.

Discussion

The aim of this chapter was to investigate whether the therapeutic effect of BA on depression during the COVID-19 pandemic can be detected objectively as a change in affective bias. We found that the intervention led to a more positive bias in our primary outcome measures of accuracy and misclassification rates in the Facial Expression Recognition Task (FERT). At the end of the 4-week intervention, the BA group showed a significant decrease in accuracy for recognising negative faces driven by an increased tendency to misclassify them as positive or neutral, both with medium to large effect sizes. Our secondary measures of emotional cognition, including the ECAT, EREC and the PILT tasks, did not show any difference between the groups at any time point.

Analysis of individual emotions in the FERT task further indicated that BA can positively shift emotional cognition in this domain. The BA group was more likely to misclassify fearful faces as displaying a different emotion, while the control group was more likely to misclassify surprise. When examining which emotions were falsely selected instead of the correct options, the BA group was significantly more likely to choose neutral, while the control group was more likely to choose fear. Moreover, exploratory analysis of accuracy to individual emotions showed that the control group was significantly better at recognising fear as well as showing trends towards better recognition of sadness and worse recognition of surprise. These findings support our hypothesis that BA shifted participants' automatic cognition towards more positive judgements, while the control group maintained or increased in their negative cognitive bias.

To our knowledge, this is the first study to find evidence of a purely behavioural treatment for depression affecting negative cognitive bias in facial expression recognition. This bias has been associated with depression and vulnerability to depression (Disner et al., 2011), possibly because it makes negative cognitive schemas more likely to be reinforced. It is therefore notable

that BA shifted the identification of negative faces towards more positive or neutral interpretations. This provides evidence for the theoretical model of BA (Farrand et al., 2013), wherein bidirectional effects occur between cognition and behaviour (as well as physical sensations). It also strengthens the support for BA being efficacious in the unique context of the pandemic, since objective cognitive markers are assumed to be resistant to placebo and other demand effects that may bias self-report (Huneke et al., 2017).

While significant between-group differences in emotional cognition only emerged at week 4, rather than at week 2 as predicted, early cognitive changes still appeared to be important. BA participants who showed early cognitive changes in three of our cognitive measures at week 2 were found to be more likely to show clinical improvement at week 4. This suggests a mechanistic role of affective cognition in BA treatment, possibly because positive shifts in unconscious bias take time to accumulate to eventually cause a conscious mood change. Such effects have previously been shown mainly for pharmacological treatment for depression (see Harmer et al., 2017 for review), and while some evidence from other psychological treatments exists, it has been inconsistent (Porter et al., 2016; Yilmaz et al., 2019). This work highlights how emotional cognition may serve as an early marker of response across different interventions, although the present findings need to be further replicated in larger samples that would allow for well-powered mediation analysis.

Interestingly, the primary effects on emotional cognition were no longer significant after excluding participants taking antidepressants. Since the rate of antidepressant use was similar between the two groups (15% vs 24%), it doesn't seem likely that the effect of medication could be driving the differences reported between groups. While the lack of effect after exclusion could be due to reduced statistical power, it is also possible that the combination of BA and antidepressant medication could have a stronger effect on emotional cognition than BA on its own and could thus be driving the effect. Preclinical research has led to reports that the combination of antidepressants with environmental enrichment supersedes either intervention alone in its beneficial effects; this has been found on several depressive endophenotypes, including reward sensitivity, BDNF, corticosterone and inflammatory markers (Branchi et al., 2013; Alboni et al., 2016). As a result, it has been hypothesised that while antidepressant medication may facilitate neuroplasticity, this effect will be strongest when it is administered in an enriching environment that facilitates positive learning (Harmer et al., 2017). It should

therefore be experimentally investigated in humans whether the environmental enrichment resulting from BA treatment interacts with antidepressant medication on measures of emotional cognition to produce a larger effect than either treatment alone.

Unexpectedly, BA was not found to affect any of the outcome measures in the probabilistic reward learning task, despite our hypothesis that reward decision-making and reinforcement learning could be one of the relevant mechanisms of change. This contrasts with our finding that BA had a significant effect on anhedonia (see Chapter 2), as well as other evidence linking BA to increased reward seeking and reward response (Karimpour-Vazifehkhorani et al., 2020). Since these other findings come from self-report measures rather than objective behavioural tasks, they may reflect a placebo demand effect rather than a genuine change in unconscious cognition, indicating that reward processing is not a relevant mechanism in this intervention. Alternatively, the SHAPS questionnaire may be more likely to reflect consummatory anhedonia as opposed to reward learning. Another explanation for the disparity may lie in the length of treatment. While the PILT task has previously been shown as sensitive to the antidepressant drug bupropion (Walsh et al., 2018), this was found only after six weeks of administration and not at an early stage of treatment at just two weeks (when other measures of emotional cognition already showed effects). It is therefore possible that changes in performance on this task would emerge only after a longer course of BA treatment.

Speculatively, another reason for the lack of effect on reward processing could be that participants significantly vary in their baseline emotional cognition profiles (and the subsequent response of these profiles to treatment), which could lead to variance in findings from different treatment paradigms and outcome measures. Rather than being a unitary neurocognitive construct, affective bias may arise in different domains and with different magnitude and persistence based on each person's unique combination of emotional symptoms, as well as their causes, chronicity, and variability. For example, as discussed in Chapter 1, several models have now proposed the existence of core cognitive schemata that may drive automatic cognition and emotion. However, it is far from clear how these schemata arise, how stable they are, whether they should be viewed as a categorical or a dimensional construct and how they are affected by different types of treatment. Such variance could also explain why research studies on emotional cognition, including our present experiment, tend to find effects on specific emotional cognitive domains rather than in all domains at the same time.

One possible solution to this would be to recruit participants based on high levels of self-reported anhedonia at baseline. It has now been numerously proposed that anhedonia may form a separate class of depressive disorder, since not all patients with MDD experience such symptoms and those who do seem to differ in neurobiology as well as in treatment response (Treadway & Zald, 2011). Moreover, anhedonia itself may require a more granular operationalisation defined by specific reward processing deficits (see Chapter 1). It is possible that our study participants differed in their level of baseline “impairment” across different reward processing domains, which could explain why we didn’t find a clear effect on our measure of probabilistic reward learning. Future research could perhaps recruit participants based on a specific reward processing profile to closely examine the neurocognitive effects of BA treatment on this domain.

Possible solutions for future research could include closer linking of different emotional cognitive domains to specific depressive symptoms and constructing a detailed mechanistic model of how they arise, relate, and change with treatment. A hypothetical example could be a patient reporting the symptom of anticipatory anhedonia, who could consequently struggle with “wanting” to perform well on a probabilistic reward learning task. This motivational deficit would prevent a cognitive improvement on this measure. However, if they didn’t have consummatory anhedonia, they could still take pleasure in the activities they performed as part of BA, since the techniques help patients circumvent struggles with motivation, and thus their depression could improve.

The individual symptom profile is therefore likely to influence both the response to the intervention as well as to the cognitive tasks. Such granularity gets lost when we relate emotional cognitive measures to summation scores from symptom scales, since the same summation score can mean a completely different set of symptoms for each participant. Future research could thus pay closer attention to specific symptom profiles to explain the variability in emotional cognitive findings in psychological treatment research. This approach would be in line with recent proposals to focus on symptom networks in an idiographic manner, rather than assuming a latent depression construct, which may be too variable to serve as a useful phenotype (Fried and Nesse, 2016).

Strengths and limitations

The strength of this analysis includes our use of an established emotional cognitive battery that has been validated in a large body of existing research. The use of objective behavioural tasks should reduce placebo and demand effects that can arise in self-report measures (Huneke et al., 2017) and therefore offers a new possibility of validating psychological treatment efficacy. Our randomised design aimed to prevent confounding effects by third variables and the comparison with a passive control group allowed us to control for spontaneous recovery and regression to the mean. This is particularly important in the context of the pandemic, where external factors affecting people's mental health (e.g., whether they can exercise or interact with family and friends) may fluctuate significantly.

Our use of an online platform for cognitive testing allowed all participants to carry out data collection in a uniform and blinded way and also enabled us to test a more representative sample from all over the UK. However, we could not control the exact conditions under which the cognitive tasks were completed. While it is possible that participants may have been more distracted when performing tasks at home, they also could have been more comfortable and relaxed than if they were being observed by a researcher in a lab environment. It is therefore unclear how the online task administration affected participants' focus and motivation, but we aimed to maximise task validity by emphasising instructions to each participant before each data collection and including comprehension checks.

It is important to bear in mind that correlations between depression symptoms and emotional cognitive outcomes were performed in an exploratory fashion. As our study was not powered enough to detect a between-group difference in correlations, we only examined this association in the BA group to explore a possible mechanism of treatment response. Future studies should investigate this relationship in a larger sample sufficiently powered for a mediation analysis. It could also be argued that the study was limited by not correcting for multiple comparisons; since our outcome measures were chosen a-priori, we considered such a correction to be overly conservative in this context.

Lastly, there are several factors that may have affected participants' emotional cognitive profiles that we weren't able to control for in this study. For example, the magnitude and stability of

affective bias may be influenced by chronicity of depressive symptoms, which has been linked to younger age of onset, longer duration of episodes and family history of mental illness (Holzel et al., 2011). We did not take these factors into account in our analysis. It may be that our sample included participants with chronically recurring low mood, as well as participants whose depression was more “situational” due to the extraordinary circumstances of the COVID-19 pandemic. While we were not able to control for these factors in our analysis, future research could explore how this affects the emotional cognitive profile of patients with depression to further elucidate the nature of the emotional cognition construct.

Conclusion of Chapter 3

This chapter presented first evidence of a behavioural treatment for depression leading to a more positive affective bias in the domain of facial emotion recognition, which has previously been found mainly in pharmacological interventions. Emotional cognitive tests may provide a way to test the efficacy of psychological treatments objectively as well as highlighting a potential mechanism by which their clinical effects occur. While we did not find effects on our other measures of affective bias, hypotheses have been put forward to explain this and to suggest future research that could further elucidate how emotional cognition interacts with different depression symptoms.

Chapter 4: Examining the effects of behavioural activation on mood in comparison to an active and a passive control group

Introduction

In Chapter 2, a 4-week BA programme delivered by non-specialists was shown to reduce self-reported symptoms of depression and increase activation when compared to a passive control group. The use of a design whereby participants were randomised to either active or passive conditions allowed a degree of experimental control over a number of non-specific factors, including: spontaneous recovery with time; repeated testing and regression to the mean; and other non-specific mood-altering factors (such as the weather, changes in lockdown restrictions etc). However, to fully ensure the validity of the treatment protocol, it is also important to examine the effectiveness of its individual components. The present study further addresses these factors to extend our understanding of the intervention.

Importantly, it is currently unclear how much of the effects of BA are driven by its “active ingredients”, such as psychoeducation, activity grading and goal setting, as opposed to the so-called “non-specific factors” of psychotherapy. These non-specific factors, also known as the “common factors” or the “psychological placebo effect”, include the patient’s beliefs and expectations, the therapeutic context, as well as the therapeutic alliance. The three major meta-analyses of BA (Cuijpers et al., 2007; Mazzucchelli et al., 2009; Ekers et al., 2014) reported that the majority of studies have used waitlist or treatment-as-usual control groups, where patients’ expectations for improvement are likely to be significantly lower than in the active condition, and where no therapeutic context or alliance is provided. This lack of control over non-specific factors may undermine the current body of BA evidence.

Non-specific factors are widely acknowledged to be important (Bjornsson, 2011) and some have even argued that they explain most of the variance in psychological treatments (Messer and Wampold, 2002). These conclusions tend to be based on findings that show different psychological treatments to be similarly effective despite differences in their specific content. However, these similarities could also be explained by different interventions being effective through different mechanisms, since, for example, BA, antidepressants, and mindfulness can be

effective for depression but could work in different ways. Moreover, individual patient characteristics, which are not evident in group-level analysis, may interact with treatment type. As a result, different treatments could be effective for different patients due to their specific factors, rather than all being similarly effective for everyone due to non-specific factors.

There is some evidence to support the idea that it is the specific content of BA, rather than non-specific contextual factors, that are primarily responsible for its therapeutic effects. The assumed importance of active ingredients has been at the core of the BA treatment development and dissemination since its simple behavioural techniques can be administered by junior mental health workers or even non-specialists. Additionally, some evidence suggests that BA interventions may outperform cognitive techniques, such as restructuring, when predicting clinical improvement (Webb et al., 2019), despite similar levels of non-specific factors. These findings support the importance of specific BA techniques in its effectiveness for treating depression.

The use of an active control group allows some of the non-specific components (such as therapist contact) to be kept equivalent across groups and therefore the effect of specific aspects of an intervention to be elucidated. In past psychotherapy research, active control groups have ranged from reflective listening (Borkovec and Costello, 1993), “non-directive” therapy (Shear et al., 1994), placebo pills (Klein, 1997) or more recently, the befriending programme (Freeman et al., 2021). If such a control group can establish positive expectations and a therapeutic alliance in a credible context, it allows us to test the importance of specific factors that should only be present in the psychological intervention of choice.

Another possibility when devising an active control group is to only use a small subset of the specific techniques used in the intervention, while keeping the non-specific factors constant. If specific factors play an important role, the intervention which utilises them significantly more should be more beneficial. If non-specific factors have stronger influence over the outcome, we shouldn’t see too much of a difference between the two groups. When investigating BA, a good candidate for such a control condition could be activity monitoring, which is a credible technique that is likely to establish good patient expectancy. At the same time, it does not impart any techniques directly related to behaviour change, which would be expected to play the most important role within the BA model of specific factors.

In this randomised controlled study, we investigated whether a 4-week course of non-specialist administered BA can reduce depression and increase activation in comparison to a passive control group (PC), as well as an active control group carrying out activity monitoring (AM) only. We hypothesised that BA would have the largest effect on depression and activation, AM would show a lesser effect and PC would experience the least significant change in symptoms. As secondary outcomes for exploratory analysis, we also examined symptoms of anxiety, levels of social support and environmental reward, where we similarly expected BA to have the largest beneficial effect, followed by AM and then PC.

The main aims of the study were to test the efficacy of the intervention as well as examining its possible mechanisms of effect. The mechanistic analysis, which examined participants' emotional cognition and actigraphy measurements, will be presented in Chapters 5 and 6.

Methods

Participants

One hundred and one participants were recruited online via advertising on social media, University of Oxford's Department of Psychiatry website, the MQ charity website and Oxford's Daily Info platform. Three participants dropped out after baseline testing (all from the activity monitoring group) and the final sample included 77 women, 20 men and 1 who preferred not to disclose gender, see Fig. 14 below. To be recruited into the study, participants had to score at least 10 points on the BDI-2 but no more than 20 points on the PHQ-9 (based on the PHQ-9 guidelines recommending that anyone over the score of 20 should be treated with pharmacotherapy under clinical supervision). All participants were aged between 18 and 65 years.

Exclusion criteria included past participation in other studies using the Emotional Testing Battery, frequent use of recreational drugs (once a month or more), undergoing any other psychological treatment, having a past or present diagnosis of psychosis or bipolar disorder, or having a current diagnosis of an eating disorder, borderline personality disorder or a substance abuse disorder. Participants could be taking antidepressant medication if they had been taking it for 3

months or more and hadn't changed dosage in the past month. Any other medication or condition was evaluated on a case-by-case basis by a study medic.

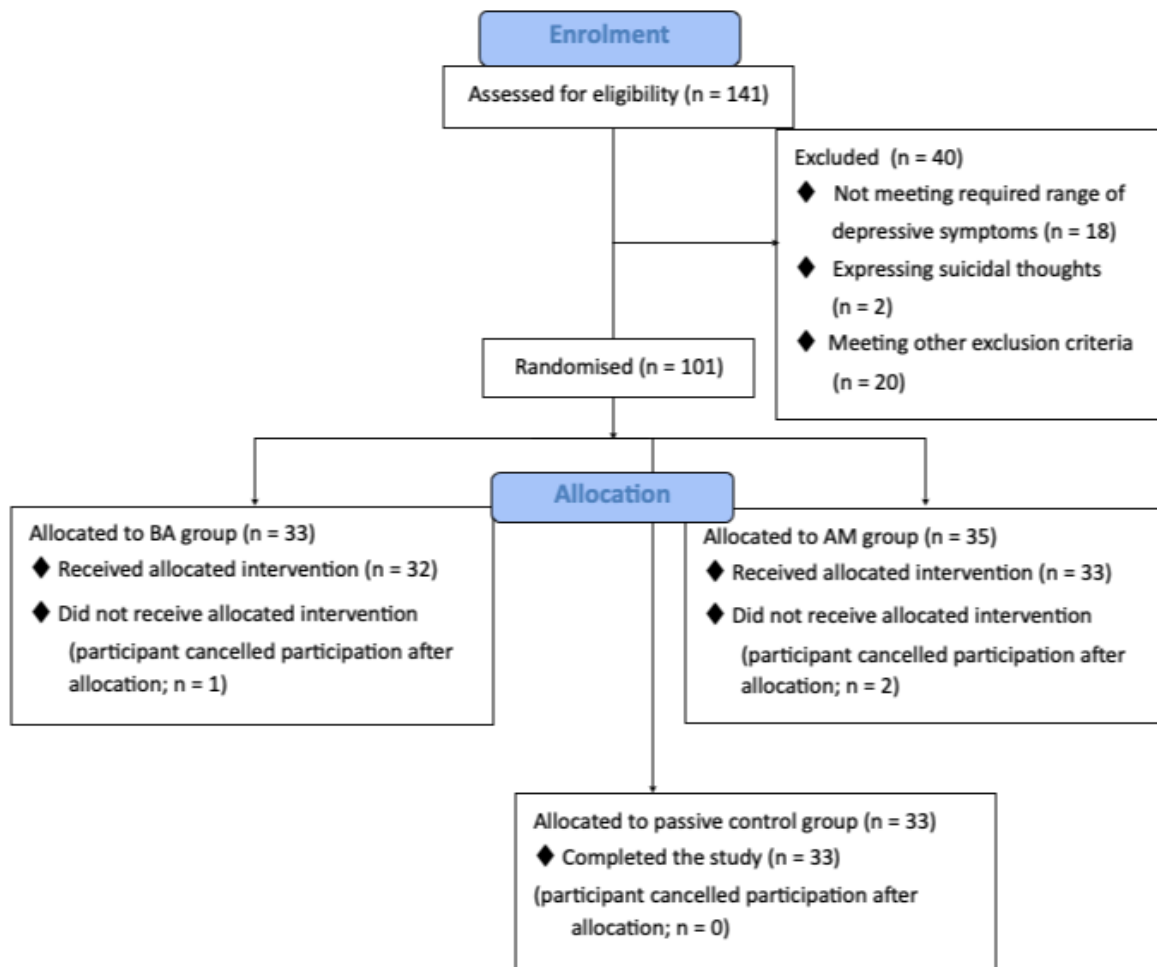


Figure 14. A flowchart showing the numbers of participants assessed for eligibility through the SCID interview as well as the numbers of participants excluded and randomised.

Power calculation

As described in Chapter 2, a meta-analysis by Ekers et al. (2014) found the average effect size of BA interventions in reducing depression symptoms to be medium-to-large (Hedge's g of 0.75). Correcting for multiple pairwise comparisons in the context of three study groups ($p = .017$), we would need a sample size of 38 to detect a difference between two independent means with an effect size of 0.7 (Cohen's d) with 80% power. As outlined in Chapter 1, we were not able to recruit our originally intended sample size due to disruptions caused by the COVID-19 pandemic.

Procedure

The recruitment was carried out in two phases. The first phase took place between April 2019 and February 2020. At that point, the study was suspended between March 2020 and April 2021 due to the COVID-19 pandemic. The second phase took place between May 2021 and September 2021, at which point these changes were made to the protocol:

- Participants were not able to join the study if they were currently isolating due to COVID-19, diagnosed with COVID-19 or experiencing any symptoms
- A new COVID-19-specific information sheet was given to participants outlining additional infection control measures implemented in the building
- All screening and intervention meetings took place remotely using Microsoft Teams
- Participants were asked to wear a face covering during all data collection sessions and have their temperature checked before entering the research building
- The pre-screening and screening procedures remained the same as in the previous study described in Chapter 2. Participants were informed that the study was comparing two behavioural interventions focused on activity and mood. Eligible participants were invited into Oxford's Department of Psychiatry to complete baseline demographic and questionnaire measures (BDI-2, BADS, STAI, EROS, MSPSS).

Participants were randomly allocated to either the BA intervention, the activity monitoring group (AM) or to a passive control group (PC) and randomisation was stratified by sex. Participants were informed about their group allocation after the first baseline week (week 0) during which baseline actigraphy and Mood Zoom measurements were being taken. The rest of the study schedule is outlined in Figure 15. *As opposed to the online study presented in Chapter 2, this study protocol included an extra baseline week (week 0) to collect baseline actigraphy data.* Data collection sessions were carried out by research assistants who were blinded to group allocation.

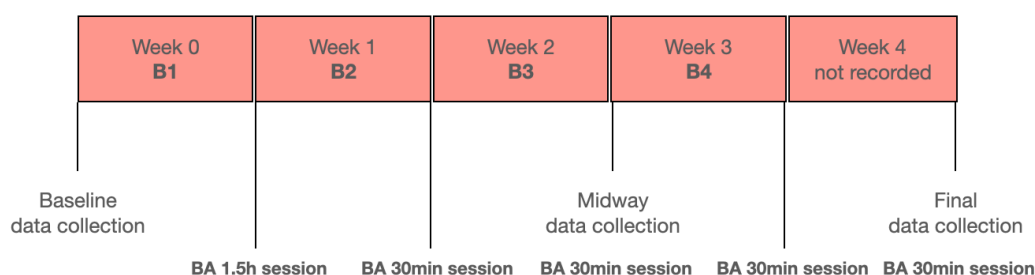


Figure 15. The study schedule outlining the data collection sessions as well as the intervention sessions applying to BA and AM groups only.

Interventions

Behavioural activation

The BA intervention was the same as described in Chapter 2, including psychoeducation, activity monitoring, goal setting, activity grading and problem solving (for more details see Table 3 below).

Activity monitoring

The AM group had an initial introductory session explaining the protocol, carried out by the same researcher who provided the BA intervention (the author). Participants were instructed to carry out daily activity monitoring using the same monitoring protocol as BA participants. They were instructed to reflect on how their activities link to their mood each day and to write down any insights into the comment box on their monitoring sheets. *As in the BA group, participants were informed that full participation in every session of the intervention and weekly submission of activity monitoring was required to receive the full study reimbursement, and all participants (except for the three who dropped out) fulfilled this requirement.*

The researcher explained that they would not be making any evaluations or suggestions in relation to how the participant spends their time. If the participants wanted to make changes to their activities based on insights gained while monitoring, they were encouraged to do so. Every week, participants had a check-in meeting with the practitioner to present their monitoring and to discuss any insights they gained. The two interventions are compared in detail below in Table

20.

Table 20. Comparison of components in the behavioural activation and activity monitoring interventions.

	Behavioural activation group	Activity monitoring group
Length of initial session to explain protocol	60-90 minutes	15-30 minutes
Psychoeducation about low mood and activity	Yes	No
Daily activity monitoring	Yes	Yes
Explanation of routine, pleasurable and necessary activities	Yes	No
Goal setting	Yes	No
Activity grading	Yes	No
Explanation of strategies to overcome low mood and rumination	Yes	No
Length of weekly check-in	15-30 minutes	5-15 minutes
Problem-solving	Yes	No
Accountability for activity completion	Yes	No
Relapse prevention strategies	Yes	No

Passive control group

The control group did not receive any intervention but received materials about BA at the end of the study, which included the intervention booklet, a video about a patient who is undergoing BA as well as resources on where to find BA in their area.

All participants were compensated with £100 for their full participation together with travel expenses and any amount won in the Probabilistic and Instrumental Learning Task (see Chapter 5).

Questionnaire measures

We measured depression symptoms using the Beck Depression Inventory 2 (BDI-2; Beck, Steer, & Brown, 1996), anxiety using the State-Trait Anxiety Inventory (STAI; Spielberger et al., 1970), activation using the long form of the Behavioural Activation for Depression Scale (BADS; Kanter

et al., 2007) and social support using the Multidimensional Scale of Perceived Social Support (MSPSS; Zimet et al., 1988). We also sent out electronic daily Mood Zoom questionnaires (Tsanas et al., 2016) via text. Psychometric details on these instruments can be found in Chapter 2.

Additionally, we measured environmental reinforcement using the Environmental Reward Observation Scale (EROS, Armento & Hopko, 2007). This is a 10-item scale designed to assess the magnitude and availability of positive reinforcement in the person's environment (e.g. "It is easy for me to find enjoyment in life") as well as the individual's ability to elicit environmental reward themselves ("The activities I engage in usually have positive consequences"). It has been previously found that the level of increased environmental reward strongly correlated with the decrease in depression and anxiety levels and an increase in levels of social support (Gawrysiak et al., 2009).

Statistical analysis

All questionnaire measures were analysed using 3x3 two-way mixed ANOVAs with group as a between-subject factor (BA, AM, and PC) and time as a within-subject factor (baseline, midway and final). Mood Zoom data were analysed using a 6x3 two-way mixed ANOVAs with group as a between-subject factor and emotion (happy, energetic, sad, anxious, and irritable) as a within-subject factor. Mood Zoom data were analysed as change-from-baseline scores at week 3 and week 4 separately due to significant differences between groups at baseline (see Results).

Significant interactions were followed up with simple main effect analyses; independent t-tests were used for simple main effects of group and pairwise comparisons, separated by group, were used for simple main effects of time.

Normality was assessed using the Shapiro–Wilk test. Outliers were assessed by the inspection of box plots and studentised residuals; they were removed if they were 3 box-lengths away from the edge of the box or if the studentised residual was greater than ± 3 standard deviations. Mauchly's test was used to assess the assumption of sphericity; if sphericity was violated, Greenhouse–Geisser correction was applied. Box's test was used to assess the equality of covariance.

All data were analysed SPSS software. For a sensitivity analysis, all analyses were compared with and without the 33 participants who were taking antidepressants to investigate whether any effects may have been driven by the concurrent pharmacological treatment.

Results

Demographic and clinical baseline information

The three groups were moderately well matched in their baseline characteristics (see Table 21). There appeared to be differences in the rates of antidepressant treatment, living arrangements, marital status and estimates of major depressive disorder. Diagnostic categories were assessed using the Structured Clinical Interview for DSM-5.

