Investigating Interventions for Affect and Cognition:

Three Randomized Controlled Trials

– Kumulative Arbeit –

Inaugural-Dissertation

zur Erlangung der Doktorwürde

der

Philosophischen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität zu Bonn

vorgelegt von

Julia Fabienne Sandkühler

aus Essen

Bonn, 2025

Gedruckt mit der Genehmigung der Philosophischen Fakultät

der Rheinischen Friedrich-Wilhelms-Universität Bonn

Zusammensetzung der Prüfungskommission:

Prof. Dr. André Beauducel (Vorsitzender)

Prof. Dr. Ulrich Ettinger (Betreuer und Gutachter)

Prof. Dr. Michael Wagner (Gutachter)

PD Dr. Dr. Cornelia McCormick (weiteres prüfungsberechtigtes Mitglied)

Tag der mündlichen Prüfung: 29.08.2025

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I am deeply thankful to my supervisor Ulrich Ettinger for his intellectual support, words of encouragement, and openness and flexibility with my PhD projects and my work location. I would also like to thank Prof. Dr. André Beauducel, Prof. Dr. Michael Wagner, and PD Dr. Dr. Cornelia McCormick for being part of my examination board. Thank you to my Bonn colleagues for warmly including me during every visit, for the laughter, and for the great conversations. A big thank you goes to all of my collaborators, in particular Jan Brauner, who greatly helped me become a productive researcher. I would like to thank our meticulous and hard-working double data entry assistants and other helpers of the studies. Thank you to all the study participants. I sincerely thank the funders of the studies – Paul Christiano, Matt Wage, the Effective Ventures Foundation, and Mind Ease Labs – for making these studies possible.

I would like to thank anyone who has reached out and shown interest in my studies – this has been very encouraging. Thank you to my 12-week-year buddies Elspeth and Marly for our very helpful weekly productivity meetings. I thank my parents and my sister for their curiosity about and enthusiasm for my work. I thank my husband and best friend for his loving and caring support throughout the PhD, for celebrating every success, and for always being there for me.

ABSTRACT

Improving mental health and cognition is highly desirable due to the important impact they have on individual quality of life and societal functioning. Making mental health and cognitive enhancement interventions accessible, affordable, and low-effort is crucial to ensuring that they reach as many people as possible. There are many open questions in this area. Some of them are: How can we improve the effectiveness of light therapy against seasonal affective disorder? Does “more light” help more? Does supplementing creatine improve cognition? How effective are self-help apps against anxiety? How do different psychotherapeutic exercises compare to each other?

This thesis addresses these questions by presenting results from three studies designed to test accessible interventions for mental health and cognitive improvement. All studies were randomised, controlled, and double-blind or partially blind, and followed the CONSORT reporting guidelines and Open Science principles. Data were analysed using frequentist (standard and robust) statistics as well as Bayesian statistics.

Study 1, a feasibility study, found that BRight, whole-ROom, All-Day (BROAD) light therapy against seasonal affective disorder was feasible and that the effectiveness of the therapy correlated positively with illuminance at eye level. Study 2 was somewhat inconclusive, with no significant effect of creatine on cognition (frequentist statistics), weak evidence for a small effect, and strong evidence against a large effect (Bayesian statistics). Study 3 found that the 12 exercises of the Mind Ease self-help app against anxiety, based on cognitive behavioural therapy, mindfulness, and acceptance and commitment therapy, had large effects on immediate anxiety levels. Mindfulness exercises had particularly large effects. Overall, a broad picture of interventions to improve mental health and cognition is provided and best practices in methods and statistics shared by the three studies are discussed.

ZUSAMMENFASSUNG

Die Verbesserung der psychischen Gesundheit und der kognitiven Fähigkeiten ist äußerst wünschenswert, da sie einen großen Einfluss auf die individuelle Lebensqualität und das Funktionieren der Gesellschaft haben. Um sicherzustellen, dass möglichst viele Menschen davon profitieren, ist es wichtig, dass Maßnahmen zur Verbesserung der psychischen Gesundheit und der kognitiven Fähigkeiten zugänglich, erschwinglich und mit geringem Aufwand verbunden sind. In diesem Bereich gibt es viele offene Fragen. Einige davon sind: Wie können wir die Wirksamkeit der Lichttherapie bei der saisonal-affektiven Störung verbessern? Hilft „mehr Licht“ mehr? Verbessert die Einnahme von Kreatin die Kognition? Wie wirksam sind Selbsthilfe-Apps gegen Angstzustände? Wie schneiden verschiedene psychotherapeutische Übungen im Vergleich zueinander ab?

Die vorliegende Arbeit nimmt sich dieser Fragen an, indem sie die Ergebnisse von drei Studien vorstellt, in denen zugängliche Interventionen zur Verbesserung der psychischen Gesundheit und der kognitiven Fähigkeiten getestet wurden. Alle Studien waren randomisiert, kontrolliert und doppelblind oder teilverblindet und folgten den CONSORT-Berichterstattungsrichtlinien und den Open Science Grundsätzen. Die Daten wurden mit Hilfe der frequentistischen (Standard- und robusten) Statistik sowie der Bayesischen Statistik analysiert.