Table 21. Demographic and clinical baseline characteristics

Variable	BA group (n = 32)	AM group (n = 33)	Control group (n = 33)
Age (mean, SD)	33.91 (13.33)	30.85 (11.77)	36.85 (13.35)
Years of full-time education (mean, SD)	17.09 (3.80)	16.58 (3.52)	16.88 (3.53)
Gender	Female 78.1% Male 21.9%	Female 75.8% Male 21.2% Prefer not to say 3%	Female 81.8% Male 18.2%
Highest education level attained	GCSE 6.3% A-level / GNVQ 31.3% Undergraduate 34.4% Postgraduate 28.1%	GCSE 6.1% A-level / GNVQ 33.3% Undergraduate 27.3% Postgraduate 33.3%	GCSE 6.1% A-level / GNVQ 24.2% Undergraduate 24.2% Postgraduate 45.5%
Ethnicity	Asian 12.5% Black 0% Caucasian 78.1% Mixed 6.3% Other 3.1%	Asian 9.1% Black 3% Caucasian 78.8% Mixed 6.1% Other 3%	Asian 6.1% Black 6.1% Caucasian 75.8% Mixed 6.1% Other 6.1%
Employment status	Student 43.8% Employed 50% Unemployed 6.3% Retired 0%	Student 51.5% Employed 42.4% Unemployed 6.1% Retired 0%	Student 27.3% Employed 57.6% Unemployed 12.1% Retired 3%

Native English speaker	Yes 81.3% No 18.8%	Yes 72.7% No 27.3%	Yes 66.7% No 33.3%
Living arrangements	Alone 12.5% Shared flat or house 37.5% With a partner/family 50% Other 0%	Alone 9.1% Shared flat or house 54.5% With a partner/family 30.3% Other 6.1%	Alone 18.2% Shared flat or house 30.3% With a partner/family 48.5% Other 3%
Marital status	Single 68.8% Married / Civil partner 25% Divorced / Civil partnership dissolved 6.3% Separated 0% Widowed 0%	Single 75.8% Married / Civil partner 21.2% Divorced / Civil partnership dissolved 0% Separated 3% Widowed 0%	Single 54.5% Married / Civil partner 36.4% Divorced / Civil partnership dissolved 3% Separated 0% Widowed 6.1%
Current antidepressants	Yes 21.9% No 78.1%	Yes 33.3% No 66.7%	Yes 45.5% No 54.5%
Shift worker	Yes 6.3% No 93.8%	Yes 6.1% No 93.9%	Yes 3% No 97%
Current major depressive episode	Yes 62.5% No 37.5%	Yes 42.4% No 57.6%	Yes 46.9% No 53.1%
Current persistent depressive disorder	Yes 31.3% No 68.8%	Yes 33.3% No 66.7%	Yes 53.1% No 46.9%
Current panic disorder	Yes 3.1% No 96.9%	Yes 3% No 97%	Yes 3.1% No 96.9%
Current agoraphobia	Yes 0% No 100%	Yes 6.1% No 93.9%	Yes 3.1% No 96.9%
Current social anxiety disorder	Yes 21.9% No 78.1%	Yes 12.1% No 87.9%	Yes 12.5% No 87.5%

Current generalised anxiety disorder	Yes 37.5% No 62.5%	Yes 27.3% No 72.7%	Yes 43.8% No 56.3%
Current obsessive-compulsive disorder	Yes 0% No 100%	Yes 0% No 100%	Yes 0% No 100%
Current post-traumatic stress disorder	Yes 3.1% No 96.9%	Yes 0% No 100%	Yes 6.3% No 93.8%

Questionnaire measures

Baseline descriptive statistics for all questionnaires can be found in the Appendix (Table 53).

Depression

A significant time x group interaction was found for depression scores as measured by the BDI-2

[$F(3.48, 161.71) = 10.11, p < .001, \eta^2 = 0.18$], as shown in Fig. 16.

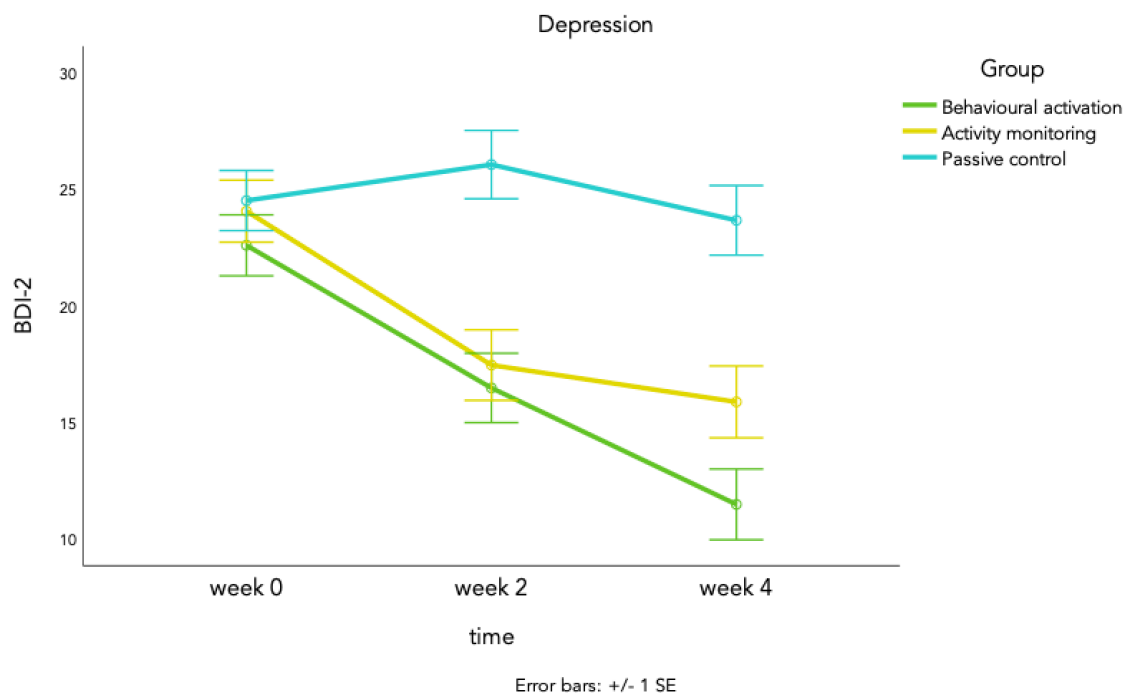


Figure 16. Depression scores as measured by the BDI-2 at baseline (week 0), halfway through the interventions (week 2) and at the end (week 4).

Simple main effect of group comparing the BA and AM conditions revealed a significant difference at the final time point (week 4), with BA showing significantly lower depression scores (see Table 22).

Table 22. Detailed statistics for post-hoc t-tests comparing the behavioural activation (BA) and activity monitoring groups (AM) in depression scores. One asterisk indicates $p < .05$.

BA vs AM	T-test statistics	Group M and SD
Week 0	$t(63) = -0.70, p = .48, d = 0.18$	BA: $M = 22.56, SD = 7.01$ AM: $M = 23.82, SD = 7.36$
Week 2	$t(63) = -0.67, p = .51, d = 0.17$	BA: $M = 16.44, SD = 8.45$ AM: $M = 17.76, SD = 7.52$
Week 4 *	$t(61) = -2.24, p = .03, d = 0.56$	BA: $M = 11.44, SD = 7.57$ AM: $M = 15.84, SD = 8.05$

When comparing the BA and PC conditions, there was a significant difference both at the midway (week 2) and final (week 4) time points, with BA showing significantly lower depression scores at both times (see Table 23).

Table 23. Detailed statistics for post-hoc t-tests comparing the behavioural activation (BA) and passive control (PC) groups in depression scores. One asterisk indicates $p < .05$, two asterisks indicate $p < .001$

BA vs PC	T-test statistics	Group M and SD
Week 0	$t(63) = -1.03, p = .31, d = 0.25$	BA: $M = 22.56, SD = 7.01$ PC: $M = 24.48, SD = 8.03$
Week 2 **	$t(63) = -4.26, p < .001, d = 1.06$	BA: $M = 16.44, SD = 8.45$ PC: $M = 26.03, SD = 9.65$
Week 4 **	$t(63) = -5.57, p < .001, d = 1.38$	BA: $M = 11.44, SD = 7.57$ PC: $M = 23.64, SD = 9.90$

When comparing the AM and PC conditions, there was again a significant difference both at the midway and final time points, with the AM group showing significantly lower depression scores in both cases (see Table 24).

Table 24. Detailed statistics for post-hoc t-tests comparing the activity monitoring and passive control groups in depression scores. One asterisk indicates $p < .05$, two asterisks indicate $p < .001$

AM vs PC	T-test statistics	Group M and SD
Week 0	$t(64) = -0.35, p = .73, d = 0.09$	AM: $M = 23.82, SD = 7.36$ PC: $M = 24.48, SD = 8.03$
Week 2 **	$t(64) = -3.89, p < .001, d = 0.96$	AM: $M = 17.76, SD = 7.52$ PC: $M = 26.03, SD = 9.65$
Week 4 *	$t(62) = -3.45, p = .001, d = 0.87$	AM: $M = 15.84, SD = 8.05$ PC: $M = 23.64, SD = 9.90$

All of these findings persisted after excluding participants taking antidepressant medications.

Activation

There was a significant time by group interaction for activation [$F(3.42, 155.57) = 7.02, p < .001, \eta^2 = 0.13$], see Fig. 17.

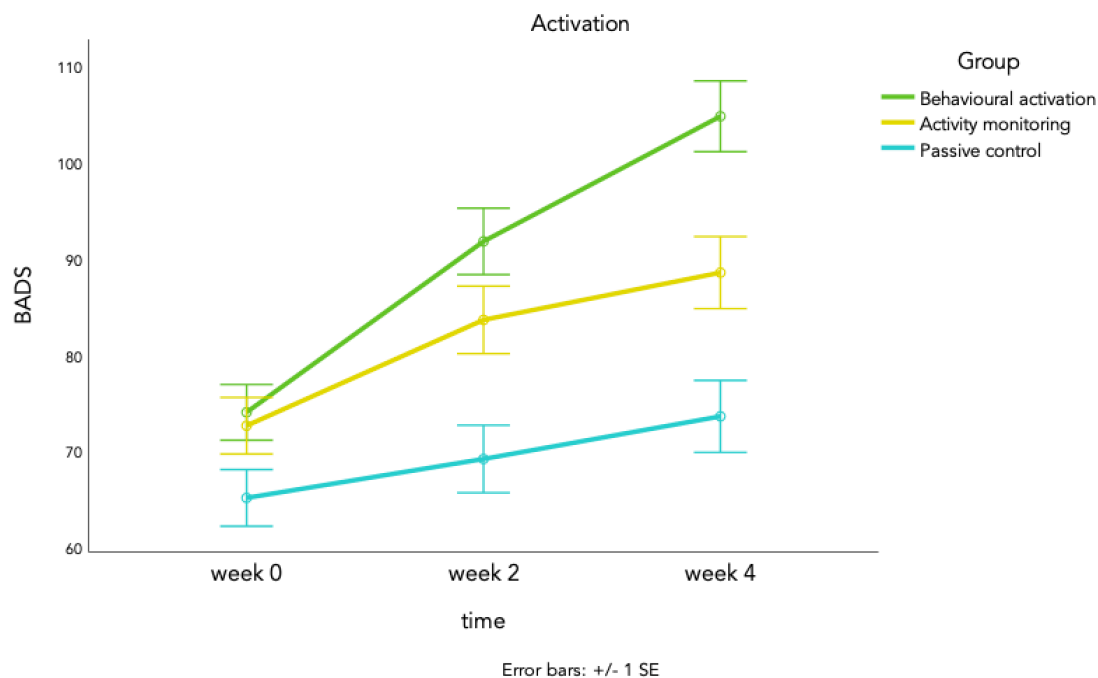


Figure 17. Activation scores as measured by the BADS at baseline (week 0), halfway through the interventions (week 2) and at the end (week 4).

Simple main effect of group analysis comparing BA and AM revealed a significant difference at the final time point, with BA showing significantly higher activation scores (see Table 25).

Table 25. Detailed statistics for post-hoc t-tests comparing the behavioural activation (BA) and activity monitoring (AM) groups in activation levels. Two asterisks indicate $p < .001$

BA vs AM	T-test statistics	Group M and SD
Week 0	$t(63) = 0.33, p = .75, d = 0.08$	BA: $M = 74.00, SD = 16.48$ AM: $M = 72.67, SD = 16.59$
Week 2	$t(62) = 1.87, p = .07, d = 0.47$	BA: $M = 91.72, SD = 18.30$ AM: $M = 82.38, SD = 21.53$
Week 4 **	$t(62) = 3.36, p = .001, d = 0.84$	BA: $M = 104.69, SD = 17.55$ AM: $M = 87.75, SD = 22.49$

When comparing BA and PC, there was a significant difference at the baseline, midway and final time points, with BA showing significantly higher activation scores (see Table 26).

Table 26. Detailed statistics for post-hoc t-tests comparing the behavioural activation (BA) and passive control (PC) groups in activation levels. One asterisk indicates $p < .05$, two asterisks indicate $p < .001$

BA vs PC	T-test statistics	Group M and SD
Week 0 *	$t(61) = 2.19, p = .03, d = 0.55$	BA: $M = 74.00, SD = 16.48$ PC: $M = 65.13, SD = 15.70$
Week 2 **	$t(63) = 4.08, p < .001, d = 1.02$	BA: $M = 91.72, SD = 18.30$ PC: $M = 71.48, SD = 21.47$
Week 4 **	$t(63) = 5.46, p < .001, d = 1.36$	BA: $M = 104.69, SD = 17.55$ PC: $M = 76.15, SD = 24.00$

When comparing AM and PC groups, there was a trend towards a significant difference at all time points, with AM showing higher activation scores (see Table 27).

Table 27. Detailed statistics for post-hoc t-tests comparing the activity monitoring (AM) and passive control (PC) groups in activation levels.

AM vs PC	T-test statistics	Group M and SD
----------	-------------------	----------------

Week 0	$t(62) = 1.86, p = .07., d = 0.47$	AM: $M = 72.67, SD = 16.59$ PC: $M = 65.13, SD = 15.70$
Week 2	$t(63) = 2.04, p = .05, d = 0.51$	AM: $M = 82.38, SD = 21.53$ PC: $M = 71.48, SD = 21.47$
Week 4	$t(63) = 2.01, p = .05, d = 0.51$	AM: $M = 87.75, SD = 22.49$ PC: $M = 76.15, SD = 24.00$

After excluding participants taking antidepressants, there was no longer a significant difference between the BA and PC groups at baseline. Moreover, a significant difference emerged between the AM and PC groups at final time point. All other findings remained the same.

State anxiety

There was a significant time x group interaction for state anxiety as measured by the STAI-S [$F(4, 184) = 3.55, p = .008, \eta^2 = .07$] see Fig. 18.

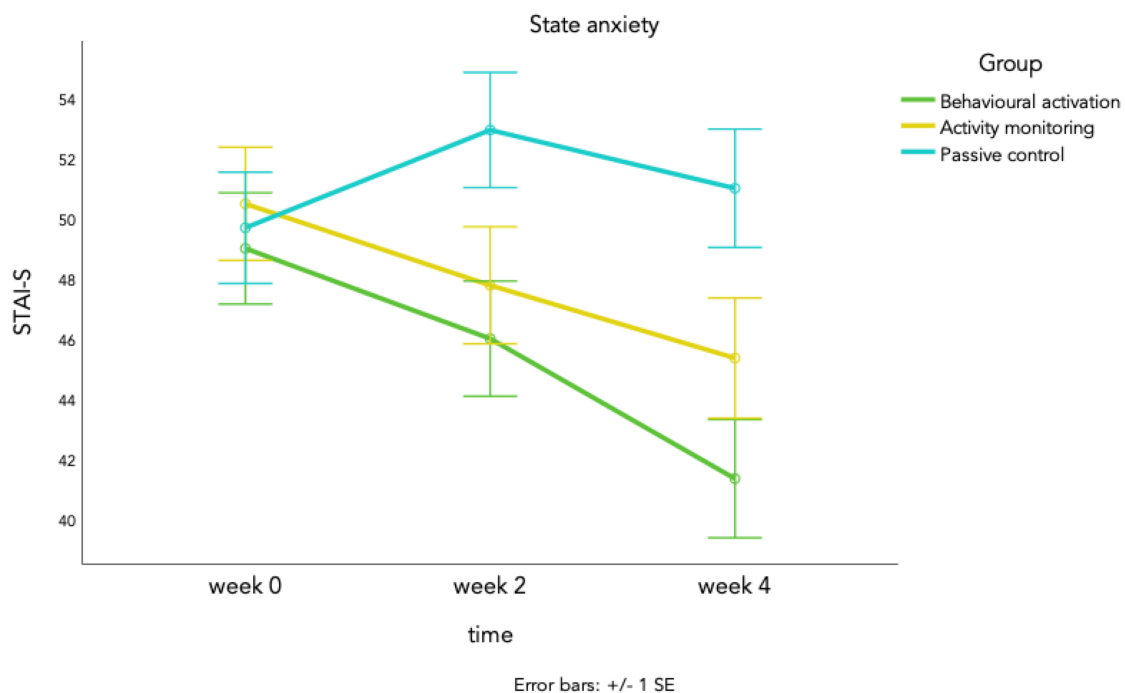


Figure 18. State anxiety scores as measured by the STAI-S at baseline (week 0), halfway through the interventions (week 2) and at the end (week 4).

When comparing BA and AM groups, there were no significant differences at any time point, [see Table 28](#).

Table 28. Detailed statistics for post-hoc t-tests comparing the behavioural activation (BA) and activity monitoring (AM) groups in state anxiety levels.

BA vs AM	T-test statistics	Group M and SD
Week 0	$t(63) = -0.79, p = .44, d = 0.20$	BA: $M = 48.97, SD = 9.79$ AM: $M = 50.85, SD = 9.50$
Week 2	$t(62) = -0.88, p = .38, d = 0.22$	BA: $M = 45.97, SD = 11.33$ AM: $M = 48.38, SD = 10.54$
Week 4	$t(62) = -1.68, p = .10, d = 0.42$	BA: $M = 41.31, SD = 8.46$ AM: $M = 45.56, SD = 11.53$

When comparing BA and PC, there was a significant difference at the midway point with the BA group being significantly lower in state anxiety than the passive control. The significant difference remained at final time point with BA again showing lower state anxiety than the passive control group, [see Table 29](#).

Table 29. Detailed statistics for post-hoc t-tests comparing the behavioural activation (BA) and passive control (PC) groups in state anxiety levels.

BA vs PC	T-test statistics	Group M and SD
Week 0	$t(62) = -0.25, p = .80, d = 0.06$	BA: $M = 48.97, SD = 9.79$ PC: $M = 49.66, SD = 11.80$
Week 2 *	$t(63) = -2.48, p = .02, d = 0.61$	BA: $M = 45.97, SD = 11.33$ PC: $M = 52.79, SD = 10.87$
Week 4 *	$t(55.85) = -3.50, p = .001, d = 0.87$	BA: $M = 41.31, SD = 8.46$ PC: $M = 50.73, SD = 12.72$

When comparing the AM and PC groups, there were no significant group differences at any time point ($p > .05$). After excluding participants taking antidepressants, a significant difference emerged between the AM and PC groups at the final time point, [see Table 30](#).

Table 30. Detailed statistics for post-hoc t-tests comparing the activity monitoring (AM) and passive control (PC) groups in state anxiety levels.

AM vs PC	T-test statistics	Group M and SD
Week 0	$t(63) = 0.45, p = .66, d = 0.11$	AM: $M = 50.85, SD = 9.50$ PC: $M = 49.66, SD = 11.80$
Week 2	$t(63) = -1.66, p = .10, d = 0.41$	AM: $M = 48.38, SD = 10.54$ PC: $M = 52.79, SD = 10.87$
Week 4	$t(63) = -1.71, p = .09, d = 0.43$	AM: $M = 45.56, SD = 11.53$ PC: $M = 50.73, SD = 12.72$

Trait anxiety

There was a significant time x group interaction for trait anxiety as measured by the STAI-T [$F(3.46, 155.49) = 6.69, p < .001, \eta^2 = .13$], see Fig. 19 below.

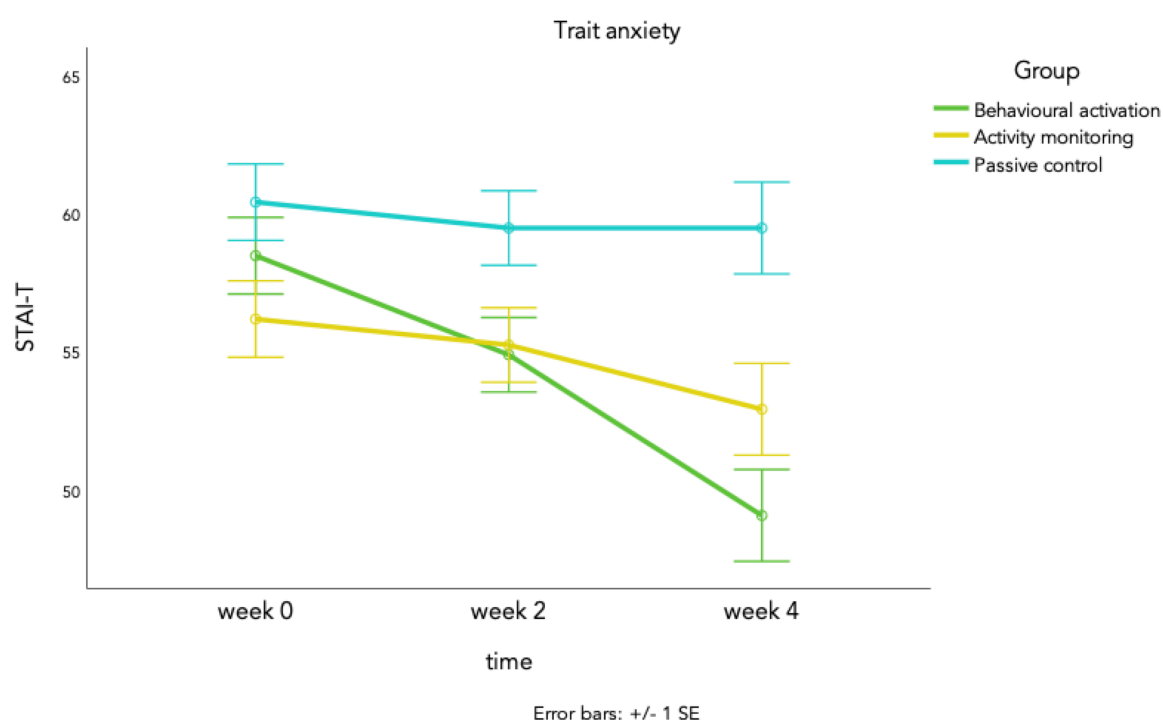


Figure 19. Trait anxiety levels as measured by the STAI-T at baseline (week 0), halfway through the interventions (week 2) and at the end (week 4).

When comparing trait anxiety between the BA and AM groups, there were no significant differences between groups, see Table 31.

Table 31. Detailed statistics for post-hoc t-tests comparing the behavioural activation and activity monitoring groups in trait anxiety levels.

BA vs AM	T-test statistics	Group M and SD
Week 0	$t(63) = 0.98, p = .33, d = 0.24$	BA: $M = 58.13, SD = 8.52$ AM: $M = 56.24, SD = 6.95$
Week 2	$t(61) = -0.38, p = .71, d = 0.1$	BA: $M = 54.87, SD = 7.90$ AM: $M = 55.59, SD = 7.16$
Week 4	$t(62) = -1.96, p = .05, d = 0.49$	BA: $M = 48.50, SD = 9.02$ AM: $M = 52.88, SD = 8.81$

When comparing BA with the PC group, there was a significant difference both at the midway and final time points, with the BA group showing significantly lower trait anxiety, see Table 32.

Table 32. Detailed statistics for post-hoc t-tests comparing the behavioural activation and passive control groups in trait anxiety levels. One asterisk indicates $p < .05$, two asterisks indicate $p < .001$

BA vs PC	T-test statistics	Group M and SD
Week 0	$t(61) = -1.11, p = .27, d = 0.28$	BA: $M = 58.13, SD = 8.52$ PC: $M = 60.39, SD = 7.63$
Week 2 *	$t(61) = -2.31, p = .02, d = 0.59$	BA: $M = 54.87, SD = 7.90$ PC: $M = 59.34, SD = 7.49$
Week 4 **	$t(62) = -4.58, p < .001, d = 1.15$	BA: $M = 48.50, SD = 9.02$ PC: $M = 59.38, SD = 9.96$

When comparing the AM and PC groups, there were significant differences at both baseline, midway and final time points, with AM showing lower trait anxiety in all cases, see Table 33.

Table 33. Detailed statistics for post-hoc t-tests comparing the activity monitoring and passive control groups in trait anxiety levels. One asterisk indicates $p < .05$.

AM vs PC	T-test statistics	Group M and SD
Week 0 *	$t(62) = -2.27, p = .03, d = 0.57$	AM: $M = 56.24, SD = 6.95$ PC: $M = 60.39, SD = 7.63$
Week 2	$t(62) = -2.05, p = .05, d = 0.51$	AM: $M = 55.59, SD = 7.16$ PC: $M = 59.34, SD = 7.49$

Week 4 *	$t(62) = -2.77, p = .007, d = 0.69$	AM: $M = 52.88, SD = 8.81$ PC: $M = 59.38, SD = 9.96$
----------	-------------------------------------	--

After excluding participants taking antidepressants, there was no longer a significant difference between the AM and PC groups at baseline.

Social support

There was no significant time x group interaction for social support as measured by the MSPSS [$F(3.48, 160.24) = 1.03, p = .39, \eta^2 = .02$].

After excluding participants taking antidepressants, there was no longer a significant difference between baseline and final levels of social support across all participants.

Environmental reward

There was a significant time x group interaction for levels of environmental reward as measured by the EROS [$F(3.56, 162.13) = 5.86, p < .001, \eta^2 = .11$], see Fig. 20 below.