Studie 1, eine Machbarkeitsstudie, ergab, dass die BROAD-Lichttherapie (BRight, whole-ROom, All-Day) gegen die saisonal-affektive Störung machbar war und die Wirksamkeit der Therapie positiv mit der Beleuchtungsstärke auf Augenhöhe korrelierte. Studie 2 war nicht eindeutig, denn sie ergab keine signifikante Wirkung von Kreatin auf die Kognition (frequentistische Statistik), aber schwache Evidenz für eine kleine Wirkung und starke Evidenz gegen eine große Wirkung (Bayesische Statistik). Studie 3 ergab, dass die 12 Übungen der Mind Ease App zur Selbsthilfe gegen Angst, die auf der kognitiven

Verhaltenstherapie, der Achtsamkeit und der Akzeptanz- und Commitment-Therapie basieren, große Auswirkungen auf das unmittelbare Angstniveau hatten. Achtsamkeitsübungen hatten besonders große Effekte. Insgesamt wird ein breites Bild von Interventionen zur Verbesserung der psychischen Gesundheit und der Kognition vorgestellt, und es werden best practice Methoden und Statistiken aus den drei Studien diskutiert.

1 INTRODUCTION

1.1 Improving mental health and cognition

Improving mental health and cognition is highly desirable due to the important impact they have on functioning and quality of life. The American Psychological Association defines mental health as “a state of mind characterized by emotional well-being, good behavioral adjustment, relative freedom from anxiety and disabling symptoms, and a capacity to establish constructive relationships and cope with the ordinary demands and stresses of life” (American Psychological Association, 2018b). Cognition, on the other hand, refers to “all forms of knowing and awareness, such as perceiving, conceiving, remembering, reasoning, judging, imagining, and problem solving” (American Psychological Association, 2018a). Together, mental health and cognition are essential for functioning across all domains of life, from personal goals and relationships to professional performance, and improving them may offer benefits both on an individual and societal level. This dissertation aims to investigate innovative and accessible interventions for improving mental health and cognition.

1.1.1 Mental health

Mental health conditions are highly prevalent. About 1 in 8 people, almost 1 billion people worldwide, live with a mental health condition. Depression (over 300 million people in 2017) and anxiety are the most common mental health conditions and increased by 25% during the first year of the COVID-19 pandemic (World Health Organization, 2022).

Mental health is essential to quality of life. Years of healthy life lost to disability (YLDs) measure loss of quality of life by assigning weights to different conditions multiplying the weight by the duration. Globally, depression is responsible for 7.5% and anxiety for 3.4% of all YLDs (World Health Organization, 2017). In total, mental health conditions are

responsible for 1 in 6 years of YLDs (World Health Organization, 2022). In 2015, the United Nations included improving mental health in their Sustainable Development Goals, which means they consider it a priority (Votruba et al., 2016).

Mental health treatment works to some extent. There are a number of interventions that have been found to improve mental health conditions, such as psychotherapy and medication (Roshanaei-Moghaddam et al., 2011). Overall, psychotherapy and psychiatric medication have been found to be similarly effective with effect sizes of about half a standard deviation, although the exact size of the effect and which treatment performs best varies between disorders (Huhn et al., 2014). This is similar to the effectiveness of general medical drugs, which also produce an effect of half a standard deviation on average (Leucht et al., 2012).

Many people do not have access to mental health treatments, so scaling them up is important (more on this in sections 1.2 and 1.4.3). However, these interventions are not always fully successful, so improving them is also desirable. For example, a recurrent type of depression called seasonal affective disorder has been found to improve with light therapy, but in many cases, this treatment only partially restores mental health (more on this in section 1.4.1).

It is more expensive *not* to treat mental health conditions than to treat them. It is estimated that for every euro we as a society invest into scaling up the treatment of depression and anxiety, we save 2.3-3.0 euros (Chisholm et al., 2016). Hence, expanding treatment provision is not only defensible on compassionate grounds but also constitutes a sound economic decision.

1.1.2 Cognition

Cognition is crucial to functioning. Cognitive performance predicts academic, professional, and financial success (Bertua et al., 2005; Ng et al., 2005; Roth et al., 2015; Strenze, 2007;

Zaboski et al., 2018). Better cognitive performance is also associated with better health, longevity, and well-being (Gottfredson & Deary, 2004; Sternberg et al., 2001) (reviewed by Breit et al. (2022)).