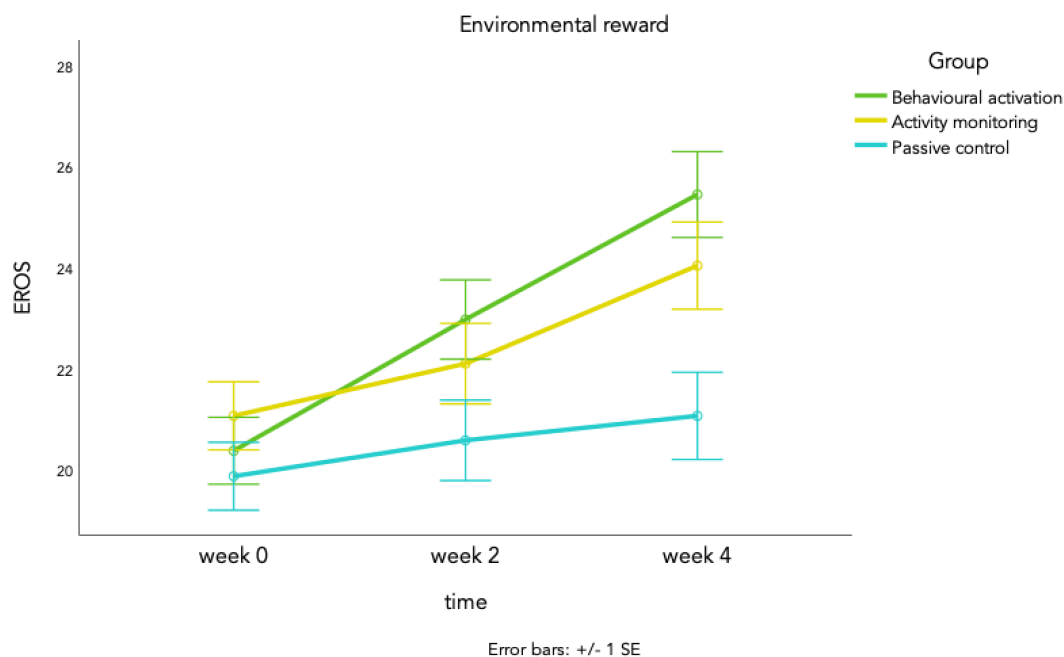


Figure 20. Environmental reward levels as measured by the EROS at baseline (week 0), halfway through the interventions (week 2) and at the end (week 4).

Simple main effect of group analysis did not reveal any significant difference between BA and AM groups at any time points, [see Table 34 below](#).

Table 34. Detailed statistics for post-hoc t-tests comparing the behavioural activation and activity monitoring groups in environmental reward levels.

BA vs AM	T-test statistics	Group M and SD
Week 0	$t(63) = -0.76, p = .45, d = 0.19$	BA: $M = 20.38, SD = 3.93$ AM: $M = 21.09, SD = 3.63$
Week 2	$t(63) = 1.02, p = .31, d = 0.26$	BA: $M = 22.97, SD = 4.11$ AM: $M = 21.88, SD = 4.43$
Week 4	$t(62) = 1.39, p = .17, d = 0.35$	BA: $M = 25.44, SD = 3.96$ AM: $M = 23.94, SD = 4.68$

When comparing BA with the PC group, there was a significant difference at the final time point with the BA group showing much higher environmental reward than the passive control group, [see Table 35 below](#).

Table 35. Detailed statistics for post-hoc t-tests comparing the behavioural activation and passive control groups in environmental reward levels. One asterisk indicates $p < .05$.

BA vs PC	T-test statistics	Group M and SD
Week 0	$t(61) = 0.53, p = .60, d = 0.14$	BA: $M = 20.38, SD = 3.93$ PC: $M = 19.87, SD = 3.58$
Week 2	$t(63) = 1.82, p = .07, d = 0.45$	BA: $M = 22.97, SD = 4.11$ PC: $M = 20.91, SD = 4.97$
Week 4 *	$t(62) = 3.59, p = .001, d = 0.90$	BA: $M = 25.44, SD = 3.96$ PC: $M = 21.13, SD = 5.52$

Furthermore, a significant difference was only found at the final time point when comparing the AM and the PC groups, with the AM group reporting higher environmental reward than the passive control group, [see Table 36 below](#).

Table 36. Detailed statistics for post-hoc t-tests comparing the activity monitoring and passive control groups in environmental reward levels. One asterisk indicates $p < .05$.

AM vs PC	T-test statistics	Group M and SD
Week 0	$t(62) = 1.35, p = .18, d = 0.34$	AM: $M = 21.09, SD = 3.63$ PC: $M = 19.87, SD = 3.58$
Week 2	$t(63) = 0.83, p = .41, d = 0.21$	AM: $M = 21.88, SD = 4.43$ PC: $M = 20.91, SD = 4.97$
Week 4 *	$t(62) = 2.20, p = .03, d = 0.55$	AM: $M = 23.94, SD = 4.68$ PC: $M = 21.13, SD = 5.52$

The BA group showed a significant negative correlation between early change in environmental reward and later change in depression symptoms [$r(28) = -0.37, p = .04$], with increases in environmental reward predicting decreases in depression.

After excluding participants taking antidepressants all results remained the same.

Mood Zoom

There were significant differences at baseline (average daily scores over week 0) between the BA and PC groups and between the AM and PC groups in the ratings of “angry” and “irritable” emotions, see Table 14 below.

Table 37. Baseline differences between the groups in daily ratings of “angry” and “irritable” emotions on the Mood Zoom questionnaire during the baseline week (week 0) of the study.

	T-test statistics	Group M and SD
BA vs PC: ratings of “angry”	$t(46.22) = -2.19, p = .03, d = 0.56$	BA: $M = 1.59, SD = 0.64$ PC: $M = 2.12, SD = 1.18$
BA vs PC: ratings of “irritable”	$t(52.54) = -2.16, p = .04, d = 0.54$	BA: $M = 2.30, SD = 0.92$ PC: $M = 2.96, SD = 1.45$
AM vs PC: ratings of “angry”	$t(47.22) = -2.14, p = .04, d = 0.30$	AM: $M = 1.60, SD = 0.68$ PC: $M = 2.12, SD = 1.19$

AM vs PC: ratings of “irritable”	$t(54.32) = -2.17, p = .03, d = 0.54$	AM: $M = 2.29, SD = 0.98$ PC: $M = 2.96, SD = 1.45$
----------------------------------	---------------------------------------	--

When analysing change-from-baseline scores at week 3 for individual emotions, there was a trend towards a significant interaction of emotion and group [$F(10, 440) = 1.73, p = .07, \eta^2 = .04$], see Fig.21.

Exploratory post-hoc t-tests revealed a significant difference in the rating of “energetic” between the BA and AM groups [$t(60) = 2.51, p = .02, d = 0.63$], with the BA group increasing significantly more ($M = 0.54, SD = 0.92$) than the AM group ($M = 0.05, SD = 0.60$).

There was also a significant difference in the rating of “energetic” between the BA and PC [$t(61) = 3.25, p = .002, d = 0.81$], again with the BA group increasing (BA: $M = 0.54, SD = 0.92$) and the control group decreasing ($M = -0.09, SD = 0.61$).

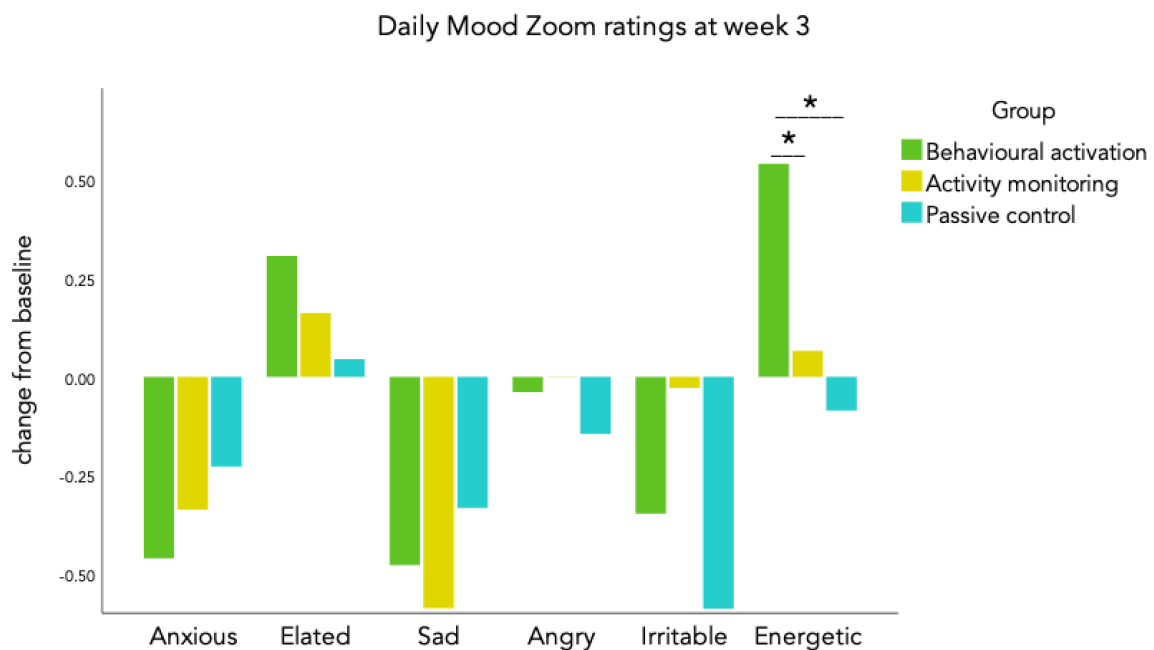


Figure 21. Change-from-baseline ratings of different emotions at week 3. One asterisk indicates $p < .05$.

After removing participants taking antidepressants, there was no longer a significant difference between BA and AM groups in the rating of “energetic” ($p = .11$). A new difference emerged in the rating of “irritable” [$t(44) = -2.64, p = .01, d = 0.78$], reflecting a decrease in the BA group ($M = -0.39, SD = 0.70$) but an increase in the AM group ($M = 0.14, SD = 0.66$).

When analysing change-from-baseline scores at week 4 for individual emotions, there was no significant interaction of emotion and group [$F(6.51, 253.70) = 1.42, p = .20, \eta^2 = .04$]. This persisted after removing participants taking antidepressants.

Discussion

This chapter investigated whether a 4-week course of non-specialist BA has a stronger effect on symptoms of depression and levels of activation than an active control group (carrying out only one of BA techniques, activity monitoring; AM) and a passive control group with no intervention (PC). As predicted, we found that BA led to the largest decrease in depression symptoms and the largest increase in activation levels, followed by AM and then PC groups, all significantly different from each other. Additionally, the BA intervention led to the lowest state and trait anxiety scores, the highest levels of environmental reward and the largest increase in daily ratings of feeling energetic.

These findings have largely replicated the effects on self-report measures from Chapter 2, further supporting the beneficial effects of a 4-week BA paradigm. Since BA led to greater decreases in depression symptoms than the AM intervention, despite their non-specific factors being similar, it is possible that the specific techniques used in BA have an important influence over its effectiveness. This would be in line with other findings supporting the importance of specific factors, such as the increasing evidence that smartphone apps can effectively deliver BA without any therapeutic relationship (Ly et al., 2014; Dahne et al., 2019). However, it is important to note that the AM group spent less time with the practitioner overall, since the process of this intervention was simpler and required less practitioner input and support. It is therefore possible that the difference in the amount of practitioner contact, which is a non-specific factor, contributed to the difference found between the two group’s results. Future

research should aim to ensure complete equivalence between the time spent with the practitioner of a psychological treatment and any active control group.

Importantly, our findings suggest that the AM protocol on its own may serve as a helpful intervention for decreasing depression and increasing activation, leading to moderate-to-large effect sizes in comparison with the PC group. While not as effective as the full BA protocol, it may be a more viable intervention when resources are low as it is even simpler for training and dissemination. From anecdotal participant insights, the mere act of monitoring may be helpful for gaining greater awareness of how one spends their time and what behavioural changes can be made to facilitate better wellbeing. Similar findings were reported in one experience sampling paradigm (Kramer et al., 2014), which provided participants with automated insights linking activities and affect. Nonetheless, some of our participants have also reported that activity monitoring without direct techniques for behaviour change was not helpful and only made them more aware of dissatisfying aspects of their lifestyle. Further research should explore whether such an intervention may have better effects if there is emphasis placed on monitoring positive activities specifically.

Interestingly, in an exploratory analysis, we found that early increases in environmental reward predicted later improvements in depression symptoms during the BA intervention. This indicates that the experience of environmental reinforcement may be another important mechanistic factor in the effectiveness of this intervention. This supports the theoretical model of BA, wherein positive activities are necessary for improvements in mood, creating a positive feedback loop wherein higher activation leads to better emotional outcomes. Future research should explore what aspects of patient's environment may be most important for BA efficacy and whether these could predict which patients will benefit from the intervention most significantly or most quickly.

Strengths and limitations

The strengths of this study include its randomised controlled design, the use of both an active and a passive control group as well as blinded data collection. We used an established BA intervention as well as validated questionnaires, replicating the protocol used in the Chapter 2 study.

Based on the descriptive data collected at baseline, our sample should be similar to samples in other studies investigating moderate low mood. The self-reported activation scores and levels of social support were similar to previous findings on depression (Kanter et al., 2009; Gladstone et al., 2007). The participants in this study showed somewhat lower state anxiety scores than previous findings in similar studies (Godlewska et al., 2012), but trait anxiety was generally the same (Fisher and Durham, 1999).

As opposed to the previous study, we did not collect follow-up data, so we weren't able to compare the two protocols directly in this regard. We also didn't collect COVID-related measures in our second phase of testing, so we couldn't account for any effects that the pandemic may have had on this part of our sample. Moreover, the two phases of our study recruitment differed in the delivery format of the interventions, which were carried out in-person during the first phase of the study, but remotely during the second phase. This may have introduced variability in the intervention effects.

We did not collect any ratings of the intervention or expectancy effects, so we weren't able to directly confirm that the non-specific factors of the two intervention groups were equivalent. While the two groups differed significantly in the amount of time spent with the intervention practitioner, participants were not aware of this difference. Nonetheless, this may have affected the influence of non-specific factors on the group and rendered the BA group with more therapeutic alliance. Direct assessment of expectations in both the intervention and active control groups has been recommended (Boot et al., 2013) and future research would benefit from this data.

Another possible weakness is that some significant differences were detected between our groups at baseline. Firstly, the BA and PC groups differed in their activation levels, with the PC group showing lower initial activity. Secondly, the PC group showed significantly higher ratings of "angry" and "irritable" emotions in the Mood Zoom questionnaire during the baseline week of the study, as well as showing higher baseline levels of trait anxiety when compared to the AM group only. All of these factors suggest the intervention groups may have been slightly "better off" in their emotional profiles at baseline, which may have aided their improvements in depression and anxiety symptoms over the course of the study. Another difference between the

groups that may have confounded our effects related to the levels of antidepressant treatment, living arrangements, marital status and estimates of major depressive disorder.

Conclusion of Chapter 4

In conclusion, this study presents further evidence that BA can be an effective intervention even when administered by non-specialists over just 4 weeks, and that this effect is likely the result of the specific treatment factors, as argued by the BA model of depression. We also showed that activity monitoring can be a helpful active control group as well as a possible standalone intervention for treating low mood with weaker, yet significant effects. Future research using active control groups should assess treatment expectancy directly to ensure it is equivalent with the experimental group of interest. We also recommend further exploration of the activity monitoring intervention to determine which components make it most clinically beneficial.

Chapter 5: Examining emotional cognition as a mechanism of behavioural activation in comparison to an active and a passive control group

Introduction

The data presented in Chapter 2 showed evidence that a short course of BA treatment delivered by non-specialists can significantly reduce depression symptoms as measured by self-report. Furthermore, Chapter 4 showed that an even simpler intervention, activity monitoring (AM), which merely instructs participants to carry out daily reflection on how they spend their time and how they feel, may have similarly beneficial effects. This chapter will explore whether BA and AM have an effect on measures of emotional cognition, which would replicate and extend the findings presented in Chapter 3. Furthermore, it will investigate whether early changes on these cognitive measures seem to predict future symptomatic improvement. This would indicate a possible mechanistic role in depression recovery, as suggested by previous findings with pharmacological treatments for depression (Tranter et al., 2009).

Emotional cognition has been widely researched as a promising treatment mechanism in pharmacological interventions for depression (see Harmer et al. 2017, for review). However, it is currently unknown whether these mechanisms are also relevant to the action of psychological treatments, since there are fewer studies on this topic and their findings have been inconsistent. A recent systematic review by Groves et al. (2021) examined the effects of evidence-based psychotherapies on various markers of neurocognitive function in depression, including “hot” emotional cognition, and concluded that it was not possible to draw any conclusions due to significant heterogeneity of research protocols and findings.

Sources of such variability include the different cognitive domains that have been investigated, the behavioural parameters used to do so as well as the psychological interventions and their length. Promising findings have come from eye tracking paradigms examining attention; Vazquez et al. (2018) found reduced fixation time on negative stimuli after 10 weeks of CBT, an effect that was also seen by Holas et al. (2020) after 8 weeks of MBCT. Several other studies also examined attentional biases with the use of reaction times. For example, patients with depression showed quicker responses to positive stimuli after 12 weeks of CBT (Yoshimura et al., 2014) and 8 weeks

of mindfulness-based cognitive therapy (MBCT; De Raedt et al., 2012). However, Porter et al. (2016) found no effects on reaction times for facial emotion recognition after 16 weeks of CBT or schema therapy and neither did Fu et al. (2008).

There appears to be only one study that examined emotional memory in this context, finding that 8 weeks of MBCT increased positive and decreased negative recall bias (Van Vugt et al., 2012). However, a number of studies investigated the effects on “cold” memory using the Auditory Verbal Learning Task (AVLT). It is known that various cold cognitive deficits persist after remission of depressive symptoms (Bora et al., 2013) and it is unknown which treatments may exert beneficial effects on this domain. Encouraging findings came from He et al. (2019), who found that 12 weeks of CBT led to improved AVLT performance when compared with an active control group. However, Dannehl et al. (2019) found no significant effects on AVLT after 16 weeks of CBT when compared to healthy controls. This demonstrates the vast heterogeneity of not only behavioural measures but also the experimental designs and treatment formats, preventing straightforward comparison and replication of the existing findings.

Another factor contributing to the disparate findings may lie in the number of techniques and approaches included in any given psychological treatment. This variety can make it difficult to isolate therapeutic methods that affect specific cognitive domains. Behavioural activation and activity monitoring may be useful in this respect in that they only use a very specific set of standardised techniques. To our knowledge, there have only been two studies investigating the effect of BA on cognition. Dichter et al. (2009, 2010) found no effects on reward learning after an 8–14-week BA intervention, nor any effects on cognitive control when exposed to affective stimuli. The other study by Nasrin et al. (2017) only examined a single-session BA paradigm using measures of approach and avoidance tendencies, finding a trend towards increased approach to positive stimuli. Having used different cognitive measures and BA paradigms that were very different from each other, these findings allow for little comparison with the other findings above.

In Chapter 3, we found that a 4-week BA intervention during the COVID-19 pandemic appeared to shift facial perception towards a more positive bias. There was also some preliminary evidence that early positive changes in emotional memory could predict future symptomatic improvement. This chapter aims to replicate this analysis using data from our second

experiment, maintaining the same BA intervention and the same battery of emotionally cognitive tasks. Moreover, we aimed to extend this investigation further. Firstly, we added an additional comparison with the AM group to test whether the monitoring component of behavioural activation could be driving its cognitive effects. Secondly, we added the AVLIT measure to extend the research on non-emotional verbal memory effects in psychological treatments. Since an increase in physical activity is known to have beneficial neurotrophic effects and to improve cognitive function (Etnier et al., 2016), we hypothesised that increasing general activation could lead to similar benefits on this task.

Methods

Participants

Detailed description of the study participants is presented in Chapter 4.

Several participants' data was not included in this chapter. Five participants did not attend week 2 testing and seven participants did not attend week 4 testing due to COVID-19 isolation. One participant did not attend week 2 testing due to another illness. The final sample included 86 participants (30 in the BA group, 29 in the AM group and 27 in the passive control group).

In the FERT task, 8 participants' week 2 data and 36 participants' week 4 data could not be used due to a technical error. For this reason, we only included week 0 and week 2 in the analysis to examine early change in facial recognition. For these two time points, the final sample included 78 participants (26 in the BA group, 27 in the AM group and 24 in the passive control group).

Power calculation

As described in previous chapters, we would need a sample size of 38 to detect a difference between two independent means with an effect size of 0.7 (Cohen's d) with 80% power, while correcting for multiple pairwise comparisons in the context of three study groups ($p = .017$). It was difficult to estimate the expected effect size for our primary measure of emotional cognition (FERT accuracy), as most of the past research used pharmacological interventions and either placebo pill or healthy participant controls. Nonetheless, one study with a passive control group testing facial recognition after CBT in panic disorder patients found a large effect size (Cohen's d of 1.5; Reinecke et al., 2013).

Cognitive tasks

The **Facial Emotion Recognition Task** (FERT; see Harmer et al., 2003) was the same as described in Chapter 3, except for a few differences. Firstly, the 250 experimental trials were separated into 3 blocks, as opposed to 4; the first two blocks contained 83 trials and the final contained 84 trials. Secondly, participants were given an unlimited time to make a response, as opposed to the previous limit of 9500ms. Finally, all tasks in this study were performed in the lab under the supervision of a researcher, as opposed to being completed remotely online. Participants used keys on a keyboard labelled with the relevant emotion to make their responses instead of clicking on screen buttons using a cursor. Furthermore, there were no engagement checks in the task itself; the task instructions were read out by the researcher and participants were able to ask any questions directly.

The **Emotion Categorisation Task** (ECAT; see Chan et al., 2008) was the same as described in Chapter 3, except that the task was performed in the lab under the supervision of a researcher.

The **Emotional Recall Task** (EREC; see Harmer et al., 2004) was the same as described in Chapter 3. The only difference was that, in this study, participants had 4 minutes to write down their responses on a paper, as opposed to having 3 minutes to type them on a computer.

The **Probabilistic Instrumental Learning Task** (PILT; see Walsh et al., 2018) was the same as described in Chapter 3, except that the task was performed in the lab under the supervision of a researcher.

In the **Auditory Visual Learning Test** (AVLT; see Murphy et al., 2020) task, participants were asked to memorise and recall lists of words. In the condition of List A, they were read out a list of 15 words and were asked to repeat back as many as possible. The instruction was: “I’m going to read a list of words. When I’m through, I want you to tell me as many of the words as you can. You can say them in any order.”. The list was read out and recalled five times in a row with the same instruction beforehand.

Afterwards, participants were read out List B, which was another set of 15 words. They were asked to recall it immediately afterwards with the instruction: “Don’t tell me any words from the first list, just this second list”. This list was read and recalled only once. Straight after, participants

were asked to do a “List A short delay free recall”, recalling as many words from List A as possible without hearing the list first. After 15-20 minutes of distraction by other study tasks, participants were asked for a “List A long delay free recall”, once again recalling words from List A freely.

Procedure

All tasks (FERT, ECAT, EREC, PILT and AVLT) were administered in the lab at baseline (week 0), halfway through the intervention (week 2) and at the end (week 4). To ensure blinding, the tasks were administered by research assistants who were not involved in randomisation or intervention delivery.

Statistical analysis

All tasks were analysed using three-way mixed ANOVAs with group as a between-subject factor, time as a within-subject factor and valence (positive or negative), emotion (anger, disgust, fear, happy, sad, surprise or neutral) or trial type (win or loss) as a within subject factor. Partial correlations were carried out between early change in cognition at week 2 and later change in depression at week 4 while controlling for baseline scores.

If groups significantly differed at baseline, the task data was analysed as change-from-baseline scores (calculated as baseline scores subtracted from midway or final scores) using a two-way mixed ANOVA (omitting the within-subject factor of time).

Separate comparisons were carried out just for the BA and passive control groups in order to replicate the analysis from Chapter 3 with maximum available power. The rest of the analysis procedure remained the same as described in Chapter 4.

Results

Baseline descriptive statistics for all cognitive tasks can be found in the Appendix (Table 55).

Facial Emotion Recognition Task (FERT)

Accuracy for recognising positive and negative emotions

There was no main effect of group at baseline in accuracy for either positive [$F(2,91) = 0.43$, $p = .65$, $\eta^2 = .009$] or negative faces [$F(2,90) = 0.46$, $p = .63$, $\eta^2 = .01$], indicating that the groups were equivalent.

When comparing the BA group only with the passive control, there was no time x valence x group interaction for the first two time points [$F(1,47) = 2.17$, $p = .15$, $\eta^2 = .04$]. When comparing all three groups together, there was no time x valence x group interaction for the first two time points [$F(2,74) = 1.08$, $p = .35$, $\eta^2 = .03$].

Accuracy for recognising individual emotions

There was no main effect of group at baseline in accuracy for individual emotions, apart from the identification of sadness, as shown in Table 41 below. As a result, the data was subsequently analysed as change-from-baseline scores at the midway time point.

Table 41. Detailed statistics for the main effect of group when comparing accuracy for recognising individual emotions in the Facial Emotion Recognition Task at baseline. Asterisk indicates $p < .05$.

	$F(df)$	sig.	effect size
anger	$F(2,94) = 0.06$	$p = .94$	$\eta^2 = .001$
disgust	$F(2,94) = 1.68$	$p = .19$	$\eta^2 = .03$
fear	$F(2,94) = 0.02$	$p = .98$	$\eta^2 < .001$
happy	$F(2,94) = 0.29$	$p = .75$	$\eta^2 = .006$
sad *	$F(2,92) = 3.28$	$p = .04$	$\eta^2 = .07$
surprise	$F(2,94) = 0.47$	$p = .63$	$\eta^2 = .01$
neutral	$F(2,91) = 0.75$	$p = .48$	$\eta^2 = .02$

When comparing the BA group only with the passive control, there was no emotion x group interaction for the change-from-baseline scores at the midway timepoint [$F(4.61,202.67) = 0.77$, $p = .57$, $\eta^2 = .02$].

When comparing all three groups together, there was no emotion x group interaction for the change-from-baseline scores at the midway timepoint [$F(9.54,333.77) = 0.59, p = .81, \eta^2 = .02$].

Reaction times to identify positive and negative faces

There was no main effect of group at baseline in reaction times for either positive [$F(2,90) = 0.31, p = .74, \eta^2 = .007$] or negative emotions [$F(2,91) = 0.07, p = .93, \eta^2 = .002$], indicating that the groups were equivalent.

When comparing the BA group only with the passive control, there was no time x valence x group interaction for the first two time points [$F(1,46) = 0.13, p = .72, \eta^2 = .003$].

When comparing all three groups together, there was no time x valence x group interaction for the first two time points [$F(2,73) = 1.25, p = .29, \eta^2 = .03$].

Reaction times to identify individual emotions

There was no main effect of group at baseline in reaction times for individual emotions (see Table 42 below), indicating that the groups were equivalent.

Table 42. Detailed statistics for the main effect of group when comparing reactions times toward individual emotions in the Facial Emotion Recognition Task at baseline.

	$F(df)$	sig.	effect size
anger	$F(2,91) = 0.50$	$p = .61$	$\eta^2 = .01$
disgust	$F(2,94) = 0.03$	$p = .97$	$\eta^2 = .001$
fear	$F(2,92) = 1.53$	$p = .22$	$\eta^2 = .03$
happy	$F(2,91) = 0.16$	$p = .86$	$\eta^2 = .003$
sad	$F(2,93) = 0.67$	$p = .52$	$\eta^2 = .01$
surprise	$F(2,92) = 0.32$	$p = .73$	$\eta^2 = .007$
neutral	$F(2,92) = 0.27$	$p = .77$	$\eta^2 = .006$

When comparing the BA group only with the passive control, there was no time x valence x group interaction for the first two time points [$F(6,276) = 0.74, p = .62, \eta^2 = .02$]. When comparing all three groups together, there was no time x valence x group interaction for the first two time points [$F(12,432) = 1.18, p = .30, \eta^2 = .03$].

Misclassification of positive and negative emotions

There was no main effect of group at baseline in misclassifications for either positive [$F(2,93) = 0.30, p = .74, \eta^2 = .006$] or negative emotions [$F(2,91) = 0.33, p = .72, \eta^2 = .007$], indicating that the groups were equivalent.

When comparing the BA group only with the passive control, there was no time x valence x group interaction for the first two time points [$F(1,47) = 0.71, p = .41, \eta^2 = .02$].

When comparing all three groups together, there was no time x valence x group interaction for the first two time points [$F(2,75) = 0.37, p = .69, \eta^2 = .01$].

Misclassification of individual emotions

There was no main effect of group at baseline in reaction times for individual emotions (see Table 44), indicating that the groups were equivalent.

Table 44. Detailed statistics for the main effect of group when comparing rates of misclassifications when recognising individual emotions in the Facial Emotion Recognition Task at baseline.

	$F(df)$	sig.	effect size
anger	$F(2,93) = 0.60$	$p = .55$	$\eta^2 = .01$
disgust	$F(2,93) = 0.53$	$p = .59$	$\eta^2 = .01$
fear	$F(2,93) = 0.44$	$p = .64$	$\eta^2 = .009$
happy	$F(2,93) = 1.52$	$p = .23$	$\eta^2 = .03$
sad	$F(2,94) = 1.06$	$p = .35$	$\eta^2 = .02$
surprise	$F(2,94) = 0.09$	$p = .91$	$\eta^2 = .002$
neutral	$F(2,93) = 0.75$	$p = .47$	$\eta^2 = .02$

When comparing the BA group only with the passive control, there was no time x valence x group interaction for the first two time points [$F(6,282) = 0.36, p = .90, \eta^2 = .008$].

When comparing all three groups together, there was no time x valence x group interaction for the first two time points [$F(12,444) = 0.69, p = .76, \eta^2 = .02$].

In the BA group, there was a significant correlation between an early change in misclassification of happy faces and a later change in depression symptoms, see Fig. 22 below.

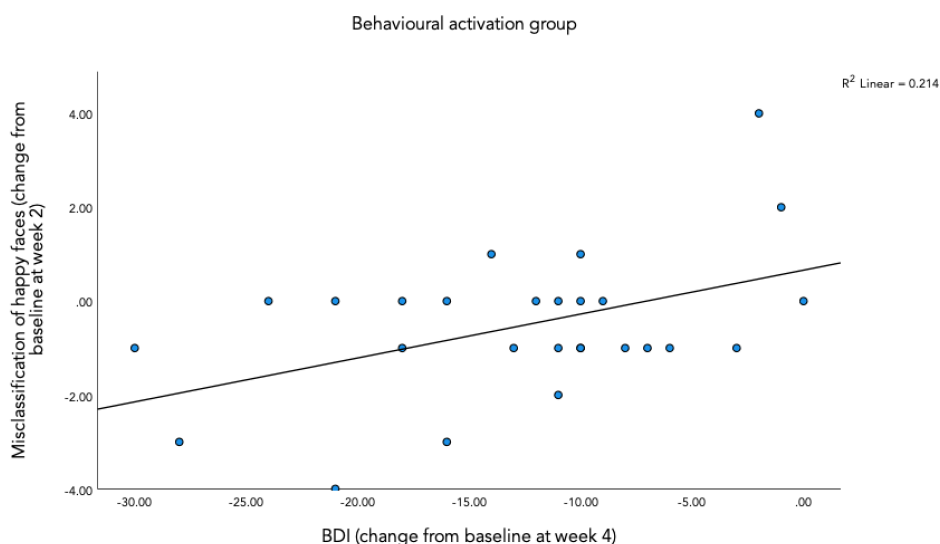


Figure 22. Positive correlations in the behavioural activation group between early change in emotional cognition at week 2 (halfway through the intervention) and later change in depression symptoms at week 4 (end of the intervention).

All correlation results for FERT parameters can be seen below in Table 45.

Table 45. Detailed statistics for correlations of early change in FERT parameters at week 2 and later change in depression symptoms at week 4 in the BA group. Asterisk indicates $p < .05$.

Emotional cognition parameter at Week 2	Correlation statistics
Accuracy for positive emotions	$r(23) = 0.03, p = .88$
Accuracy for negative emotions	$r(22) = -0.04, p = .85$
Accuracy for anger	$r(23) = 0.31, p = .13$
Accuracy for disgust	$r(23) = -0.01, p = .98$
Accuracy for fear	$r(23) = -0.07, p = .75$
Accuracy for happiness	$r(23) = 0.05, p = .83$
Accuracy for sadness	$r(22) = -0.07, p = .74$
Accuracy for surprise	$r(23) = -0.05, p = .82$
Accuracy for neutral	$r(22) = -0.30, p = .16$
Misclassification of positive faces	$r(23) = 0.03, p = .90$
Misclassification of negative faces	$r(21) = 0.06, p = .77$
Misclassification of anger	$r(21) = -0.20, p = .36$
Misclassification of disgust	$r(23) = -0.03, p = .90$
Misclassification of fear	$r(23) = -0.10, p = .62$
Misclassification of happiness *	$r(23) = 0.49, p = .01$
Misclassification of sadness	$r(23) = 0.18, p = .40$
Misclassification of surprise	$r(23) = -0.12, p = .58$
Misclassification of neutral	$r(23) = -0.04, p = .85$

Emotional categorisation task (ECAT)

Accuracy for positive and negative words

There was no significant difference between the three groups in baseline accuracy to either positive [$F(2,91) = 1.59, p = .21, \eta^2 = .03$] or negative words [$F(2,90) = 0.76, p = .47, \eta^2 = .02$], indicating that the groups were equivalent.

When comparing the BA group only with the passive control, there was no time x valence x group interaction [$F(2,96) = 0.40, p = .67, \eta^2 = .008$]. There was a main effect of group

$[F(1,48) = 3.89, p = .05, \eta^2 = .08]$, wherein the BA group was overall more accurate ($M = 0.99, SE = 0.004$) than the passive control group ($M = 0.98, SE = 0.006$), see Fig.23.

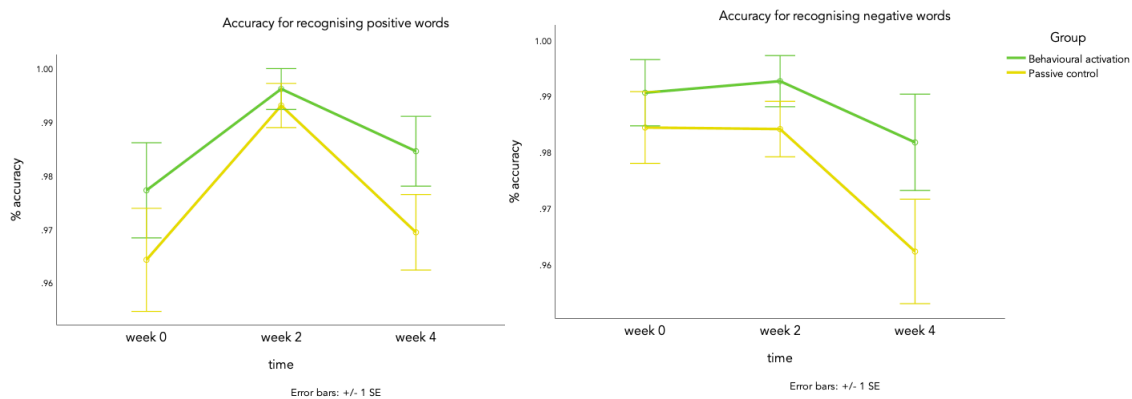


Figure 23. Accuracy for positive and negative words in the ECAT task at baseline (week 0), halfway through the interventions (week 2) and at the end (week 4).

When analysing the three groups together, there was no significant time x group x valence interaction for accuracy [$F(4,142) = 0.44, p = 0.78, \eta^2 = 0.01$].

Reaction times for positive and negative words

There was a significant difference between the three groups in baseline reaction times to both positive [$F(2,95) = 3.14, p = .05, \eta^2 = .06$] and negative words [$F(2,95) = 7.03, p = .001, \eta^2 = .13$], see Fig.25 below. As a result, reaction times were analysed as change-from-baseline scores.

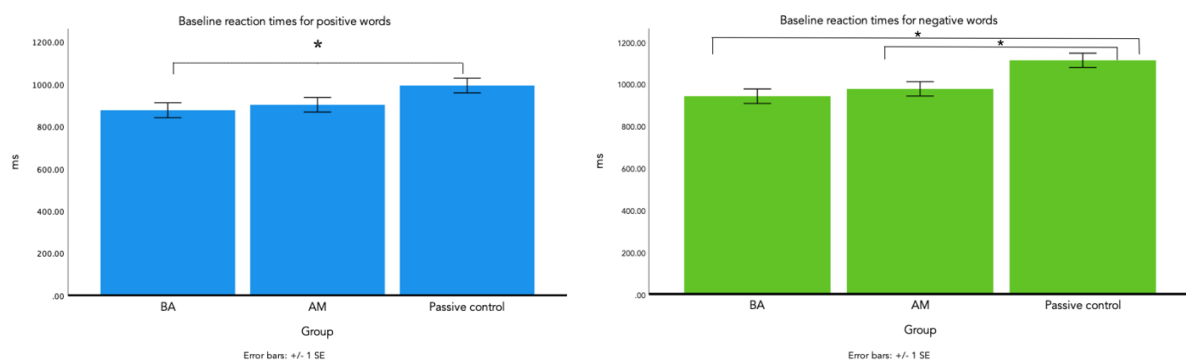


Figure 25. Baseline reaction times of the three experimental groups in response to positive and negative words in the Emotional Categorisation Task (ECAT).

When comparing the BA group only with the passive control, there was no significant valence x group interaction for the change-from-baseline scores at midway [$F(1,58) = 1.76, p = .19, \eta^2 =$

.03]. At the final timepoint, there was a trend towards a significant valence x group interaction [$F(1,55) = 3.36, p = .07, \eta^2 = .06$] and a significant main effect of time [$F(1,55) = 6.18, p = .02, \eta^2 = .10$].

When analysing the three groups together, there was no valence x group interaction for the change-from-baseline scores at midway [$F(2,87) = 1.39, p = .26, \eta^2 = .03$] or final timepoint [$F(2,84) = 1.90, p = .16, \eta^2 = .04$]. For the change scores at the final timepoint, there was a main effect of group [$F(2,84) = 3.77, p = .03, \eta^2 = .08$].

Post-hoc tests showed a significant difference between the BA group and the passive control at the final time point across both valences (*Mean difference* = 119.57, *SE* = 43.85, $p = .02$), see Fig.

26.

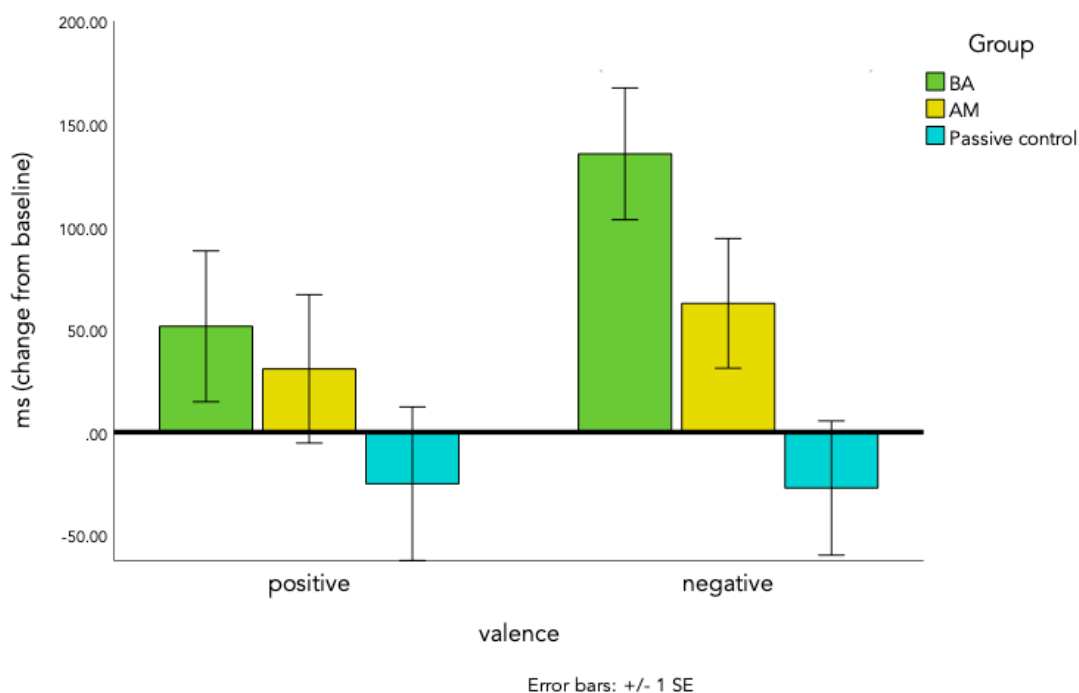


Figure 26. Change-from-baseline scores at the final time point for reaction times in the Emotion Categorisation Task (ECAT).

All correlation results for ECAT parameters can be seen below in Table 46.

Table 46. Detailed statistics for correlations of early change in ECAT parameters at week 2 and later change in depression symptoms at week 4 in the BA group. Asterisk indicates $p < .05$.

Emotional cognition parameter at Week 2	Correlation statistics
Accuracy for positive words	$r(24) = -0.19, p = .34$
Accuracy for negative words	$r(25) = -0.08, p = .70$
Reaction times for positive words	$r(26) = -0.32, p = .10$
Reaction times for negative words	$r(26) = -0.34, p = .08$

Emotional Recall Task (EREC)

Accuracy

There was no significant difference between the three groups in baseline recall accuracy for either positive [$F(2,93) = 0.55, p = 0.58, \eta^2 = 0.01$] or negative words [$F(2,94) = 0.85, p = 0.43, \eta^2 = 0.02$], indicating that the groups were equivalent.

When comparing the BA group only with the passive control, there was no significant time x group x valence interaction in recall accuracy [$F(2,108) = 0.58, p = 0.56, \eta^2 = 0.01$].

When comparing the three groups together, there was no significant time x group x valence interaction in recall accuracy [$F(4,160) = 1.08, p = 0.37, \eta^2 = 0.03$].

False alarms

There was no significant difference between the three groups in baseline rate of either positive [$F(2,93) = 0.42, p = 0.66, \eta^2 = 0.009$] or negative false alarms [$F(2,91) = 1.36, p = 0.26, \eta^2 = 0.03$], indicating that the groups were equivalent.

When comparing the BA group only with the passive control, there was no significant time x group x valence interaction when analysing false alarms [$F(2,100) = 1.06, p = 0.35, \eta^2 = 0.02$].

When comparing the three groups together, there was no significant time x group x valence interaction when analysing false alarms [$F(4,152) = 0.81, p = 0.52, \eta^2 = 0.02$].

All correlation results for EREC parameters can be seen below in Table 47.

Table 47. Detailed statistics for correlations of early change in EREC parameters at week 2 and later change in depression symptoms at week 4 in the BA group. Asterisk indicates $p < .05$.

Emotional cognition parameter at Week 2	Correlation statistics
Accuracy for positive words	$r(26) = -0.33, p = .09$
Accuracy for negative words	$r(26) = 0.08, p = .70$
False alarm rate for positive words	$r(25) = 0.04, p = .84$
False alarm rate for negative words	$r(25) = 0.22, p = .27$

Probabilistic Instrumental Learning Task (PILT)

Choice behaviour – full task

When comparing the BA group only with the passive control, there was no significant time x trial x group interaction for choice behaviour [$F(2,104) = 0.09, p = 0.92, \eta^2 = 0.002$].

When comparing the three groups together, there was no significant time x trial x group interaction for choice behaviour [$F(4,158) = 0.38, p = 0.83, \eta^2 = 0.009$].

Choice behaviour – last 30 trials

When comparing the BA group only with the passive control, there was no significant time x trial x group interaction for the last 30 times [$F(2,104) = 0.21, p = 0.81, \eta^2 = 0.004$].

When comparing the three groups together, there was no significant time x trial x group interaction for the last 30 trials [$F(4,158) = 0.18, p = 0.95, \eta^2 = 0.004$].

Auditory Visual Learning Task (AVLT)

Free recall List A – trial by trial

There was no significant difference between the three groups at baseline [$F(2,90) = 1.13, p = .33, \eta^2 = .03$], indicating that the groups were equivalent.

When comparing the BA group only with the passive control, there was no significant time x trial x group interaction for trial-wise free recall of list A [$F(8,392) = 0.33, p = .95, \eta^2 = .007$].

When comparing the three groups together, there was no significant time x trial x group interaction for trial-wise free recall of list A [$F(16,592) = 1.02, p = 0.44, \eta^2 = 0.03$].

Total free recall List A

There was no significant difference between the three groups at baseline [$F(2,93) = 1.38, p = .26, \eta^2 = .03$].

When comparing the BA group only with the passive control, there was no significant time x trial x group interaction for total free recall of list A [$F(2,106) = 0.53, p = .59, \eta^2 = .01$].

When comparing the three groups together, there was no significant time x group interaction [$F(4,156) = 1.06, p = 0.38, \eta^2 = 0.03$].

Total free recall List B

There was no significant difference between the three groups at baseline [$F(2,95) = 2.44, p = .09, \eta^2 = .05$].

When comparing the BA group only with the passive control, there was no significant time x trial x group interaction for total free recall of list A [$F(2,106) = 1.11, p = .33, \eta^2 = .02$].

When comparing the three groups together, there was no significant time x group interaction when analysing total free recall for list B [$F(4,156) = 0.53, p = 0.72, \eta^2 = 0.01$].

Short delay recall List A

There was no significant difference between the three groups at baseline [$F(2,94) = 1.28, p = .28, \eta^2 = .03$].

When comparing the BA group only with the passive control, there was no significant time x trial x group interaction for short delay recall of list A [$F(2,104) = 0.11, p = .89, \eta^2 = .002$].

When comparing the three groups together, there was no significant time x group interaction when analysing short delay recall for list A [$F(4,154) = 1.15, p = 0.33, \eta^2 = 0.03$].

Long delay recall List A

There was no significant difference between the three groups at baseline [$F(2,91) = 0.19, p = .82, \eta^2 = .004$].

When comparing the BA group only with the passive control, there was no significant time x trial x group interaction for long delay recall of list A [$F(2,100) = 0.07, p = .93, \eta^2 = .001$].

When comparing the three groups together, there was no significant time x group interaction when analysing long delay recall for list A [$F(3.58,132.60) = 0.70, p = 0.58, \eta^2 = 0.02$].

Intrusions

There was a significant difference between the three groups at baseline [$F(2,91) = 3.25, p = .04, \eta^2 = .07$]. As a result, the data was subsequently analysed as change-from-baseline scores at the final time point.

When comparing the BA group only with the passive control, there was no significant main effect of group for the change-from-baseline scores [$F(1,51) = 0.02, p = .88, \eta^2 < .001$].

When comparing the three groups together, there was no significant main effect of group for the change-from-baseline scores [$F(2,78) = 0.09, p = .91, \eta^2 = .002$].

Repetitions

There was no significant difference between the three groups at baseline [$F(2,93) = 1.93, p = .15, \eta^2 = .04$].

When comparing the BA group only with the passive control, there was no significant time x group interaction when analysing repetitions [$F(2,102) = 0.91, p = 0.41, \eta^2 = 0.02$].

When comparing the three groups together, there was no significant time x group interaction when analysing repetitions [$F(3.58,137.82) = 0.58, p = 0.66, \eta^2 = 0.02$].

Discussion

The aim of this chapter was to examine the hot and cold cognitive effects of behavioural activation and activity monitoring in comparison to a passive control group. We examined several cognitive domains, including perception of emotional faces, attention and memory for emotional words and reward learning, as well as memory for non-emotional stimuli, but found no significant effects on any of those domains. This fails to replicate the emotional cognitive effects of BA reported in Chapter 3 and shows no such effects for the AM intervention.

These null findings are surprising for a number of reasons. Firstly, this study included participants with a higher severity range of depression scores than the previous online study, leading us to expect stronger negative cognitive biases and a broader opportunity for change. Secondly, all participants carried out tests of emotional cognition in the laboratory under direct supervision of a researcher, as opposed to remote data collection at home, which should lead to higher validity of the data collected. Lastly, this study collected two thirds of the data before the

COVID-19 pandemic and the last third during a later pandemic phase free of most societal restrictions. In comparison to the sample collected in 2020, we expected the depressive symptoms in this study to be less circumstantial and more likely to result from persistent negative biases in cognition.

There are several possible explanations for these findings. Firstly, it could merely be a result of reduced statistical power, particularly for the three-group comparisons expected to show smaller effect sizes. Due to a significant technical error in the coding of the facial expression recognition task, we lost more than a third of the facial recognition data, and the rest of the data collection was also impeded, albeit to a lesser extent, by participants having to isolate due to COVID-19 restrictions. Since sample sizes have generally been small in this research area, it is advisable for a future study to test this hypothesis using a much larger sample. Secondly, it is possible that neither BA nor AM exert their effects through early cognitive changes and that our previous findings were false positives. Groves et al. (2021) have pointed out the risks of increased type 1 error rate when using multiple measures of cognition and proposed a possible solution by using *a priori* groupings of variables to create composite scores, which could be a promising strategy for future research.

Furthermore, these findings highlight how little we currently understand about the construct of emotional cognition and its resulting properties. For example, it is unknown how “stable” negative cognitive biases may be in the face of different mood factors, such as the different phases of a pandemic, how they relate to different severity scores on self-report measures of depression, how responsive they may be to different treatment types and lengths and how long-lasting any beneficial effects are. Ruhe et al. (2019) showed that even remitted patients free of antidepressant medication maintained a stable negative cognitive bias on several domains, some of which were associated with future recurrence. However, they found that the best predictive variables related to demographic factors and levels of childhood adversity. The personality dimension of neuroticism is also known to cause biased cognition (Chan et al., 2007) and could play an important role in its stability. It is possible that these factors affect how ingrained the cognitive biases may be and thus how resistant they could be to treatment. This could result in significant variability among research participants that should be investigated in future research.

Strengths and limitations

A significant strength of this analysis was the close replication of a previous paradigm allowing for a repeated examination of the same BA intervention, at the same length, and examined with the same battery of cognitive tasks. These allowed us to investigate several cognitive domains using different behavioural parameters, thus exploring whether any of the previously found effects weren't merely the artefacts of specific experimental paradigms. Moreover, the comparison with a passive control group, as well as the use of different stimuli at each data collection session, allowed us to control for any confounding effects from regression to the mean, practice effects or general task familiarity.

In addition to the aforementioned issues with statistical power, our study was limited in several ways. Firstly, we did not collect data on participants' general cognitive function at baseline, apart from verbal memory. While we excluded participants with any formally diagnosed cognitive impairment, it is possible that our analysis was confounded by variable cognitive ability, particularly in relation to participants with later-onset depression where neurovascular factors may play a role (Bora et al., 2013). Secondly, our sample included participants taking antidepressants. While we only included those that have been taking the medication for three months or more and haven't gone through any recent changes in dosage, this could still create variability in the effects on emotional cognition. Nonetheless, it allowed us to test a more ecologically valid sample of participants. We also did not include a never-depressed control group that would allow for comparisons with a "normative" profile of emotional cognition.

It is important to bear in mind that the comparison with the online study was limited by differences in sample, task administration and general study protocol as all data collection took place in our Oxford laboratory. While in-person testing should lead to higher validity as mentioned above, since all participants performed the tasks under the exact same conditions using the same computer, participants in the online experiment also may have been more comfortable and less anxious during their data collections. Future research should compare emotional cognitive findings when tests are performed online vs in-person while keeping all other factors equivalent to examine the effect of this protocol change.

Lastly, as discussed in Chapter 4, the study was limited by some baseline differences between groups that may have affected the emotional cognition findings. Several factors suggest that the passive control group may have been clinically “worse off” at baseline; participants in this group showed lower levels of initial activation, higher daily ratings of “angry” and “irritable” emotions, greater prevalence of persisted depressive symptoms and higher use of antidepressants. However, it is not clear how this could contribute to the lack of significant differences found between the groups’ cognition over time.

Conclusion of Chapter 5

In conclusion, this chapter found no significant effects of the two behavioural interventions on measures of cognition, possibly due to limited statistical power. Future research should recruit larger samples of participants with depression and use a wide battery of cognitive tasks to investigate multiple domains, perhaps utilising analytical methods based on composite scores to prevent an inflated rate of false positive findings. Further efforts to clarify the properties of emotional cognition as a neuropsychological construct are recommended.

Chapter 6: Examining the mechanism of behavioural activation using actigraphy measurement in comparison to an active and a passive control group

Introduction

The previous chapter presented findings from several cognitive measurements in order to explore whether they play an early mechanistic role in behavioural treatments and, if so, whether this could serve as an objective marker of their effectiveness. The current chapter further extends this investigation to actigraphy measurements, examining whether an early increase in motor activity could be a mechanism through which these interventions exert their effects, possibly predicting future clinical improvements and helping to resolve issues with demand effects in self-report.

Depression has long been associated with lower psychomotor activity and reduced physical fitness (Martinsen et al., 1989). A recent systematic review and meta-analysis of both self-report and objective measures confirmed that depressed individuals show lower levels of exercise and increased sedentary behaviour (Schuch et al., 2017). Additionally, low mood and activity appear to have a bidirectional relationship, wherein inactivity predicts future depression (Mammen and Faulkner, 2013) and depressive symptoms themselves further reduce motivation and self-efficacy (Vancampfort et al., 2015). This “vicious cycle” is likely to play an important part in the economic burden of depression, as shown in the significant loss of productivity associated with the condition (Mykletun et al., 2006) as well as the physical health burden. As a result, physical activity should be an important clinical marker in depression treatment.

Importantly, the association between activity and mood may be mediated by circadian processes. Exercise can improve sleep length and quality by increasing time spent in stage 4 and REM phases (Gerber et al., 2014), thereby alleviating the common symptom of insomnia in depression. Moreover, sleep quality and circadian rhythmicity appear to predict future depressive symptoms (Baglioni et al., 2011; Smagula et al., 2015) and interventions targeting circadian processes directly, such as bright light exposure or CBT for insomnia, can be effective depression treatments (Tao et al., 2020; Manber et al., 2008). Affective and circadian processes also have some commonalities in their neurotransmitter and genetic profiles (Pace-Schott &

Hobson, 2002; McClung, 2007). Together, these findings suggest that circadian changes are likely to play an important role in depression recovery.

Actigraphy measurements have long been utilised in psychiatry as they provide a helpful insight both into general activity levels as well as the associated circadian rhythmicity. They are also known to be more accurate than self-reported estimates of activity and sleep (Pye et al., 2021). A recent systematic review and meta-analysis of actigraphic studies by Wuthrich et al. (2022) reported that patients with mood disorders showed overall lower levels of activity when compared to healthy controls, even during remission, providing objective evidence of psychomotor slowing in depression.

The most notable parameters that have been utilised in research include the onset and activity levels during the 10 most active hours (M10), used as a proxy for wakefulness. For example, Pye et al. (2021) found that patients with remitted depression showed lower M10 activity when compared to healthy controls. Another important marker is the onset and activity levels during the 5 least active hours (L5), used as a proxy for sleep. Since mood disorders have been associated with lower sleep quality than healthy controls, depressed participants have shown longer sleep onset latency (Slyepchenko et al., 2019) as well as higher activity during the night (Pye et al., 2021) on actigraphic measurements.

Other markers, such as the relative amplitude (RA), interdaily stability (IS) and intradaily variability (IV), aim to capture the circadian variation of daytime and night-time activity over time. Lower RA has been associated with current major depression in older adults (Pye et al., 2021) as well as neuroticism, loneliness and lower health satisfaction (Lyll et al., 2018) in unipolar and bipolar depression. Inter-daily stability has been negatively associated with impulsiveness in borderline personality disorder, while intra-daily variability was positively associated with mood instability in the same patient group (McGowan et al., 2020).

Previous studies have shown that early changes in actigraphy markers can predict clinical improvements during pharmacological treatment of depression (Raoux et al., 1993; Baune et al., 2006; Todder et al., 2009), but to our knowledge, this has not been investigated with psychological interventions. Since behavioural activation aims to help patients increase their level of value-driven activities, while reducing avoidance and rumination, it appears that it would

be likely to impact patients' general motor activity levels and possibly their sleep quality, too. In fact, some authors perceive BA as one of the chronotherapeutic interventions (Pye et al., 2021), even though its impact on circadian markers has not been empirically tested using objective measures.

Furthermore, comparing BA with activity monitoring, the active control group in this study, would help inform the debate on the importance of specific and non-specific factors in this treatment (see Chapter 4). If BA leads to significant improvements on actigraphy markers while AM does not, it would suggest that the “active ingredients” of BA, such as psychoeducation, activity grading and goal setting, play an important role in its clinical effect. If BA and AM lead to similar improvements in activity, it would indicate that the clinical effect results from the non-specific factors, such as the patient expectations and the therapeutic alliance, or the specific factor of activity monitoring, which both groups share (or perhaps a combination of both).

If neither group improves in actigraphy markers in comparison to the passive control, it is possible that increased physical activity does not play an important role in the effects of BA and AM. Alternatively, we could argue that any effects found on self-report measures for these interventions may merely result from demand effects since they are not reflected in any of the objective markers.

In this analysis, we hypothesised that BA would lead to higher physical activity, reflected in significantly higher M10 and RA scores in comparison to both the active and passive control groups. We further predicted that early changes in these markers at week 2 of the intervention would significantly correlate with later changes in depression symptoms at week 4. As secondary outcomes, we predicted that BA would show higher interdaily stability scores, lower intradaily variability scores and lower activity during sleep, reflected in lower L5 scores.

Methods

Participants

Participant characteristics were previously described in Chapter 4. Out of the total sample of 97 participants, 11 participants' actigraphy data had to be excluded due to various forms of

malfunction, either due to damage to the watch at the time of wearing or due to errors at the data extraction and processing stages. The final sample included 86 participants, consisting of 68 females, 17 males and 1 who preferred not to disclose their gender.

Procedure

The selection criteria, screening process, data collections and interventions have been described in detail in Chapter 4.

We used the waterproof GeneActive Original watches, which were set up to record continuous raw data for one month at 25Hz frequency. Participants were given the watches at their first meeting with the researcher (start of week 0, see Fig.1 below). They were instructed to wear them on their non-dominant wrist and keep them “comfortably tight”, so that they wouldn’t move too much on their own. Participants were asked to wear the watches as much as possible, including time spent sleeping or being exposed to water. The watches were returned to the researcher at the final testing meeting (end of week 5).

Outcome measures

Our actigraphy parameters included the following markers as defined by Witting et al. (1990):

Table 48. Actigraphy parameters extracted from the GeneActive watch recording.

L5 activity (m/s^2)	average activity during the 5 most inactive hours within a 24h period (proxy for sleep)
L5 onset (hrs.mins)	the start of the L5 period
M10 activity (m/s^2)	average activity during the 10 most active hours within a 24h period (proxy for daytime physical activity)
M10 onset (hrs.mins)	the start of the M10 period
Relative amplitude (RA)	the difference between M10 activity and L5 activity, calculated as: $(M10-L5)/(M10+L5)$; higher values indicate a greater difference between daytime and night-time activity and a more robust circadian rhythm

Inter-daily stability (IS)	consistency across days; higher value indicates higher stability
Intra-daily variability (IV)	consistency within days; higher value indicates higher variability

Data processing

GeneActive software was used to configure, extract and erase each watch. All data files were first converted into a CSV format using the GeneActive software. Subsequently, the data was processed using the GGIR package (developed by van Hees et al., 2013; Sabia et al., 2014; Migueles et al., 2019) in R Statistical Programming Software, see the Appendix for the script. Any detected periods of non-wear were automatically ignored during processing. The final dataset contained the values for each of the five parameters at each day of recording. This daily data was separated into four epochs, represented as B1, B2, B3 and B4 as shown in Fig. 27 below. Each block had to have at least 4 days of valid data to be included.

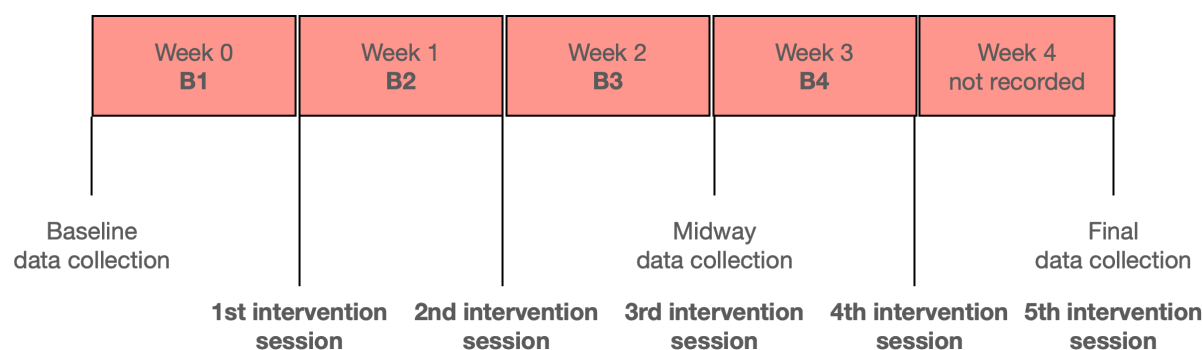


Figure 27. Correspondence of actigraphy epochs to the study schedule.

The baseline block (B1) represented participants' activity during the week between their first data collection session and their first intervention sessions (for the BA and AM groups), excluding the day of the intervention start. If this baseline period was longer than a week, only the last 7 days were used in the analysis. If it was shorter than a week, all days were included in the epoch. For the control group, the length of the B1 epoch was calculated as the average of B1 length for BA and AM groups, which was 6 days.

The first midway block (B2) represented the first week of the intervention, starting with the first intervention session. The second midway block (B3) represented the second week of the

intervention with the day of the Midway data collection session as its final day. The length of B2 and B3 blocks was determined by calculating the number of days in these two epochs and dividing them in half. If the total number of days was odd, the larger number was allocated to the B2 epoch. The final block (B4) comprised of any days left in the recording for each participant after the Midway data collection date.

The primary actigraphy measures (M10, L5, M10 onset and L5 onset) were averaged across each epoch for each participant. Additionally, the RA parameter was calculated individually for each epoch. The GGRI R package provided a single value representing the IS and IV parameters for each participant's dataset. These values were subsequently analysed using the SPSS software.

Statistical analysis

Outliers were assessed by the inspection of boxplots, and were removed if they were 3 box-lengths away from the edge of the box. This led to the removal of 3 data points.

Baseline differences, as well as the IS and IV parameters, were assessed using one-way ANOVA provided there was no violation of normality and no violation of homogeneity of error variance. Normality was assessed using the Shapiro–Wilk test; if it was violated for at least one cell of data, non-parametric tests were used instead, namely the Mann-Whitney U Test when comparing two groups and the Kruskal Wallis H Test when comparing three groups. Homogeneity of error variance was assessed using Levene's test; if it was violated, Welch's robust test of equality of mean was used instead.

We used mixed two-way ANOVAs to compare the M10 activity, M10 onset, L5 activity, L5 onset and relative amplitude (RA) parameters across groups (BA, AM and PC) and time (B1, B2 and B3). We only used the first three time blocks because only 55 out of 86 participants had valid data for B4.

Significant interaction effects were followed up with Bonferonni-corrected post-hoc t-tests; if homogeneity of error variance was violated, the Games-Howell correction was used instead.

Results

Descriptive statistics for all cognitive tasks can be found in the Appendix (Table 54).

M10 activity

A Kruskal-Wallis H test showed that there was no significant difference in M10 activity between the three groups during the baseline (B1) epoch [$\chi^2(2) = 0.96, p = .62$].

Comparison of BA and passive control

There was no significant time x group interaction in M10 activity for these two groups [$F(2,106) = 0.44, p = .65, \eta^2 = .008$], see Fig. 25.

Comparison of all three groups

There was no significant time x group interaction in M10 activity when the three groups were analysed together [$F(4, 150) = 1.002, p = .41, \eta^2 = .03$], see Fig. 28.

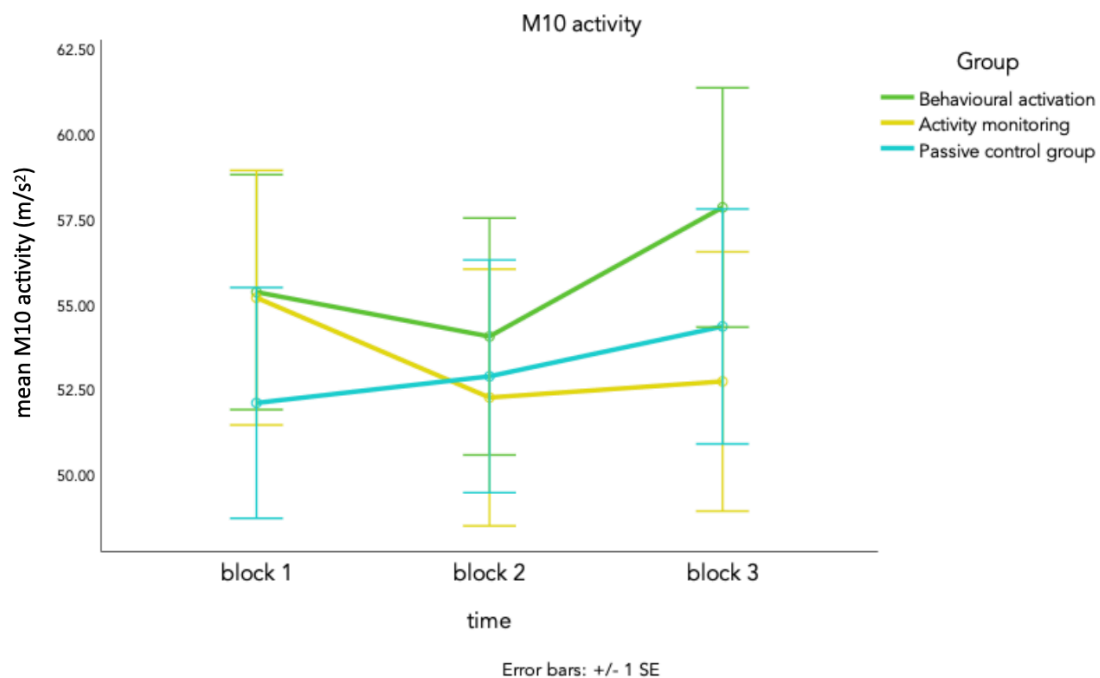


Figure 28. M10 activity as measured by the GeneActive actigraphy watches during block 1 (baseline “week 0”), block 2 (the first week of intervention) and block 3 (the second week of intervention).

M10 onset

There was no significant difference in the onset time of M10 activity between the three groups in B1 [$F(2, 83) = 0.43, p = .65, \eta^2 = .01$],

Comparison of BA and passive control

There was no significant time x group interaction in M10 onset for these two groups [$F(2,108) = 1.64, p = .20, \eta^2 = .03$]. There was no significant main effect of time [$F(2, 106) = 2.22, p = .11, \eta^2 = .04$] or group [$F(1,53) = 0.34, p = .56, \eta^2 = .006$].

Comparison of all three groups

There was no significant time x group interaction in M10 onset when the three groups were analysed together [$F(4, 150) = 1.002, p = .41, \eta^2 = .03$]. There was no significant main effect of time [$F(2, 150) = 1.22, p = .30, \eta^2 = .02$] or group [$F(2, 75) = 0.20, p = .82, \eta^2 = .005$].

L5 activity

There was no significant difference in L5 activity between the three groups at baseline in B1 [$\chi^2(2) = 3.16, p = .21$].

Comparison of BA and passive control

There was no significant time x group interaction in L5 activity when the two groups were compared [$F(1.69, 91.46) = 0.44, p = .62, \eta^2 = .008$]. There was no significant main effect of time [$F(1.69, 91.46) = 1.66, p = .20, \eta^2 = .03$] or group [$F(1, 54) = 0.45, p = .51, \eta^2 = .008$].

Comparison of all three groups

There was no significant time x group interaction in L5 activity when the three groups were analysed together [$F(3.35, 127.16) = 1.02, p = .39, \eta^2 = .03$].

L5 onset

There was no significant difference in L5 onset between the three groups at baseline in B1 [$\chi^2(2) = 2.17, p = .34$].

Comparison of BA and passive control

There was no significant time x group interaction in L5 onset when the two groups were compared [$F(2, 108) = 0.22, p = .80, \eta^2 = .004$]. There was no significant main effect of time [$F(2, 108) = 0.32, p = .73, \eta^2 = .006$] or group [$F(1, 54) = 0.31, p = .58, \eta^2 = .006$].

Comparison of all three groups

There was no significant time x group interaction in L5 onset when the three groups were analysed together [$F(4, 152) = 0.13, p = .97, \eta^2 = .003$]. There was no significant main effect of time [$F(2, 152) = 0.65, p = .52, \eta^2 = .008$] or group [$F(2, 76) = 0.23, p = .80, \eta^2 = .006$].

Relative amplitude (RA)

There was no significant difference in relative amplitude between the three groups at baseline in B1 [$\chi^2(2) = 3.80, p = .15$].

Comparison of BA and passive control

There was no significant time x group interaction in relative amplitude for these two groups [$F(1.71, 92.45) = 0.32, p = .67, \eta^2 = .006$], see Fig.4.

Comparison of all three groups

There was no significant time x group interaction in relative amplitude when the three groups were analysed together [$F(3.36, 126.10) = 1.17, p = .33, \eta^2 = .03$], see Fig.29 below. There was no significant main effect of time [$F(1.68, 126.10) = 1.83, p = .17, \eta^2 = .02$].

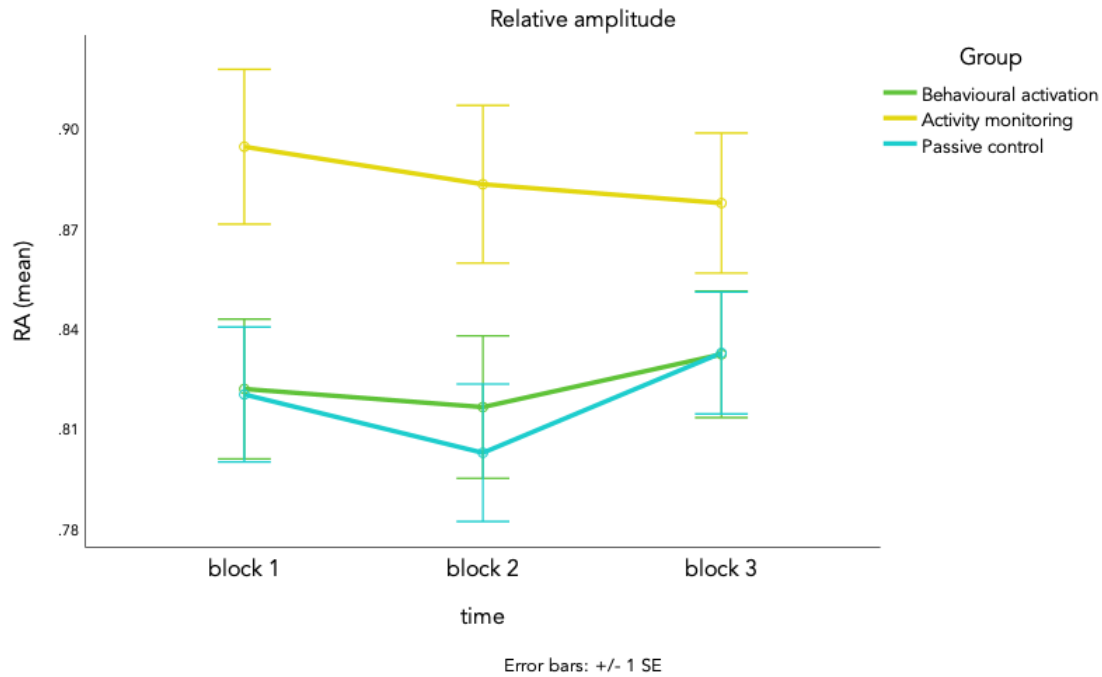


Figure 29. Relative amplitude (calculated using the M10 and L5 activity values measured by the GeneActive actigraphy watches) during block 1 (baseline “week 0”), block 2 (the first week of intervention) and block 3 (the second week of intervention).

There was a main effect of group [$F(2, 75) = 3.33, p = .04, \eta^2 = .08$], which was driven by a significant difference between the BA group and the AM group and by a significant difference between the AM group and the passive control group, see Table 49.

Table 49. Detailed statistics for post-hoc t-tests comparing relative amplitude for all participants between B2 and B3.

all time points	mean difference	significance	group M and SE
BA vs AM	$M = -0.06$	$p = .03$	BA: $M = 0.82, SE = 0.02$ AM: $M = 0.89, SE = 0.02$
AM vs PC	$M = 0.07$	$p = .03$	AM: $M = 0.89, SE = 0.02$ PC: $M = 0.82, SE = 0.02$
BA vs PC	$M = 0.005$	$p = 1.00$	BA: $M = 0.82, SE = 0.02$ PC: $M = 0.82, SE = 0.02$

Inter-daily stability (IS)

Comparison of BA and passive control

There was no significant difference in the IS scores of the two groups ($U = 541, z = 1.10, p = .27$).

Comparison of all three groups

There was no significant difference between the three groups in IS scores [$\chi^2(2) = 1.80, p = .41$].

Intra-daily variability (IV)

Comparison of BA and passive control

There was no significant difference in IV scores when directly comparing the BA and the passive control group [$t(59) = 1.21, p = .23, d = .29$].

Comparison of all three groups

There was no significant difference in IV scores when comparing the three groups [$F(2,88) = 1.45, p = .24, \eta^2 = .03$].

Correlating change in depression with change in actigraphy parameters

There were no significant correlations between change in depression scores at the final timepoint and early change in M10 activity at block 3 [$r(72) = -.08, p = .48$], early change in RA at block 3 [$r(73) = .06, p = .63$] or early change in L5 activity at block 3 [$r(73) = -.04, p = .75$].

Discussion

The aim of this chapter was to examine how actigraphic markers change during the early stages of behavioural activation (BA) and activity monitoring (AM) interventions when compared to a passive control group, in order to investigate physical activity as a potential mechanism of action. Our results indicate that there were no significant differences between the three groups in levels of day and night-time activity or their time of onset over the first three weeks of the study. There was also no significant differences on any of the circadian parameters, including the relative amplitude, inter-daily stability, and intra-daily variability. Finally, early changes in the main activity parameters did not correlate with later changes in depression scores.

These findings suggest that early changes in physical activity and circadian rhythm do not play a role in behavioural activation treatment for depression. This contradicts our prediction that the “active ingredients” of BA, such as psychoeducation about low mood and activity, activity grading and goal setting, would help participants become more physically active early in the intervention. A possible explanation for this is that any effects on physical activity were highly dependent on the types of goals participants chose to focus on. While many participants selected goals relating to physical activity, others preferred creative, social, or necessary activities, such as work or study, which were still likely to have beneficial mental health effects. Moreover, even when participants focused on exercise goals, the BA principle of avoiding “boom and bust” by increasing exercise slowly and gradually may have prevented these activities from having a significant effect on the actigraphy markers within the first few weeks.

While we assumed that the benefits of accomplishing any value-driven goals could increase participants’ general activity levels, this may not be the case, or at least not early on in the BA process. Future research could deploy experience sampling methods to examine the nature of participants’ daily activities in detail and to analyse if different types of activities lead to a different mechanism of effectiveness. Alternatively, specialist BA protocols could encourage particular activity types more than others. Some BA paradigms have already been tested with specific focus on physical activity promotion (Pentecost et al., 2015; Lambert et al., 2017) since even very low levels of activity appear to have a protective effect against depression (Mammen and Faulkner, 2013). Emphasis could also be placed on other activities with circadian benefits; for example, spending more time outside to increase exposure to natural light and other

zeitgebers. This would help direct the BA process towards a specific mechanism assumed to increase its effectiveness.

Hofmann & Hayes (2019) have recently argued for a move towards more idiographic approaches in psychiatry, wherein each individual's symptoms and treatment process are viewed as a unique network of many variables. According to this conceptualisation, increasing physical activity could be an important mechanism of effect for some individuals, but it would likely interact with other mechanistic factors at the same time. Rather than assuming a linear, unidirectional mediation for mechanisms of psychotherapy, the authors suggest the use of complex network analysis to capture the process of clinical improvement. For example, small increases in physical activity could lead to small improvements in mood, which would in turn reduce anhedonia and increase reward learning, leading to further benefits for mood and activity. Considering the complexity of each individual patient experience, research on this topic would likely benefit from this analytical approach.

Another explanation for our findings may be that neither behavioural activation nor activity monitoring led to actual clinical benefit and the effects we observed on self-report measures were merely the result of experimental demand effects. As it is currently unknown how common demand effects are, it is not possible to establish what the “true marker of efficacy” in psychological treatments might be. While there are other objective measures that can be utilised in future depression research, such as immunological markers (Zorrilla et al., 2001), voice acoustics (Cannizzaro et al., 2004) or even typing patterns (Stange et al., 2018), these will always have to be interpreted in the context of the patients' own self-report.

Strengths and limitations

The strengths of this analysis include a clearly defined *apriori* hypothesis, our use of validated actigraphy parameters as well as establishing baseline equivalence of activity levels between our three groups to prevent confounding effects.

Several limitations of this investigation should also be noted. Firstly, the three-group comparison was likely limited in power, as we weren't able to collect the originally intended sample size due to COVID-19 restrictions. Additionally, the AM intervention appeared to have a stronger effect

that we originally expected, likely leading to larger effect size that needed to be detected when comparing it with the BA group. Furthermore, several technological issues affected our analysis, both in the number of participants we were able to include as well as in the length of recording provided by the watches. It is possible that significant differences in actigraphy markers would emerge in block 4 of the study if we had more data points to include for this period.

Lastly, since all of our participants wore an actigraphy watch, we weren't able to control for the effects that this act alone may have on activation levels. While some findings suggest that wearing an actigraphy device, such as a Fitbit, may aid in the process of activation (Chum et al., 2017), we don't consider this likely in our study since the GeneActive watches gave participants no information about their activity levels and the passive control group showed minimal improvements in self-reported depression and activation (see Chapter 4).

It is notable that the participants of our study had somewhat different actigraphic findings than those found in other research on depression (Pye et al., 2021; Slyepchenko et al., 2019). Our participants showed lower inter-daily stability and higher intra-daily variability as well as lower activity for both L5 and M10 at baseline, while the relative amplitude was comparable. In comparison to a healthy control group (Pye et al., 2021), our participants also showed lower IS, IV, M10 as well as L5. This may be due to differences in factors such as impulsivity or mood instability (see introduction), which we did not control for. Future research should take these factors into account to clarify the generalisability of actigraphy findings to both healthy and depressed populations.

Conclusion of Chapter 6

In conclusion, the results of this chapter suggest that an early change in actigraphy markers is not seen with behavioural activation treatment or in any beneficial effects which may arise as a result of activity monitoring, at least not when examined with the present sample size. Future research should deploy experience sampling methods to examine the effects of specific activity types on actigraphy markers, and complex network analysis could help elucidate how multiple factors interact on an individual level. A larger sample size and longer actigraphy recordings would also help clarify whether changes in physical activity emerge later on in the treatment.

Chapter 7: General discussion

Synopsis

This thesis examined the efficacy and potential mechanisms of a short course of non-specialist behavioural activation as a treatment for depression. The first aim was to examine whether the intervention can be effective in a 4-week format when administered by a non-specialist after a short training and supervision. This is particularly relevant during the COVID-19 pandemic where incidence of depression is on the rise and mental health services are limited in capacity. The second aim was to examine possible mechanistic factors through which behavioural activation exerts its beneficial effects. By understanding the mechanism of this psychological treatment, it becomes possible to optimise the best format and conditions in which it should be delivered and perhaps even predict which patient the intervention will be most suitable for.

As summarised in Table 21 below, our two experiments provided a number of valuable insights to address those aims. Perhaps the most robust finding is that 4 weeks of non-specialist behavioural activation appear to be very effective for reducing self-reported symptoms of depression as well as increasing activation. This applied even in the context of the first and second waves of the COVID-19 pandemic in the UK, where options for positive activities were significantly constrained, and the beneficial effects remained at one-month follow-up. Moreover, an even simpler intervention that only includes activity monitoring may also provide beneficial effects despite very low time commitment on the side of the practitioner.

These are very encouraging findings, considering that this treatment length has not been examined before and there is an urgent need to provide quick and cost-effective depression treatments worldwide, particularly during crisis periods such as a pandemic. Since the training required to deliver these interventions as part of this research only lasted 3 days, new practitioners could possibly be recruited and educated relatively quickly during times of heightened need for mental health care. However, it is important to bear in mind that both of the practitioners in our experiments were psychology researchers and may have had pre-existing psychological skills that aided their effective delivery of the treatment.

Nonetheless, despite the promising findings regarding the efficacy of BA, our findings on the possible mechanisms of this treatment are less clear. To our knowledge, the two experiments of this thesis were the first ones to investigate whether behavioural activation works through the mechanism of early change in emotional cognition, as has been proposed for pharmacological treatment of depression (see Harmer et al., 2017 for review). Our first experiment found that the intervention had moderate-to-large effects on facial emotion recognition, increasing positive affective bias on measures of accuracy and misclassification. However, these changes only emerged at the end of the intervention, at the same time as finding a significant difference between groups in depression symptoms. As a result, it was not possible to disentangle which of these changes occurred first and thus infer whether one could cause the other.

While the first study found that early changes in facial emotion recognition as well as emotional recall significantly correlated with future changes in depression symptoms, these correlations had a high chance of false positive findings and were performed in the BA group only. A larger sample powered for detecting a difference between the two groups' correlations would help clarify whether this association was unique to the BA intervention. Future studies with larger power could also use structural equation modelling, path analysis and the Sobel test to understand the pathway of effect more closely.

While the large number of outcome measures in the Emotional Testing Battery increases the likelihood of finding false positives, it also makes it difficult to interpret how the many processes represented by these parameters may work together in the brain. For example, the cognitive measures that showed significant correlations at week 2 of our first study were different from the ones that showed a significant between-group difference at the end of the study (increasing recognition of positive faces versus decreasing recognition of negative ones). While both of these findings are in line with having a more positive affective bias, it is unclear whether they are two distinct mechanisms or represent the same process of change. Further theoretical development of the emotional cognition construct is necessary to understand how the different facets of cognition impact different symptoms and how they respond to different treatments. Alternatively, composite scores of multiple measures could be used (as suggested by Groves et al., 2021) to reduce the likelihood of false positives.

Moreover, it is difficult to determine what would be the appropriate normative values for these emotional cognitive measures. Past research has mostly used relative comparisons to healthy or placebo control groups, but these scores may always be affected by the expectancy related to the different treatments tested in these studies. A large, cross-sectional sample of healthy participants would be useful for developing normative scores, similar to the normative standards developed for executive function in older adults by CANTAB (Abbott et al., 2019).

In addition to these limitations, other factors may have affected why we did not replicate our emotional cognition findings in the second experiment. While it may be because the study had reduced statistical power due to technical and COVID-19 related disruptions, it is also possible that the higher severity of depression found in the sample made it more challenging to shift negative emotional bias with just 4 weeks of treatment. A comparison with a longer BA protocol, as well as other longer psychological treatments, could help us delineate what intervention is necessary to affect cognitive bias in different severities of depression.

This second study also identified another mechanistic factor that may play an important role in BA effectiveness - environmental reward. Chapter 4 showed that early increases in this variable at the midway time point predicted decreases in depression symptoms at the end of the study. This may be because participants with the most positively reinforcing environments, be it through their social surroundings or their access to their favourite activities, would be more likely to put their goals into practice quickly and thus could benefit from the positive emotional effects of their activities for longer. Future research should explore what environmental factors may be most important for this benefit and whether this knowledge could predict which patients will be most suitable for these behavioural treatments.

We found no evidence supporting the other mechanistic factors considered in Chapter 1. Neither of the two experiments showed significant effects on measures of reward processing as tested in the Probabilistic Instrumental Learning Task. An important variable that we did not take into account when analysing this task is participants' household income. It is possible that those with a higher income did not find the relatively small monetary reward of the task sufficiently reinforcing, which impeded their reward learning. Further research should revisit this investigation alongside exploring the influence of this sociodemographic factor on the performance of this task.

It is possible that, as opposed to the findings from pharmacological treatment, emotional cognition does not play an important role in BA efficacy. In a recent meta-analysis of neuro-imaging findings, Nord et al. (2021) concluded that the effects of psychotherapy and pharmacotherapy in depression appear to occur through distinct neural mechanisms. While antidepressant medication causes changes in limbic regions such as the amygdala, psychotherapy evokes changes in the medial prefrontal cortex, an area linked to cognitive control. This could correspond to the bottom-up and top-down theories of cognitive bias mentioned in Chapter 1, wherein antidepressants affect unconscious automatic processing of stimuli while psychotherapy enhances our ability to control our emotional responses after they occur. Larger studies investigating cognitive bias in different formats of psychological treatments are necessary to resolve this.

We also found no effects on the motor and circadian parameters explored in the actigraphy analysis of Chapter 6. This may be due to reduced statistical power or technical issues with the length of our actigraphy recording, since significant effects may have emerged later on in the interventions. Future research could utilise experience sampling methods together with complex network analysis to investigate individual variation in BA mechanism, as it is unlikely that all patients will have the exact same process of recovery (Hofmann and Hayes, 2019).

Another analytical approach that would improve the rigour of this research involves linear mixed models. These statistical tools significantly improve the handling of missing data in the context of repeated measures (Molenberghs et al., 2004), since using ANOVA in combination with post-hoc t-tests can lead to differences in the data points included that may bias results. As it can also effectively adjust for baseline covariates, including between-group differences that remain after randomisation, it is a popular method of choice in clinical trials and would be worth deploying in experimental medicine research investigating treatment mechanisms.

Table 50. Summary of main study findings

Study	Chapter	Main findings	Conclusion
Online behavioural activation compared to a passive control group during COVID-19	Chapter 2	<ul style="list-style-type: none"> Online behavioural activation was effective in significantly reducing self-reported symptoms of depression when compared to the passive control group Benefits remained at one-month follow-up 	<p>Behavioural activation appears to be a helpful intervention for depression when administered over video calls during a societal crisis period.</p> <p>New practitioners may be able to get trained in just a few days when demand on mental health services is high.</p>
	Chapter 3	<ul style="list-style-type: none"> Online behavioural activation led to a decrease in accuracy for negative facial expressions driven by an increased tendency to misclassify them as positive or neutral in the Facial Emotion Recognition Task Other measures of emotional cognition, including verbal memory and reward learning, showed no significant effects 	We found first evidence of a purely behavioural treatment for depression shifting affective bias in a manner similar to antidepressant medication.
Behavioural activation compared to activity monitoring and a passive control group	Chapter 4	<ul style="list-style-type: none"> Behavioural activation was effective in significantly reducing self-reported symptoms of depression when compared to both the activity monitoring group and the passive control group Activity monitoring on its own also led to significant, albeit smaller, reductions in depression Early changes in environmental reward predicted later improvements in depression symptoms 	<p>This study replicated the efficacy of our BA paradigm first presented in Chapter 2.</p> <p>Activity monitoring, a much simpler intervention with less time required for the practitioner, also appears to have beneficial effects.</p> <p>Levels of environmental reward may be an important mechanistic factor in BA effectiveness.</p>

	Chapter 5	<ul style="list-style-type: none"> ○ We did not replicate the previous effects of behavioural activation on facial emotion recognition as presented in Chapter 3 ○ We found no other effects on hot or cold cognition. This may be due to reduced statistical power of the study. 	<p>This study found no significant effects of the two behavioural interventions on cognition. Future research should investigate this in a larger sample using analytical methods that prevent an inflated rate of false positive findings.</p>
	Chapter 6	<ul style="list-style-type: none"> ○ There was no significant change in levels of day and night-time activity during the early stages of behavioural activation and activity monitoring when compared to the passive control group. ○ There was no change in circadian parameters ○ Early changes in actigraphy markers did not correlate with later changes in depression symptoms 	<p>Early change in actigraphy parameters does not appear to play a mechanistic role in behavioural activation treatment or in any benefits resulting from activity monitoring.</p> <p>Future research could utilise experience sampling methods together with complex network analysis to examine the process of change on an individual level.</p>

Conclusion

To conclude, this thesis brings new insights regarding the effectiveness of behavioural treatments for depression as well as the mechanistic factors that might influence their effects. Behavioural activation, as well as activity monitoring, appear to be very promising interventions for quick dissemination during societal times of crisis. By improving our understanding of their mechanism, we can make the most informed decisions as to who they are going to be most beneficial for and in what context.

Appendix

Chapter 2

Table 51. Descriptive statistics at baseline

Measure	BA group (mean, SD)	Control group (mean, SD)
BDI (depression)	20.38 (5.54)	19.41 (5.19)
BADS (activation)	72.36 (16.24)	72.97 (18.57)
SHAPS (anhedonia)	4.15 (3.10)	3.00 (2.17)
STAI-S (state anxiety)	51.91 (9.26)	49.64 (9.59)
STAI-T (trait anxiety)	53.91 (8.55)	51.52 (9.45)
MSPSS (social support)	58.73 (12.76)	60.58 (15.11)
ATQ (automatic negative thoughts)	34.03 (9.04)	36.30 (10.30)

Chapter 3

Table 52. Descriptive statistics at baseline

Measure	BA group (mean, SD)	Control group (mean, SD)
FERT % accuracy for positive faces	72.73 (6.31)	75.48 (5.50)
FERT % accuracy for negative faces	54.80 (8.57)	54.01 (7.04)
FERT % accuracy for anger	64.14 (10.15)	62.72 (10.12)
FERT % accuracy for disgust	45.47 (13.25)	46.62 (12.25)
FERT % accuracy for fear	43.52 (18.31)	40.15 (16.19)
FERT % accuracy for happiness	74.69 (9.52)	76.32 (8.38)
FERT % accuracy for sadness	66.09 (8.54)	66.54 (9.69)
FERT % accuracy for surprise	70.78 (9.49)	74.63 (8.44)
FERT % accuracy for neutral faces	85.63 (15.01)	79.71 (17.14)
FERT misclassifications of positive faces as negative or neutral	20.25 (5.22)	17.65 (4.62)
FERT misclassifications of negative faces as positive or neutral	41.84 (14.32)	40.74 (9.54)
FERT misclassifications of positive faces as other emotions	10.58 (2.37)	9.24 (2.27)

FERT misclassifications of negative faces as other emotions	17.65 (3.33)	17.65 (2.72)
FERT misclassifications of anger as other emotions	13.94 (4.02)	14.26 (3.74)
FERT misclassifications of disgust as other emotions	21.47 (5.17)	20.82 (4.88)
FERT misclassifications of fear as other emotions	22.00 (7.47)	22.62 (6.01)
FERT misclassifications of happiness as other emotions	9.78 (3.66)	8.91 (3.34)
FERT misclassifications of sadness as other emotions	13.19 (3.26)	12.88 (3.91)
FERT misclassifications of surprise as other emotions	11.38 (3.65)	9.56 (3.40)
FERT misclassifications of neutral faces as other emotions	1.41 (1.43)	1.88 (1.63)
FERT misclassifications of other emotions as anger	8.50 (6.65)	7.47 (4.76)
FERT misclassifications of other emotions as disgust	6.19 (5.02)	6.44 (4.50)
FERT misclassifications of other emotions as fear	6.81 (6.78)	5.29 (4.23)
FERT misclassifications of other emotions as happiness	1.60 (2.60)	2.09 (4.36)
FERT misclassifications of other emotions as sadness	14.28 (7.09)	16.03 (7.20)
FERT misclassifications of other emotions as surprise	8.13 (6.38)	8.29 (4.78)
FERT misclassifications of other emotions as neutral	47.66 (13.31)	45.32 (10.71)
FERT reactions times to positive faces	1746.78 (301.70)	1773.47 (285.98)
FERT reactions times to negative faces	2013.77 (384.65)	2009.74 (286.18)
FERT reactions times to anger	1984.01 (446.35)	1932.49 (334.43)
FERT reactions times to disgust	2025.16 (463.13)	2051.81 (349.88)
FERT reactions times to fear	2385.33 (501.58)	2571.68 (540.95)
FERT reactions times to happiness	1642.96 (283.79)	1617.38 (274.49)

FERT reactions times to sadness	1861.63 (371.57)	1751.40 (262.02)
FERT reactions times to surprise	1865.84 (369.05)	1935.65 (349.64)
FERT reactions times to neutral faces	1647.81 (420.97)	1623.43 (342.30)
ECAT accuracy for positive words	93.68 (5.27)	92.21 (13.10)
ECAT accuracy for negative words	94.12 (9.09)	93.82 (12.31)
ECAT reaction times for positive words	891.94 (306.56)	847.54 (201.27)
ECAT reaction times for negative words	974.53 (333.74)	917.70 (234.17)
EREC recall accuracy for positive words	4.12 (2.13)	4.29 (2.04)
EREC recall accuracy for negative words	3.44 (1.54)	3.09 (1.75)
EREC false alarm rate for positive words	3.00 (2.34)	3.18 (1.96)
EREC false alarm rate for negative words	1.26 (1.36)	1.12 (1.12)
PILT probability of choosing high probability symbol in Win trials (full task)	69.97 (17.40)	75.35 (15.54)
PILT probability of choosing high probability symbol in Win trials (last 20 trials)	73.24 (19.95)	80.71 (18.22)
PILT probability of choosing high probability symbol in Loss trials (full task)	35.85 (12.44)	34.68 (12.63)
PILT probability of choosing high probability symbol in Loss trials (last 20 trials)	34.26 (15.48)	31.82 (13.63)
PILT total amount won	526.47 (75.75)	526.47 (65.08)
PILT total amount lost	-487.65 (68.40)	-490.59 (80.79)

Chapter 4

Table 53. Descriptive statistics at baseline

Measure	BA group	AM group	Control group
BDI (depression)	22.56 (7.01)	23.82 (7.36)	24.48 (8.03)
BADS (activation)	74.00 (16.48)	72.67 (16.59)	67.31 (19.78)
EROS (environmental reward)	20.38 (3.93)	21.09 (3.63)	20.41 (4.64)
STAI-S (state anxiety)	48.97 (9.79)	50.85 (9.50)	49.66 (11.80)
STAI-T (trait anxiety)	58.13 (8.52)	56.24 (6.95)	59.53 (8.93)
MSPSS (social support)	59.31 (13.35)	56.64 (12.21)	54.50 (15.48)

Chapter 5

Table 54. Descriptive statistics at baseline

Measure	BA group (mean, SD)	Activity monitoring group (mean, SD)	Passive control group (mean, SD)
FERT % accuracy for positive faces	70.27 (4.48)	69.84 (5.71)	69.04 (5.66)
FERT % accuracy for negative faces	51.43 (7.39)	53.10 (7.20)	51.79 (7.04)
FERT % accuracy for anger	57.19 (13.01)	57.72 (10.80)	58.06 (11.10)
FERT % accuracy for disgust	44.69 (11.46)	49.77 (12.57)	48.98 (11.91)
FERT % accuracy for fear	35.16 (16.29)	35.91 (14.41)	35.39 (15.81)
FERT % accuracy for happiness	72.81 (6.02)	72.19 (6.47)	72.34 (7.47)
FERT % accuracy for sadness	68.47 (9.14)	67.66 (7.55)	62.03 (10.09)
FERT % accuracy for surprise	67.73 (8.14)	67.35 (7.07)	66.69 (8.20)
FERT % accuracy for neutral faces	87.42 (9.99)	83.75 (11.29)	84.51 (15.67)
FERT misclassifications of positive faces as other emotions	6.61 (2.69)	6.59 (2.64)	7.06 (2.91)
FERT misclassifications of negative faces as other emotions	8.91 (3.21)	8.86 (2.62)	8.32 (3.45)
FERT misclassifications of anger as other emotions	6.77 (4.77)	6.82 (4.42)	5.78 (3.55)
FERT misclassifications of disgust as other emotions	6.44 (4.63)	7.64 (4.72)	7.42 (5.63)
FERT misclassifications of fear as other emotions	5.53 (5.11)	6.55 (5.30)	6.68 (5.55)
FERT misclassifications of happiness as other emotions	1.34 (1.73)	1.84 (2.20)	2.28 (2.47)
FERT misclassifications of sadness as other emotions	16.88 (10.22)	14.42 (6.84)	14.16 (7.34)
FERT misclassifications of surprise as other emotions	11.88 (5.15)	11.40 (4.81)	11.84 (5.22)

FERT misclassifications of neutral faces as other emotions	54.25 (12.54)	52.63 (11.65)	56.25 (11.29)
FERT reactions times to positive faces	2.66 (0.28)	2.64 (0.34)	2.69 (0.26)
FERT reactions times to negative faces	2.80 (0.30)	2.79 (0.33)	2.77 (0.27)
FERT reactions times to anger	2.82 (0.33)	2.88 (0.40)	2.80 (0.25)
FERT reactions times to disgust	2.69 (0.29)	2.70 (0.35)	2.71 (0.32)
FERT reactions times to fear	3.09 (0.46)	2.91 (0.38)	3.03 (0.40)
FERT reactions times to happiness	2.54 (0.23)	2.52 (0.28)	2.55 (0.23)
FERT reactions times to sadness	2.60 (0.27)	2.70 (0.39)	2.66 (0.40)
FERT reactions times to surprise	2.77 (0.40)	2.75 (0.44)	2.83 (0.44)
FERT reactions times to neutral faces	2.35 (0.33)	2.41 (0.29)	2.39 (0.35)
ECAT accuracy for positive words	0.98 (0.04)	0.97 (0.05)	0.96 (0.06)
ECAT accuracy for negative words	0.99 (0.03)	0.98 (0.03)	0.98 (0.04)
ECAT reaction times for positive words	877.11 (186.54)	902.80 (194.62)	994.19 (213.09)
ECAT reaction times for negative words	941.19 (176.57)	976.10 (192.65)	1112.13 (213.13)
EREC recall accuracy for positive words	4.91 (2.15)	4.97 (2.40)	4.44 (2.08)
EREC recall accuracy for negative words	3.50 (1.61)	3.76 (2.12)	3.16 (1.83)
EREC false alarm rate for positive words	1.75 (1.22)	1.75 (1.61)	1.47 (1.39)
EREC false alarm rate for negative words	0.61 (0.88)	0.94 (1.16)	0.58 (0.77)
AVLT total free recall list A	58.13 (6.56)	59.66 (6.04)	56.91 (7.28)

AVLT short delay free recall list A	12.74 (2.07)	12.73 (1.49)	12.06 (2.28)
AVLT long delay free recall list A	12.23 (2.35)	12.28 (1.69)	11.97 (2.38)
AVLT free recall list B	7.19 (2.07)	6.97 (1.83)	6.21 (1.71)
AVLT intrusions	1.03 (1.33)	1.36 (1.73)	0.50 (0.78)
AVLT repetitions	7.39 (6.96)	5.61 (4.69)	4.72 (4.57)
PILT probability of choosing high probability symbol in Win trials	0.78 (0.26)	0.72 (0.28)	0.76 (0.26)
PILT probability of choosing high probability symbol in Win trials (last 30 trials)	0.77 (0.31)	0.75 (0.30)	0.77 (0.31)
PILT probability of choosing high probability symbol in Loss trials	0.38 (0.14)	0.36 (0.16)	0.36 (0.15)
PILT probability of choosing high probability symbol in Loss trials (last 30 trials)	0.38 (0.22)	0.34 (0.22)	0.31 (0.21s)

Chapter 6

Script used for processing actigraphy data

```
library('GGIR')

g.shell.GGIR(mode=c(1),
datadir="inputfolder",
outputdir="outputfolder",
f0=1, f1=c()),
do.enmo = TRUE,      do.anglez=TRUE,
chunksize=1,        printsummary=TRUE)

g.shell.GGIR(mode=c(2),
datadir="inputfolder",
outputdir="outputfolder",
f0=1, f1=c()),
```

```

strategy = 2,          ndayswindow=7,
hrs.del.start = 0,     hrs.del.end = 0,
maxdur = 30,          includedaycrit = 21,
L5M5window = c(0,24), M5L5res = 1,
winhr = c(5,10),
qllevels = c(c(1380/1440),c(1410/1440)),
qwindow=c(0,24),
ilevels = c(seq(0,400,by=50),8000),
mvpathreshold =c(100,120),
bout.metric = 4,
closedbout=FALSE,
IVIS_windowsize_minutes = 60, IVIS_epochsize_seconds = 3600)

```

Chapter 6

Table 55. Descriptive statistics for all time points

Measure	BA group (mean, SD)	Activity monitoring group (mean, SD)	Control group (mean, SD)
IS	0.32 (0.13)	0.32 (0.11)	0.35 (0.12)
IV	1.06 (0.24)	1.09 (0.23)	0.99 (0.24)
L5 onset (block 1)	8.26 (6.19)	6.42 (5.77)	6.56 (4.98)
L5 activity (block 1)	4.98 (2.86)	3.81 (2.64)	5.50 (4.44)
M10 onset (block 1)	12.49 (2.70)	13.19 (3.53)	12.76 (2.16)
M10 activity (block 1)	54.89 (16.92)	57.08 (21.93)	51.55 (15.29)
Relative amplitude (block 1)	0.83 (0.11)	0.87 (0.09)	0.82 (0.13)
L5 onset (block 2)	6.93 (4.53)	6.21 (4.84)	6.57 (4.62)
L5 activity (block 2)	5.12 (3.16)	3.96 (2.88)	6.16 (5.08)
M10 onset (block 2)	12.37 (2.42)	12.71 (2.38)	12.88 (2.17)
M10 activity (block 2)	54.74 (17.54)	52.76 (17.83)	54.38 (19.35)
Relative amplitude (block 2)	0.82 (0.11)	0.86 (0.10)	0.81 (0.14)
L5 onset (block 3)	7.66 (5.40)	6.93 (5.08)	7.01 (4.03)
L5 activity (block 3)	4.84 (2.86)	3.88 (2.98)	5.39 (4.25)
M10 onset (block 3)	12.74 (2.18)	12.49 (2.18)	12.25 (2.40)
M10 activity (block 3)	58.88 (20.02)	52.71 (16.66)	56.03 (20.09)

Relative amplitude (block 3)	0.84 (0.10)	0.88 (0.07)	0.83 (0.11)
L5 onset (block 4)	7.16 (5.49)	6.20 (4.84)	6.71 (4.67)
L5 activity (block 4)	4.67 (2.96)	3.89 (2.42)	6.16 (4.99)
M10 onset (block 4)	11.26 (2.38)	12.88 (2.57)	13.05 (2.48)
M10 activity (block 4)	55.32 (16.93)	58.40 (24.86)	55.59 (20.00)
Relative amplitude (block 4)	0.84 (0.11)	0.89 (0.05)	0.81 (0.15)

References

Abbott, R. A., Skirrow, C., Jokisch, M., Timmers, M., Streffer, J., van Nueten, L., ... & Weimar, C. (2019). Normative data from linear and nonlinear quantile regression in CANTAB: Cognition in mid-to-late life in an epidemiological sample. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 11, 36-44.

Admon, R., & Pizzagalli, D. A. (2015). Dysfunctional reward processing in depression. *Current Opinion in Psychology*, 4, 114–118.

Ahn, H., & Wampold, B. E. (2001). Where oh where are the specific ingredients? A meta-analysis of component studies in counseling and psychotherapy. *Journal of Counseling Psychology*, 48(3), 251.

Alboni, S., Poggini, S., Garofalo, S., Milior, G., El Hajj, H., Lecours, C., Girard, I., Gagnon, S., Boisjoly-Villeneuve, S., Brunello, N., Wolfer, D. P., Limatola, C., Tremblay, M.-È., Maggi, L., & Branchi, I. (2016). Fluoxetine treatment affects the inflammatory response and microglial function according to the quality of the living environment. *Brain, Behavior, and Immunity*, 58, 261–271.

Ali, S., Rhodes, L., Moreea, O., McMillan, D., Gilbody, S., Leach, C., ... & Delgadillo, J. (2017). How durable is the effect of low intensity CBT for depression and anxiety? Remission and relapse in a longitudinal cohort study. *Behaviour research and therapy*, 94, 1-8.

Anderson, N. H. (1968). Likableness ratings of 555 personality-trait words. *Journal of Personality and Social Psychology*, 9(3), 272.

Aragona, M., Barbato, A., Cavani, A., Costanzo, G., & Mirisola, C. (2020). Negative impacts of COVID-19 lockdown on mental health service access and follow-up adherence for immigrants and individuals in socio-economic difficulties. *Public Health*, 186, 52–56.

Arjadi, R., Nauta, M. H., Scholte, W. F., Hollon, S. D., Chowdhary, N., Suryani, A. O., ... & Bockting, C. L. (2018). Internet-based behavioural activation with lay counsellor support versus online

minimal psychoeducation without support for treatment of depression: a randomised controlled trial in Indonesia. *The Lancet Psychiatry*, 5(9), 707-716.

Armento, M. E., & Hopko, D. R. (2007). The environmental reward observation scale (EROS): Development, validity, and reliability. *Behavior Therapy*, 38(2), 107–119.

Baune, B. T., Caliskan, S., & Todder, D. (2006). A case series on the development of rest–activity rhythm and quality of sleep in patients hospitalized for treatment of uni- or bipolar depression: A potential role for quetiapine. *International Journal of Psychiatry in Clinical Practice*, 10(4), 269–275.

Beck, A. T. (1979). *Cognitive therapy of depression*. Guilford press.

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck depression inventory (BDI-II)* (Vol. 10). Pearson.

Beck, A. T., Ward, C., Mendelson, M., Mock, J., & Erbaugh, J. (1961). Beck depression inventory (BDI). *Arch Gen Psychiatry*, 4(6), 561–571.

Bennett-Levy, J., Farrand, P., Christensen, H., & Griffiths, K. (2010). *Oxford guide to low intensity CBT interventions*. Oxford University Press.

Bewernick, B. H., Urbach, A. S., Bröder, A., Kayser, S., & Schlaepfer, T. E. (2017). Walking away from depression-motor activity increases ratings of mood and incentive drive in patients with major depression. *Psychiatry Research*, 247, 68–72.

Bjornsson, A. S. (2011). Beyond the “psychological placebo”: Specifying the nonspecific in psychotherapy. *Clinical Psychology: Science and Practice*, 18(2), 113–118.

Bland, A. R., Roiser, J. P., Mehta, M. A., Sahakian, B. J., Robbins, T. W., & Elliott, R. (2021). The impact of COVID-19 social isolation on aspects of emotional and social cognition. *Cognition and Emotion*, 1–9.

Bora, E., Harrison, B. J., Yücel, M., & Pantelis, C. (2013). Cognitive impairment in euthymic major depressive disorder: A meta-analysis. *Psychological Medicine*, 43(10), 2017–2026.

Borkovec, T. D., & Costello, E. (1993). Efficacy of applied relaxation and cognitive-behavioral therapy in the treatment of generalized anxiety disorder. *Journal of Consulting and Clinical Psychology*, 61(4), 611.

Bouhuys, A. L., Geerts, E., & Gordijn, M. C. (1999). Depressed patients' perceptions of facial emotions in depressed and remitted states are associated with relapse: A longitudinal study. *The Journal of Nervous and Mental Disease*, 187(10), 595–602.

Branchi, I., Santarelli, S., Capoccia, S., Poggini, S., D'Andrea, I., Cirulli, F., & Alleva, E. (2013). Antidepressant Treatment Outcome Depends on the Quality of the Living Environment: A Pre-Clinical Investigation in Mice. *PLoS ONE*, 8(4), e62226.

Brown, G. W., Bifulco, A., Harris, T., & Bridge, L. (1986). Life stress, chronic subclinical symptoms and vulnerability to clinical depression. *Journal of Affective Disorders*, 11(1), 1–19.

Bruckner, T. A., Scheffler, R. M., Shen, G., Yoon, J., Chisholm, D., Morris, J., Fulton, B. D., Dal Poz, M. R., & Saxena, S. (2011). The mental health workforce gap in low-and middle-income countries: A needs-based approach. *Bulletin of the World Health Organization*, 89, 184–194.

Bryant, R. A., Schafer, A., Dawson, K. S., Anjuri, D., Mulili, C., Ndogoni, L., ... & Van Ommeren, M. (2017). Effectiveness of a brief behavioural intervention on psychological distress among women with a history of gender-based violence in urban Kenya: a randomised clinical trial. *PLoS medicine*, 14(8), e1002371.

Burton, C., McKinstry, B., Tătar, A. S., Serrano-Blanco, A., Pagliari, C., & Wolters, M. (2013). Activity monitoring in patients with depression: A systematic review. *Journal of Affective Disorders*, 145(1), 21–28.

Cannizzaro, M., Harel, B., Reilly, N., Chappell, P., & Snyder, P. J. (2004). Voice acoustical measurement of the severity of major depression. *Brain and Cognition*, 56(1), 30–35.

Carlbring, P., Hägglund, M., Luthström, A., Dahlin, M., Kadowaki, Å., Vernmark, K., & Andersson, G. (2013). Internet-based behavioral activation and acceptance-based treatment for depression: A randomized controlled trial. *Journal of Affective Disorders*, 148(2–3), 331–337.

Chan, S. W. Y., Goodwin, G. M., & Harmer, C. J. (2007). Highly neurotic never-depressed students have negative biases in information processing. *Psychological Medicine*, 37(9), 1281–1291.

Chan, S. W. Y., Harmer, C. J., Goodwin, G. M., & Norbury, R. (2008). Risk for depression is associated with neural biases in emotional categorisation. *Neuropsychologia*, 46(12), 2896–2903.

Chum, J., Kim, M. S., Zielinski, L., Bhatt, M., Chung, D., Yeung, S., Litke, K., McCabe, K., Whattam, J., Garrick, L., O'Neill, L., Goyert, S., Merrifield, C., Patel, Y., & Samaan, Z. (2017). Acceptability of the Fitbit in behavioural activation therapy for depression: A qualitative study. *Evidence Based Mental Health*, 20(4), 128–133.

Ciharova, M., Furukawa, T. A., Efthimiou, O., Karyotaki, E., Miguel, C., Noma, H., Cipriani, A., Riper, H., & Cuijpers, P. (2021). Cognitive restructuring, behavioral activation and cognitive-behavioral therapy in the treatment of adult depression: A network meta-analysis. *Journal of Consulting and Clinical Psychology*, 89(6), 563–574.

Cipriani, A., Furukawa, T. A., Salanti, G., Chaimani, A., Atkinson, L. Z., Ogawa, Y., Leucht, S., Ruhe, H. G., Turner, E. H., Higgins, J. P. T., Egger, M., Takeshima, N., Hayasaka, Y., Imai, H., Shinohara, K., Tajika, A., Ioannidis, J. P. A., & Geddes, J. R. (2018). Comparative Efficacy and Acceptability of 21 Antidepressant Drugs for the Acute Treatment of Adults With Major Depressive Disorder: A Systematic Review and Network Meta-Analysis. *FOCUS*, 16(4), 420–429.

Clemente-Suárez, V. J., Martínez-González, M. B., Benitez-Agudelo, J. C., Navarro-Jiménez, E., Beltran-Velasco, A. I., Ruisoto, P., Diaz Arroyo, E., Laborde-Cárdenas, C. C., & Tornero-Aguilera, J. F. (2021). The Impact of the COVID-19 Pandemic on Mental Disorders. A Critical Review. *International Journal of Environmental Research and Public Health*, 18(19), 10041.

Cuijpers, P., Berking, M., Andersson, G., Quigley, L., Kleiboer, A., & Dobson, K. S. (2013). A Meta-Analysis of Cognitive-Behavioural Therapy for Adult Depression, Alone and in Comparison with other Treatments. *The Canadian Journal of Psychiatry*, 58(7), 376–385.

Cuijpers, P., van Straten, A., & Warmerdam, L. (2007). Behavioral activation treatments of depression: A meta-analysis. *Clinical Psychology Review*, 27(3), 318–326.

Dahne, J., Collado, A., Lejuez, C. W., Risco, C., Diaz, V. A., Kustanowitz, J., Zvolensky, M., & Carpenter, M. J. (2019). *Aptivate!*: A Spanish-language behavioral activation mobile application for delivery via primary care. *Psychological Services*, 16(2), 271.

Daly, M., Sutin, A. R., & Robinson, E. (2020). Longitudinal changes in mental health and the COVID-19 pandemic: Evidence from the UK Household Longitudinal Study. *Psychological Medicine*, 1–10.

Dannehl, K., Rief, W., & Euteneuer, F. (2019). Effects of cognitive behavioural therapy on verbal learning and memory in major depression: Results of a randomized controlled trial. *Clinical Psychology & Psychotherapy*, 26(3), 291–297.

De Raedt, R., Baert, S., Demeyer, I., Goeleven, E., Raes, A., Visser, A., Wismans, M., Jansen, E., Schacht, R., Van Aalderen, J. R., & Speckens, A. (2012). Changes in Attentional Processing of Emotional Information Following Mindfulness-Based Cognitive Therapy in People with a History of Depression: Towards an Open Attention for all Emotional Experiences. *Cognitive Therapy and Research*, 36(6), 612–620.

Dichter, G. S., Felder, J. N., Petty, C., Bizzell, J., Ernst, M., & Smoski, M. J. (2009). The Effects of Psychotherapy on Neural Responses to Rewards in Major Depression. *Biological Psychiatry*, 66(9), 886–897.

Dichter, G. S., Felder, J. N., & Smoski, M. J. (2010). The effects of Brief Behavioral Activation Therapy for Depression on cognitive control in affective contexts: An fMRI investigation. *Journal of Affective Disorders*, 126(1), 236–244.

Dimidjian, S., Hollon, S. D., Dobson, K. S., Schmalings, K. B., Kohlenberg, R. J., Addis, M. E., Gallop, R., McGlinchey, J. B., Markley, D. K., Gollan, J. K., Atkins, D. C., Dunner, D. L., & Jacobson, N. S. (2006). Randomized trial of behavioral activation, cognitive therapy, and antidepressant medication in the acute treatment of adults with major depression. *Journal of Consulting and Clinical Psychology, 74*(4), 658–670.

Disner, S. G., Beevers, C. G., Haigh, E. A. P., & Beck, A. T. (2011). Neural mechanisms of the cognitive model of depression. *Nature Reviews Neuroscience, 12*(8), 467–477.

Ducasse, D., Loas, G., Dassa, D., Gramaglia, C., Zeppegno, P., Guillaume, S., Olié, E., & Courtet, P. (2018). Anhedonia is associated with suicidal ideation independently of depression: A meta-analysis. *Depression and Anxiety, 35*(5), 382–392.

Ekers, D. M., Dawson, M. S., & Bailey, E. (2013). Dissemination of behavioural activation for depression to mental health nurses: Training evaluation and benchmarked clinical outcomes. *Journal of Psychiatric and Mental Health Nursing, 20*(2), 186–192.

Ekers, D., Webster, L., Van Straten, A., Cuijpers, P., Richards, D., & Gilbody, S. (2014). Behavioural Activation for Depression; An Update of Meta-Analysis of Effectiveness and Sub Group Analysis. *PLoS ONE, 9*(6).

Eldar, E., Rutledge, R. B., Dolan, R. J., & Niv, Y. (2016). Mood as Representation of Momentum. *Trends in Cognitive Sciences, 20*(1), 15–24.

Etnier, J. L., Wideman, L., Labban, J. D., Piepmeyer, A. T., Pendleton, D. M., Dvorak, K. K., & Becofsky, K. (2016). The Effects of Acute Exercise on Memory and Brain-Derived Neurotrophic Factor (BDNF). *Journal of Sport and Exercise Psychology, 38*(4), 331–340.

Ettman, C. K., Abdalla, S. M., Cohen, G. H., Sampson, L., Vivier, P. M., & Galea, S. (2020). Prevalence of Depression Symptoms in US Adults Before and During the COVID-19 Pandemic. *JAMA Network Open, 3*(9), e2019686.

Fancourt, D., Steptoe, A., & Bu, F. (2021). Trajectories of anxiety and depressive symptoms during enforced isolation due to COVID-19 in England: A longitudinal observational study. *The Lancet Psychiatry*, 8(2), 141–149.

Farashbandi, A., Farashbandi, F., Eftekharsaadi, Z., & Haidarie, A. (2020). The Mediating Role of Cognitive Bias in Explaining the Correlation Between Behavioral Activation system and Behavioral Inhibition System with Depression. *Journal of Health Promotion Management*, 9(5), 59–71.

Farashbandi, A., Hafezi, F., Eftekhar Saadi, Z., & Heidari, A. (2021). Role of Behavioral Activation and Inhibition Systems in Symptoms of Major Depression Disorder Regarding the Mediating Role of Cognitive Bias. *Avicenna Journal of Neuro Psycho Physiology*, 8(4), 215-221.

Farrand, P., Taylor A., Greaves C., & Pentecost C. (2013). *Get Active, Feel Good!* Retrieved from http://cedar.exeter.ac.uk/media/universityofexeter/schoolofpsychology/cedar/documents/BA_Control.pdf

Ferrari, A. J., Somerville, A. J., Baxter, A. J., Norman, R., Patten, S. B., Vos, T., & Whiteford, H. A. (2013). Global variation in the prevalence and incidence of major depressive disorder: A systematic review of the epidemiological literature. *Psychological Medicine*, 43(3), 471–481.

Fisher, P. L., & Durham, R. C. (1999). Recovery rates in generalized anxiety disorder following psychological therapy: an analysis of clinically significant change in the STAI-T across outcome studies since 1990. *Psychological medicine*, 29(6), 1425-1434.

Freeman, D., Emsley, R., Diamond, R., Collett, N., Bold, E., Chadwick, E., Isham, L., Bird, J. C., Edwards, D., Kingdon, D., Fitzpatrick, R., Kabir, T., Waite, F., Carr, L., Causier, C., Černis, E., Kirkham, M., Lambe, S., Lister, R., ... Twivy, E. (2021). Comparison of a theoretically driven cognitive therapy (the Feeling Safe Programme) with befriending for the treatment of persistent persecutory delusions: A parallel, single-blind, randomised controlled trial. *The Lancet Psychiatry*, 8(8), 696–707.

Fried, E. I., & Nesse, R. M. (2015). Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR*D study. *Journal of Affective Disorders*, 172, 96–102.

Fu, C. H. Y., Williams, S. C. R., Cleare, A. J., Scott, J., Mitterschiffthaler, M. T., Walsh, N. D., Donaldson, C., Suckling, J., Andrew, C., Steiner, H., & Murray, R. M. (2008). Neural Responses to Sad Facial Expressions in Major Depression Following Cognitive Behavioral Therapy. *Biological Psychiatry*, 64(6), 505–512.

Fu, Z., Burger, H., Arjadi, R., Nauta, M. H., & Bockting, C. L. (2021). Explaining the Efficacy of an Internet-Based Behavioral Activation Intervention for Major Depression: A Mechanistic Study of a Randomized-Controlled Trial. *Clinical Psychology in Europe*, 3(3), 1–24.

Fullana, M. A., Hidalgo-Mazzei, D., Vieta, E., & Radua, J. (2020). Coping behaviors associated with decreased anxiety and depressive symptoms during the COVID-19 pandemic and lockdown. *Journal of Affective Disorders*, 275, 80–81.

Gawrysiak, M., Nicholas, C., & Hopko, D. R. (2009). Behavioral activation for moderately depressed university students: Randomized controlled trial. *Journal of Counseling Psychology*, 56(3), 468–475.

Gerber, M., Brand, S., Herrmann, C., Colledge, F., Holsboer-Trachsler, E., & Pühse, U. (2014). Increased objectively assessed vigorous-intensity exercise is associated with reduced stress, increased mental health and good objective and subjective sleep in young adults. *Physiology & Behavior*, 135, 17–24.

Gillespie, A., Carson, J., Van Assche, I., Murphy, S., & Harmer, C. (2020). P. 737 Risk factors for depression vulnerability during the COVID-19 pandemic: findings from the Oxford COSIE (COVID-19, Social Isolation and Emotion) Study. *European Neuropsychopharmacology*, S417-S418.

Gladstone, G. L., Parker, G. B., Malhi, G. S., & Wilhelm, K. A. (2007). Feeling unsupported? An investigation of depressed patients' perceptions. *Journal of Affective Disorders*, 103(1-3), 147-154.

Glahn, D. C., Curran, J. E., Winkler, A. M., Carless, M. A., Kent, J. W., Charlesworth, J. C., Johnson, M. P., Göring, H. H. H., Cole, S. A., Dyer, T. D., Moses, E. K., Olvera, R. L., Kochunov, P., Duggirala, R., Fox, P. T., Almasy, L., & Blangero, J. (2012). High Dimensional Endophenotype Ranking in the Search for Major Depression Risk Genes. *Biological Psychiatry*, 71(1), 6–14.

Gortner, E. T., Gollan, J. K., Jacobson, N. S., & Dobson, K. S. (1998). Cognitive-behavioral treatment for depression: Relapse prevention. *Journal of Consulting and Clinical Psychology*, 377–384.

Grey, J. A., & McNaughton, N. (1982). *The neuropsychology of anxiety. An enquiry into the functions of the septo-hippocampal system*. New York: Clarendon Press/Oxford University Press.

Groves, S. J., Douglas, K. M., Milanovic, M., Bowie, C. R., & Porter, R. J. (2021). Systematic review of the effects of evidence-based psychotherapies on neurocognitive functioning in mood disorders. *Australian & New Zealand Journal of Psychiatry*, 55(10), 944–957.

Halahakoon, D. C., Kieslich, K., O'Driscoll, C., Nair, A., Lewis, G., & Roiser, J. P. (2020). Reward-processing behavior in depressed participants relative to healthy volunteers: A systematic review and meta-analysis. *JAMA psychiatry*, 77(12), 1286-1295.

Harmer, C. J., Bhagwagar, Z., Perrett, D. I., Völlm, B. A., Cowen, P. J., & Goodwin, G. M. (2003). Acute SSRI administration affects the processing of social cues in healthy volunteers. *Neuropsychopharmacology*, 28(1), 148–152.

Harmer, C. J., Duman, R. S., & Cowen, P. J. (2017). How do antidepressants work? New perspectives for refining future treatment approaches. *The Lancet Psychiatry*, 4(5), 409–418.

Harmer, C. J., Goodwin, G. M., & Cowen, P. J. (2009). Why do antidepressants take so long to work? A cognitive neuropsychological model of antidepressant drug action. *The British Journal of Psychiatry*, 195(2), 102–108.

Harmer, C. J., Shelley, N. C., Cowen, P. J., & Goodwin, G. M. (2004). Increased Positive Versus Negative Affective Perception and Memory in Healthy Volunteers Following Selective Serotonin and Norepinephrine Reuptake Inhibition. *American Journal of Psychiatry*, 161(7), 1256–1263.

Hawryluck, L., Gold, W. L., Robinson, S., Pogorski, S., Galea, S., & Styra, R. (2004). SARS Control and Psychological Effects of Quarantine, Toronto, Canada. *Emerging Infectious Diseases*, 10(7), 1206–1212.

He, H.-L., Zhang, M., Gu, C.-Z., Xue, R.-R., Liu, H.-X., Gao, C.-F., & Duan, H.-F. (2019). Effect of Cognitive Behavioral Therapy on Improving the Cognitive Function in Major and Minor Depression. *The Journal of Nervous and Mental Disease*, 207(4), 232–238.

Hilvert-Bruce, Z., Rossouw, P. J., Wong, N., Sunderland, M., & Andrews, G. (2012). Adherence as a determinant of effectiveness of internet cognitive behavioural therapy for anxiety and depressive disorders. *Behaviour Research and Therapy*, 50(7), 463–468.

Hofmann, S. G., & Hayes, S. C. (2019). The Future of Intervention Science: Process-Based Therapy. *Clinical Psychological Science*, 7(1), 37–50.

Holas, P., Krejtz, I., Wisiecka, K., Rusanowska, M., & Nezlek, J. B. (2020). Modification of Attentional Bias to Emotional Faces Following Mindfulness-Based Cognitive Therapy in People with a Current Depression. *Mindfulness*, 11(6), 1413–1423.

Hollon, S. D., & Kendall, P. C. (1980). Cognitive self-statements in depression: Development of an automatic thoughts questionnaire. *Cognitive Therapy and Research*, 4(4), 383–395.

Holmes, E. A., Ghaderi, A., Harmer, C. J., Ramchandani, P. G., Cuijpers, P., Morrison, A. P., Roiser, J. P., Bocking, C. L., O'Connor, R. C., & Shafran, R. (2018). The Lancet Psychiatry Commission on psychological treatments research in tomorrow's science. *The Lancet Psychiatry*, 5(3), 237–286.

Holmes, E. A., O'Connor, R. C., Perry, V. H., Tracey, I., Wessely, S., Arseneault, L., Ballard, C., Christensen, H., Silver, R. C., Everall, I., Ford, T., John, A., Kabir, T., King, K., Madan, I., Michie, S., Przybylski, A. K., Shafran, R., Sweeney, A., ... Bullmore, E. (2020). Multidisciplinary research

priorities for the COVID-19 pandemic: A call for action for mental health science. *The Lancet Psychiatry*, 7(6), 547–560.

Hölzel, L., Härter, M., Reese, C., & Kriston, L. (2011). Risk factors for chronic depression—A systematic review. *Journal of Affective Disorders*, 129(1–3), 1–13.

Hopko, D. R., Lejuez, C. W., Lepage, J. P., Hopko, S. D., & McNeil, D. W. (2003). A Brief Behavioral Activation Treatment for Depression: A Randomized Pilot Trial within an Inpatient Psychiatric Hospital. *Behavior Modification*, 27(4), 458–469.

Hopko, D. R., Lejuez, C. W., Ryba, M. M., Shorter, R. L., & Bell, J. L. (2016). Support for the efficacy of behavioural activation in treating anxiety in breast cancer patients. *Clinical Psychologist*, 20(1), 17–26.

Huang, Y., & Zhao, N. (2020). Generalized anxiety disorder, depressive symptoms and sleep quality during COVID-19 outbreak in China: A web-based cross-sectional survey. *Psychiatry Research*, 288, 112954.

Huneke, N. T., Walsh, A. E., Brown, R., Browning, M., & Harmer, C. J. (2017). No evidence for an acute placebo effect on emotional processing in healthy volunteers. *Journal of Psychopharmacology*, 31(12), 1578–1587.

Irnside, M., O'Shea, J., Cowen, P. J., & Harmer, C. J. (2016). Frontal cortex stimulation reduces vigilance to threat: Implications for the treatment of depression and anxiety. *Biological Psychiatry*, 79(10), 823–830.

Jacob, L., Tully, M. A., Barnett, Y., Lopez-Sanchez, G. F., Butler, L., Schuch, F., López-Bueno, R., McDermott, D., Firth, J., Grabovac, I., Yakkundi, A., Armstrong, N., Young, T., & Smith, L. (2020). The relationship between physical activity and mental health in a sample of the UK public: A cross-sectional study during the implementation of COVID-19 social distancing measures. *Mental Health and Physical Activity*, 19, 100345.

Jacobson, N. S., Dobson, K. S., Truax, P. A., Addis, M. E., Koerner, K., Gollan, J. K., ... & Prince, S. E. (1996). A component analysis of cognitive-behavioral treatment for depression. *Journal of consulting and clinical psychology*, 64(2), 295.

Joormann, J., Talbot, L., & Gotlib, I. H. (2007). Biased processing of emotional information in girls at risk for depression. *Journal of Abnormal Psychology*, 116(1), 135–143.

Judd, L., & Wells, B. (1996). *Depressive Symptoms and Major Depression in a Sample of the General Population*. 7.

Kaltenboeck, A., Ruzickova, T., Breunhölder, V., Zghoul, T., Cowen, P. J., & Harmer, C. J. (2022). No antidepressant-like acute effects of bright light on emotional information processing in healthy volunteers. *Psychopharmacology*, 239(1), 277–286.

Kanter, J. W., Busch, A. M., & Rusch, L. C. (2009). *Distinctive features of behavioral activation*. London: Routledge Press.

Kanter, J. W., Mulick, P. S., Busch, A. M., Berlin, K. S., & Martell, C. R. (2007). The Behavioral Activation for Depression Scale (BADs): Psychometric properties and factor structure. *Journal of Psychopathology and Behavioral Assessment*, 29(3), 191–202.

Kanter, J. W., Rusch, L. C., Busch, A. M., & Sedivy, S. K. (2009). Validation of the Behavioral Activation for Depression Scale (BADs) in a community sample with elevated depressive symptoms. *Journal of Psychopathology and Behavioral Assessment*, 31(1), 36–42.

Karimpour-Vazifehkhori, A., Bakhshipour Rudsari, A., Rezvanizadeh, A., Kehtary-Harzag, L., & Hasanizadeh, K. (2020). Behavioral Activation Therapy on Reward Seeking Behaviors in Depressed People: An Experimental study. *Journal of Caring Sciences*, 9(4), 195–202.

Karyotaki, E., Smit, Y., Holdt Henningsen, K., Huibers, M. J. H., Robays, J., de Beurs, D., & Cuijpers, P. (2016). Combining pharmacotherapy and psychotherapy or monotherapy for major depression? A meta-analysis on the long-term effects. *Journal of Affective Disorders*, 194, 144–152.

Kendler, K. S., Thornton, L. M., & Gardner, C. O. (2001). Genetic risk, number of previous depressive episodes, and stressful life events in predicting onset of major depression. *American Journal of Psychiatry*, 158(4), 582–586.

Kieslich, K., Valton, V., & Roiser, J. (2021). *Pleasure, reward value and prediction error in anhedonia*.

Klein, D. F. (1997). Control groups in pharmacotherapy and psychotherapy evaluations. *Treatment*, 1(1), 1a.

Knol, M. J., Twisk, J. W. R., Beekman, A. T. F., Heine, R. J., Snoek, F. J., & Pouwer, F. (2006). Depression as a risk factor for the onset of type 2 diabetes mellitus. A meta-analysis. *Diabetologia*, 49(5), 837.

Kramer, I., Simons, C. J. P., Hartmann, J. A., Menne-Lothmann, C., Viechtbauer, W., Peeters, F., Schruers, K., van Bemmelen, A. L., Myin-Germeys, I., Delespaul, P., van Os, J., & Wichers, M. (2014). A therapeutic application of the experience sampling method in the treatment of depression: A randomized controlled trial. *World Psychiatry*, 13(1), 68–77.

Kwong, A. S. F., Pearson, R. M., Adams, M. J., Northstone, K., Tilling, K., Smith, D., Fawns-Ritchie, C., Bould, H., Warne, N., Zammit, S., Gunnell, D. J., Moran, P. A., Micali, N., Reichenberg, A.,

Hickman, M., Rai, D., Haworth, S., Campbell, A., Altschul, D., ... Timpson, N. J. (2021). Mental health before and during the COVID-19 pandemic in two longitudinal UK population cohorts. *The British Journal of Psychiatry*, 218(6), 334–343.

Lambert, J. D., Greaves, C. J., Farrand, P., Haase, A. M., & Taylor, A. H. (2017). Development of a web-based intervention (eMotion) based on behavioural activation to promote physical activity in people with depression. *Mental Health and Physical Activity*, 13, 120–136.

Lejuez, C. W., Hopko, D. R., Acierno, R., Daughters, S. B., & Pagoto, S. L. (2011). Ten year revision of the brief behavioral activation treatment for depression: Revised treatment manual. *Behavior Modification, 35*(2), 111–161.

Lewinsohn, P. M. (1974). A behavioral approach to depression. *Essential Papers on Depression, 150–172*.

Lewinsohn, P. M., & Graf, M. (1973). Pleasant activities and depression. *Journal of Consulting and Clinical Psychology, 41*(2), 261.

Lewinsohn, P. M., & Libet, J. (1972). Pleasant events, activity schedules, and depression. *Journal of Abnormal Psychology, 291–295*.

Lewinsohn, P. M., & Talkington, J. (1979). Studies on the Measurement of Unpleasant Events and Relations with Depression. *Applied Psychological Measurement, 3*(1), 83–101.

Liu, X., Li, L., Li, M., Ren, Z., & Ma, P. (2021). Characterizing the subtype of anhedonia in major depressive disorder: A symptom-specific multimodal MRI study. *Psychiatry Research: Neuroimaging, 308*, 111239.

Lopes, B. C. da S., & Jaspal, R. (2020). Understanding the mental health burden of COVID-19 in the United Kingdom. *Psychological Trauma: Theory, Research, Practice, and Policy, 12*(5), 465.

Ly, K. H., Topooco, N., Cederlund, H., Wallin, A., Bergström, J., Molander, O., Carlbring, P., & Andersson, G. (2015). Smartphone-Supported versus Full Behavioural Activation for Depression: A Randomised Controlled Trial. *PLOS ONE, 10*(5), e0126559.

Ly, K. H., Trüschel, A., Jarl, L., Magnusson, S., Windahl, T., Johansson, R., Carlbring, P., & Andersson, G. (2014). Behavioural activation versus mindfulness-based guided self-help treatment administered through a smartphone application: A randomised controlled trial. *BMJ Open, 4*(1), e003440.

Lyall, L. M., Wyse, C. A., Graham, N., Ferguson, A., Lyall, D. M., Cullen, B., Celis Morales, C. A., Biello, S. M., Mackay, D., Ward, J., Strawbridge, R. J., Gill, J. M. R., Bailey, M. E. S., Pell, J. P., & Smith, D. J. (2018). Association of disrupted circadian rhythmicity with mood disorders, subjective wellbeing, and cognitive function: A cross-sectional study of 91 105 participants from the UK Biobank. *The Lancet Psychiatry*, 5(6), 507–514.

Magidson, J. F., Andersen, L. S., Satinsky, E. N., Myers, B., Kagee, A., Anvari, M., & Joska, J. A. (2020). “Too much boredom isn’t a good thing”: Adapting behavioral activation for substance use in a resource-limited South African HIV care setting. *Psychotherapy*, 57(1), 107–118.

Mammen, G., & Faulkner, G. (2013). Physical Activity and the Prevention of Depression: A Systematic Review of Prospective Studies. *American Journal of Preventive Medicine*, 45(5), 649–657.

Martell, C. R., Dimidjian, S., & Herman-Dunn, R. (2010). *Behavioral activation for depression*.

Martínez-Vispo, C., Martínez, Ú., López-Durán, A., Fernández del Río, E., & Becoña, E. (2018). Effects of behavioural activation on substance use and depression: A systematic review. *Substance Abuse Treatment, Prevention, and Policy*, 13(1), 36.

Mata, J., Thompson, R. J., Jaeggi, S. M., Buschkuhl, M., Jonides, J., & Gotlib, I. H. (2012). Walk on the bright side: Physical activity and affect in major depressive disorder. *Journal of Abnormal Psychology*, 121(2), 297.

Mathers, C. D., & Loncar, D. (2006). Projections of Global Mortality and Burden of Disease from 2002 to 2030. *PLoS Medicine*, 3(11), e442.

Mazzucchelli, T., Kane, R., & Rees, C. (2009). Behavioral Activation Treatments for Depression in Adults: A Meta-analysis and Review. *Clinical Psychology: Science and Practice*, 16(4), 383–411.

McDowell, I. (2006). *Measuring health: A guide to rating scales and questionnaires*. Oxford University Press, USA.

McGowan, N. M., Goodwin, G. M., Bilderbeck, A. C., & Saunders, K. E. (2020). Actigraphic patterns, impulsivity and mood instability in bipolar disorder, borderline personality disorder and healthy controls. *Acta Psychiatrica Scandinavica*, 141(4), 374-384.

Meléndez, J. C., Satorres, E., Reyes-Olmedo, M., Delhom, I., Real, E., & Lora, Y. (2020). Emotion recognition changes in a confinement situation due to COVID-19. *Journal of Environmental Psychology*, 72, 101518.

Messer, S. B., & Wampold, B. E. (2002). Let's face facts: Common factors are more potent than specific therapy ingredients. *Clinical Psychology: Science and Practice*, 9(1), 21–25.

Metzger, R. L. (1976). A reliability and validity study of the State-Trait Anxiety Inventory. *Journal of Clinical Psychology*.

Miguel JH, Rowlands AV, et al. GGIR: A Research Community–Driven Open Source R Package for Generating Physical Activity and Sleep Outcomes From Multi-Day Raw Accelerometer Data. *Journal for the Measurement of Physical Behaviour*. 2(3) 2019. doi: 10.1123/jmpb.2018-0063.

Mihashi, M., Otsubo, Y., Yinjuan, X., Nagatomi, K., Hoshiko, M., & Ishitake, T. (2009). Predictive factors of psychological disorder development during recovery following SARS outbreak. *Health Psychology*, 28(1), 91.

Milders, M., Bell, S., Platt, J., Serrano, R., & Runcie, O. (2010). Stable expression recognition abnormalities in unipolar depression. *Psychiatry Research*, 179(1), 38–42.

Molenberghs, G., Thijs, H., Jansen, I., Beunckens, C., Kenward, M. G., Mallinckrodt, C., & Carroll, R. J. (2004). Analyzing incomplete longitudinal clinical trial data. *Biostatistics*, 5(3), 445-464.

Murphy, S. E., Wright, L. C., Browning, M., Cowen, P. J., & Harmer, C. J. (2020). A role for 5-HT₄ receptors in human learning and memory. *Psychological Medicine*, 50(16), 2722–2730.

Mykletun, A., Overland, S., Dahl, A. A., Krokstad, S., Bjerkeset, O., Glozier, N., Aarø, L. E., & Prince, M. (2006). A Population-Based Cohort Study of the Effect of Common Mental Disorders on Disability Pension Awards. *American Journal of Psychiatry*, 163(8), 1412–1418.

Nakonezny, P. A., Carmody, T. J., Morris, D. W., Kurian, B. T., & Trivedi, M. H. (2010). Psychometric evaluation of the Snaith–Hamilton Pleasure Scale (SHAPS) in adult outpatients with major depressive disorder. *International Clinical Psychopharmacology*, 25(6), 328–333.

Nasrin, F., Rimes, K., Reinecke, A., Rinck, M., & Barnhofer, T. (2017). Effects of Brief Behavioural Activation on Approach and Avoidance Tendencies in Acute Depression: Preliminary Findings. *Behavioural and Cognitive Psychotherapy*, 45(1), 58–72.

National Institute of Health and Clinical Excellence. (2009). *Depression in adults: Recognition and management (Clinical guideline [CG90])*. <https://www.nice.org.uk/guidance/cg90>

Netemeyer, R. G., Williamson, D. A., Burton, S., Biswas, D., Jindal, S., Landreth, S., Mills, G., & Primeaux, S. (2002). Psychometric properties of shortened versions of the Automatic Thoughts Questionnaire. *Educational and Psychological Measurement*, 62(1), 111–129.

Oei, T. P., Evans, L., & Crook, G. M. (1990). Utility and validity of the STAI with anxiety disorder patients. *British Journal of Clinical Psychology*, 29(4), 429–432.

Orgeta, V., Brede, J., & Livingston, G. (2017). Behavioural activation for depression in older people: Systematic review and meta-analysis. *British Journal of Psychiatry*, 211(5), 274–279.

Pace-Schott, E. F., & Hobson, J. A. (2002). The neurobiology of sleep: Genetics, cellular physiology and subcortical networks. *Nature Reviews Neuroscience*, 3(8), 591–605.

Patel, A. R., Weobong, B., Patel, V. H., & Singla, D. R. (2019). Psychological treatments for depression among women experiencing intimate partner violence: findings from a randomized controlled trial for behavioral activation in Goa, India. *Archives of women's mental health*, 22(6), 779–789.

Patel, V., Weobong, B., Weiss, H. A., Anand, A., Bhat, B., Katti, B., Dimidjian, S., Araya, R., Hollon, S. D., King, M., Vijayakumar, L., Park, A.-L., McDaid, D., Wilson, T., Velleman, R., Kirkwood, B. R., & Fairburn, C. G. (2017). The Healthy Activity Program (HAP), a lay counsellor-delivered brief psychological treatment for severe depression, in primary care in India: A randomised controlled trial. *The Lancet*, 389(10065), 176–185.

Pentecost, C., Farrand, P., Greaves, C. J., Taylor, R. S., Warren, F. C., Hillsdon, M., Green, C., Welsman, J. R., Rayson, K., Evans, P. H., & Taylor, A. H. (2015). Combining behavioural activation with physical activity promotion for adults with depression: Findings of a parallel-group pilot randomised controlled trial (BACPAc). *Trials*, 16(1), 367.

Piehl, C., Budimir, S., & Probst, T. (2020). The effect of age, gender, income, work, and physical activity on mental health during coronavirus disease (COVID-19) lockdown in Austria. *Journal of Psychosomatic Research*, 136, 110186.

Pierce, M., Hope, H., Ford, T., Hatch, S., Hotopf, M., John, A., Kontopantelis, E., Webb, R., Wessely, S., McManus, S., & Abel, K. M. (2020). Mental health before and during the COVID-19 pandemic: A longitudinal probability sample survey of the UK population. *The Lancet Psychiatry*, 7(10), 883–892.

Pizzagalli, D. A., Iosifescu, D., Hallett, L. A., Ratner, K. G., & Fava, M. (2008). Reduced hedonic capacity in major depressive disorder: Evidence from a probabilistic reward task. *Journal of Psychiatric Research*, 43(1), 76–87.

Pizzagalli, D. A., Jahn, A. L., & O'Shea, J. P. (2005). Toward an Objective Characterization of an Anhedonic Phenotype. *Biological Psychiatry*, 57(4), 319–327.

Porter, R. J., Bourke, C., Carter, J. D., Douglas, K. M., McIntosh, V. V. W., Jordan, J., Joyce, P. R., & Frampton, C. M. A. (2016). No change in neuropsychological dysfunction or emotional processing during treatment of major depression with cognitive–behaviour therapy or schema therapy. *Psychological Medicine*, 46(2), 393–404.

Purrington, J., & Beail, N. (2021). The impact of Covid-19 on access to psychological services. *Advances in Mental Health and Intellectual Disabilities*.

Pye, J., Phillips, A. J., Cain, S. W., Montazerolghaem, M., Mowszowski, L., Duffy, S., ... & Naismith, S. L. (2021). Irregular sleep-wake patterns in older adults with current or remitted depression. *Journal of Affective Disorders*, 281, 431-437.

Pyne, J. M., Patterson, T. L., Kaplan, R. M., Gillin, J. C., Koch, W. L., & Grant, I. (1997). Assessment of the quality of life of patients with major depression. *Psychiatric Services*, 48(2), 224–230.

Raoux, N., Benoit, O., Dantchev, N., Denise, P., Franc, B., Allilaire, J., & Widlöcher, D. (1994). Circadian pattern of motor activity in major depressed patients undergoing antidepressant therapy: Relationship between actigraphic measures and clinical course. *Psychiatry Research*, 52, 85–98.

Rehm, L. P. (1978). Mood, pleasant events, and unpleasant events: Two pilot studies. *Journal of Consulting and Clinical Psychology*, 46(5), 854–859.

Reinecke, A., Waldenmaier, L., Cooper, M. J., & Harmer, C. J. (2013). Changes in automatic threat processing precede and predict clinical changes with exposure-based cognitive-behavior therapy for panic disorder. *Biological Psychiatry*, 73(11), 1064-1070.

Reynolds, D. L., Garay, J. R., Deamond, S. L., Moran, M. K., Gold, W., & Styra, R. (2008). Understanding, compliance and psychological impact of the SARS quarantine experience. *Epidemiology and Infection*, 136(7), 997–1007.

Richards, D. A., Ekers, D., McMillan, D., Taylor, R. S., Byford, S., Warren, F. C., Barrett, B., Farrand, P. A., Gilbody, S., Kuyken, W., O'Mahen, H., Watkins, E. R., Wright, K. A., Hollon, S. D., Reed, N., Rhodes, S., Fletcher, E., & Finning, K. (2016). Cost and Outcome of Behavioural Activation versus Cognitive Behavioural Therapy for Depression (COBRA): A randomised, controlled, non-inferiority trial. *The Lancet*, 388(10047), 871–880.

Riedel, M., Möller, H. J., Obermeier, M., Schennach-Wolff, R., Bauer, M., Adli, M., ... & Seemüller, F. (2010). Response and remission criteria in major depression—a validation of current practice. *Journal of psychiatric research*, 44(15), 1063-1068.

Roiser, J., & Sahakian, B. (2013). Hot and cold cognition in depression. *CNS Spectrums*, 18, 1–11.

Ruhe, H. G., Mocking, R. J. T., Figueroa, C. A., Seeverens, P. W. J., Ikani, N., Tyborowska, A., Browning, M., Vrijssen, J. N., Harmer, C. J., & Schene, A. H. (2019). Emotional Biases and Recurrence in Major Depressive Disorder. Results of 2.5 Years Follow-Up of Drug-Free Cohort Vulnerable for Recurrence. *Frontiers in Psychiatry*, 10.

Rush, A. J. (2007). Limitations in efficacy of antidepressant monotherapy. *Journal of Clinical Psychiatry*, 68(B), 8.

Russo, G. B., Tirrell, E., Busch, A., & Carpenter, L. L. (2018). Behavioral activation therapy during transcranial magnetic stimulation for major depressive disorder. *Journal of Affective Disorders*, 236, 101–104.

Ruzickova, T., Carson, J., Argabright, S., Gillespie, A., Guinea, C., Pearse, A., ... & Harmer, C. J. (2021). Online behavioural activation during the COVID-19 pandemic decreases depression and negative affective bias. *Psychological medicine*, 1-10.

Sabia S, van Hees VT, Shipley MJ, Trenell MI, Hagger-Johnson G, Elbaz A, Kivimaki M, Singh-Manoux A. Association between questionnaire- and accelerometer-assessed physical activity: the role of sociodemographic factors. *Am J Epidemiol*. 2014 Mar 15;179(6):781-90.

Schuch, F., Vancampfort, D., Firth, J., Rosenbaum, S., Ward, P., Reichert, T., Bagatini, N. C., Bgeginski, R., & Stubbs, B. (2017). Physical activity and sedentary behavior in people with major depressive disorder: A systematic review and meta-analysis. *Journal of Affective Disorders*, 210, 139–150.

Segal, Z. V., & Gemar, M. (1997). Changes in cognitive organisation for negative self-referent material following cognitive behaviour therapy for depression: A primed Stroop study. *Cognition & Emotion*, 11(5–6), 501–516.

Shear, M. K., Pilkonis, P. A., Cioitre, M., & Leon, A. C. (1994). Cognitive behavioral treatment compared with nonprescriptive treatment of panic disorder. *Archives of General Psychiatry*, 51(5), 395–401.

Shevlin, M., Nolan, E., Owczarek, M., McBride, O., Murphy, J., Miller, J. G., Hartman, T. K., Levita, L., Mason, L., Martinez, A. P., McKay, R., Stocks, T. V. A., Bennett, K. M., Hyland, P., & Bentall, R. P. (2020). COVID-19-related anxiety predicts somatic symptoms in the UK population. *British Journal of Health Psychology*, 25(4), 875–882.

Singla, D. R., Weobong, B., Nadkarni, A., Chowdhary, N., Shinde, S., Anand, A., Fairburn, C. G., Dimijdan, S., Velleman, R., Weiss, H., & Patel, V. (2014). Improving the scalability of psychological treatments in developing countries: An evaluation of peer-led therapy quality assessment in Goa, India. *Behaviour Research and Therapy*, 60, 53–59.

Slyepchenko, A., Allega, O. R., Leng, X., Minuzzi, L., Eltayebani, M. M., Skelly, M., ... & Frey, B. N. (2019). Association of functioning and quality of life with objective and subjective measures of sleep and biological rhythms in major depressive and bipolar disorder. *Australian & New Zealand Journal of Psychiatry*, 53(7), 683–696.

Snaith, R. P., Hamilton, M., Morley, S., Humayan, A., Hargreaves, D., & Trigwell, P. (1995). A Scale for the Assessment of Hedonic Tone the Snaith–Hamilton Pleasure Scale. *The British Journal of Psychiatry; London*, 167(1), 99–103.

Solomou, I., & Constantinidou, F. (2020). Prevalence and predictors of anxiety and depression symptoms during the COVID-19 pandemic and compliance with precautionary measures: Age and sex matter. *International Journal of Environmental Research and Public Health*, 17(14), 4924.

Spielberger, C. D. (1970). Manual for the State-Trait Anxiety, Inventory. *Consulting Psychologist*.

Spijker, J., Graaf, R., Bijl, R. V., Beekman, A. T. F., Ormel, J., & Nolen, W. A. (2004). Functional disability and depression in the general population. Results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Acta Psychiatrica Scandinavica*, 110(3), 208–214.

Stange, J. P., Zulueta, J., Langenecker, S. A., Ryan, K. A., Piscitello, A., Duffecy, J., McInnis, M. G., Nelson, P., Ajilore, O., & Leow, A. (2018). Let your fingers do the talking: Passive typing instability predicts future mood outcomes. *Bipolar Disorders*, 20(3), 285–288.

Stein, A. T., Carl, E., Cuijpers, P., Karyotaki, E., & Smits, J. A. J. (2021). Looking beyond depression: A meta-analysis of the effect of behavioral activation on depression, anxiety, and activation. *Psychological Medicine*, 51(9), 1491–1504.

Sullivan, L. E., Fiellin, D. A., & O'Connor, P. G. (2005). The prevalence and impact of alcohol problems in major depression: A systematic review. *The American Journal of Medicine*, 118(4), 330–341.

Taquet, M., Luciano, S., Geddes, J. R., & Harrison, P. J. (2021). Bidirectional associations between COVID-19 and psychiatric disorder: Retrospective cohort studies of 62 354 COVID-19 cases in the USA. *The Lancet Psychiatry*, 8(2), 130–140.

Teychenne, M., Ball, K., & Salmon, J. (2010). Sedentary behavior and depression among adults: A review. *International Journal of Behavioral Medicine*, 17(4), 246–254.

Tindall, L., Mikocka-Walus, A., McMillan, D., Wright, B., Hewitt, C., & Gascoyne, S. (2017). Is behavioural activation effective in the treatment of depression in young people? A systematic review and meta-analysis. *Psychology and Psychotherapy: Theory, Research and Practice*, 90(4), 770–796.

Tison, G. H., Avram, R., Kuhar, P., Abreau, S., Marcus, G. M., Pletcher, M. J., & Olgin, J. E. (2020). Worldwide Effect of COVID-19 on Physical Activity: A Descriptive Study. *Annals of Internal Medicine*, 173(9), 767–770.

Todder, D., Caliskan, S., & Baune, B. T. (2009). Longitudinal changes of day-time and night-time gross motor activity in clinical responders and non-responders of major depression. *The World Journal of Biological Psychiatry*, 10(4), 276–284.

Tranter, R., Bell, D., Gutting, P., Harmer, C., Healy, D., & Anderson, I. M. (2009). The effect of serotonergic and noradrenergic antidepressants on face emotion processing in depressed patients. *Journal of Affective Disorders*, 118(1), 87–93.

Treadway, M. T., & Zald, D. H. (2011). Reconsidering anhedonia in depression: Lessons from translational neuroscience. *Neuroscience & Biobehavioral Reviews*, 35(3), 537–555.

Tsanas, A., Saunders, K. E. A., Bilderbeck, A. C., Palmius, N., Osipov, M., Clifford, G. D., Goodwin, G. M., & De Vos, M. (2016). Daily longitudinal self-monitoring of mood variability in bipolar disorder and borderline personality disorder. *Journal of Affective Disorders*, 205, 225–233.

Uher, R., Perlis, R. H., Henigsberg, N., Zobel, A., Rietschel, M., Mors, O., ... & McGuffin, P. (2012). Depression symptom dimensions as predictors of antidepressant treatment outcome: replicable evidence for interest-activity symptoms. *Psychological medicine*, 42(5), 967-980.

van Hees VT, Gorzelniak L, et al. Separating Movement and Gravity Components in an Acceleration Signal and Implications for the Assessment of Human Daily Physical Activity. *PLoS ONE* 8(4) 2013.

Van Vugt, M., Hitchcock, P., Shahar, B., & Britton, W. (2012). The Effects of Mindfulness-Based Cognitive Therapy on Affective Memory Recall Dynamics in Depression: A Mechanistic Model of Rumination. *Frontiers in Human Neuroscience*, 6.

Vancampfort, D., Stubbs, B., Sienaert, P., Wyckaert, S., De Hert, M., Rosenbaum, S., & Probst, M. (2015). What are the factors that influence physical activity participation in individuals with depression? A review of physical activity correlates from 59 studies. *Psychiatria Danubina*, 27(3), 0-224.

Vazquez, C., Duque, A., Blanco, I., Pascual, T., Poyato, N., Lopez-Gomez, I., & Chaves, C. (2018). CBT and positive psychology interventions for clinical depression promote healthy attentional biases: An eye-tracking study. *Depression and Anxiety*, 35(10), 966–973.

Vinckier, F., Gourion, D., & Mouchabac, S. (2017). Anhedonia predicts poor psychosocial functioning: results from a large cohort of patients treated for major depressive disorder by general practitioners. *European Psychiatry*, 44, 1-8.

Vos, T., Abajobir, A. A., Abate, K. H., Abbafati, C., Abbas, K. M., Abd-Allah, F., Abdulkader, R. S., Abdulle, A. M., Abebo, T. A., Abera, S. F., Aboyans, V., Abu-Raddad, L. J., Ackerman, I. N., Adamu, A. A., Adetokunboh, O., Afarideh, M., Afshin, A., Agarwal, S. K., Aggarwal, R., ... Murray, C. J. L. (2017). Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet*, 390(10100), 1211–1259.

Walsh, A. E., Browning, M., Drevets, W. C., Furey, M., & Harmer, C. J. (2018). Dissociable temporal effects of bupropion on behavioural measures of emotional and reward processing in depression. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1742), 20170030.

Walsh, E. C., Eisenlohr-Moul, T. A., Minkel, J., Bizzell, J., Petty, C., Crowther, A., Carl, H., Smoski, M. J., & Dichter, G. S. (2019). Pretreatment brain connectivity during positive emotion upregulation predicts decreased anhedonia following behavioral activation therapy for depression. *Journal of Affective Disorders*, 243, 188–192.

Wang, C., Pan, R., Wan, X., Tan, Y., Xu, L., Ho, C. S., & Ho, R. C. (2020). Immediate psychological responses and associated factors during the initial stage of the 2019 coronavirus disease (COVID-19) epidemic among the general population in China. *International Journal of Environmental Research and Public Health*, 17(5), 1729.

Wang, Y.-P., & Gorenstein, C. (2013). Psychometric properties of the Beck Depression Inventory-II: A comprehensive review. *Brazilian Journal of Psychiatry*, 35, 416–431.

Webb, C. A., Stanton, C. H., Bondy, E., Singleton, P., Pizzagalli, D. A., & Auerbach, R. P. (2019). Cognitive versus behavioral skills in CBT for depressed adolescents: Disaggregating within-patient versus between-patient effects on symptom change. *Journal of Consulting and Clinical Psychology, 87*(5), 484.

Wichers, M., Peeters, F., Rutten, B. P., Jacobs, N., Derom, C., Thiery, E., Delespaul, P., & van Os, J. (2012). A time-lagged momentary assessment study on daily life physical activity and affect. *Health Psychology, 31*(2), 135.

Winkler, P., Formanek, T., Mlada, K., Kagstrom, A., Mohrova, Z., Mohr, P., & Csemy, L. (2020). Increase in prevalence of current mental disorders in the context of COVID-19: Analysis of repeated nationwide cross-sectional surveys. *Epidemiology and Psychiatric Sciences, 29*, e173.

Witting, W., Kwa, I. H., Eikelenboom, P., Mirmiran, M., & Swaab, D. F. (1990). Alterations in the circadian rest-activity rhythm in aging and Alzheimer's disease. *Biological Psychiatry, 27*(6), 563–572.

World Health Organization. (2008). *mhGAP: Mental Health Gap Action Programme: scaling up care for mental, neurological and substance use disorders*. World Health Organization.

Wüthrich, F., Nabb, C. B., Mittal, V. A., Shankman, S. A., & Walther, S. (2022). Actigraphically measured psychomotor slowing in depression: systematic review and meta-analysis. *Psychological medicine, 1-14*.

Xu, X., Kwan, R., & Leung, A. Y. M. (2019). Behavioural activation for family caregivers of people with dementia: a systematic review and meta-analysis. *Innovation in Aging, 3* (Suppl 1), S109.

Yılmaz, O., Mirçık, A. B., Kunduz, M., Çombaş, M., Öztürk, A., Deveci, E., & Kırpınar, İ. (2019). Effects of Cognitive Behavioral Therapy, Existential Psychotherapy and Supportive Counselling on Facial Emotion Recognition Among Patients with Mild or Moderate Depression. *Psychiatry Investigation, 16*(7), 491–503.

Yoshimura, S., Okamoto, Y., Onoda, K., Matsunaga, M., Okada, G., Kunisato, Y., Yoshino, A., Ueda, K., Suzuki, S., & Yamawaki, S. (2014). Cognitive behavioral therapy for depression changes medial prefrontal and ventral anterior cingulate cortex activity associated with self-referential processing. *Social Cognitive and Affective Neuroscience*, 9.

Zabihi, S., Lemmel, F. K., & Orgeta, V. (2020). Behavioural Activation for Depression in Informal Caregivers: A Systematic Review and Meta-Analysis of Randomised Controlled Clinical Trials. *Journal of Affective Disorders*, 274, 1173–1183.

Zimet, G. D., Dahlem, N. W., Zimet, S. G., & Farley, G. K. (1988). The multidimensional scale of perceived social support. *Journal of Personality Assessment*, 52(1), 30–41.

Zimet, G. D., Powell, S. S., Farley, G. K., Werkman, S., & Berkoff, K. A. (1990). Psychometric characteristics of the multidimensional scale of perceived social support. *Journal of Personality Assessment*, 55(3–4), 610–617.