Given how important cognitive performance is, it is not surprising that there have been many attempts to improve it. Some of these attempts were very successful. For example, governments and non-governmental organisations have successfully improved the cognitive performance of their citizens with large scale interventions aimed at increasing iodine intake (Qian et al., 2005), eliminating lead exposure (Soong et al., 1999), preventing diseases such as malaria (Fernando et al., 2010; Pierre-Louis et al., 2010), and of course improving (access to) education (Sternberg et al., 2001). Some of these interventions might be considered to alleviate cognitive impairment, while interventions such as education might be considered to enhance cognition. While this distinction is currently widely used e.g. in medical insurance, it has also been contested as hard to draw, arbitrary and not a good guide for how to prioritise interventions (Bostrom & Sandberg, 2009; Daniels, 2000). For example, an individual with a “naturally” lower ability using an “enhancement” method might still perform worse than an individual with a “naturally” high ability (possibly even if the latter acquires a nutritional deficiency, illness, or injury that lowers performance). A nutritional supplement with only preliminary evidence supporting its effectiveness might be regarded as an optional enhancer, but if the evidence strengthened, and especially if it was a nutrient found “naturally” in some diets, it seems conceivable that it would not be long until a low level of this nutrient would be considered a deficiency. I use the term “improve” in a neutral way that applies to all levels of functioning. Interventions individuals can take to improve their cognitive performance include exercise (Chang et al., 2012; Etnier & Chang, 2019), good sleep (Alhola & Polo-Kantola, 2007; Killgore & Weber, 2014), and ventilation (Du et al., 2020). A question that has received considerable public interest is whether cognitive performance can be improved with other supplements. Many compounds have been

suggested but only very few, such as creatine (more on this in section 1.4.3), seem at least somewhat promising.

1.1.3 Link between mental health and cognition

As mentioned before, mental health and cognitive performance correlate positively with each other. For example, cognitive performance has been found to be associated with positive affect and life satisfaction (Ali et al., 2013), lower self-rated levels of depression and other mental health problems (Khandaker et al., 2018) and fewer mental health conditions (Batty et al., 2005) (reviewed in Jokela (2022)).

Many interventions improve both mental health and cognitive performance. A good example for this is exercise, which has considerable effects on both cognitive performance and mental health conditions, especially anxiety and depression (Mikkelsen et al., 2017; Parry et al., 2018). Mindfulness meditation, well-known for improving mental health, has also been found to improve cognitive performance (Zeidan et al., 2010). Light therapy, a first-line treatment for seasonal affective disorder and possibly of help in other depressive disorders, has been found to also improve alertness and executive function in healthy individuals as well as in individuals with seasonal affective disorder (Huang et al., 2024). There is some very preliminary evidence that creatine might improve both cognitive performance in healthy adults (more on this in section 1.4.2) and depression (Bakian et al., 2020).

1.2 Accessibility

Making tools to improve mental health and cognition widely available, low-cost, and low effort is valuable for several reasons, including increasing inclusivity, improving adherence, and offering scalable solutions to large-scale public health challenges.

Many individuals face barriers to mental health and cognition interventions (World Health Organization, 2022). Psychotherapy can be unaffordable if not covered by insurance. Due to

the low supply of psychotherapists relative to demand, appointments can be hard to obtain. This is true in high-income countries and even more so in middle- and low-income countries. In particular in rural areas, distance to the provider can hinder treatment. Stigma is another main barrier to mental health treatment. Overcoming these barriers is the topic of Study 3.

Seasonal affective disorder, which is depression recurring in winter, might be avoided by moving to a different climate zone (Rosen et al., 1990). However, for many individuals, this solution is not feasible due to cost or personal or professional attachments and commitments. Standard light therapy lamps do not have these barriers, but they limit movement for a while in the morning, as patients have to sit with their eyes very close (e.g. 10-20cm) to the lamp, and they often do not cause full remittance (Pjrek et al., 2020). More summer-like conditions might be achieved by spending all day in a clinic light room, but for many patients this is not available or practical. Study 1 investigated an accessible alternative to the clinic light room (more on this in section 1.4.3).

Supplements for cognitive enhancement can be expensive (and often at the same time, lacking in evidence) (Global Council on Brain Health, 2019). Other interventions, such as transcranial magnetic stimulation or mental training, are similarly unavailable to many. Exercise and improvements in sleep are in theory more accessible, but many find them hard to implement. Study 2 tests a cheap, widely available and low effort potential intervention for cognitive enhancement – creatine supplementation.

1.3 Neural mechanisms

In the following, I will describe the neural mechanisms underlying the interventions used in the studies included in this thesis. All three interventions target the prefrontal cortex (PFC).

1.3.1 Light therapy for seasonal affective disorder

Major depressive disorder, like all depressive disorders, is characterised by “sad, empty, or irritable mood, accompanied by somatic and cognitive changes that significantly affect the individual's capacity to function” (American Psychiatric Association, 2013). These symptoms have to last for at least two weeks at a time. Study 1 is about light therapy for a subtype of major depressive disorder: major depressive disorder (recurrent episode) with a winter seasonal pattern, also known as seasonal affective disorder (SAD). SAD patients have most of their episodes of major depressive disorder in autumn/winter and remit in spring/summer. SAD occurs in places with periods of little sunlight, i.e. in places that are far from the equator. Diagnostic interviews suggest the rate of SAD stands at around 3.4% in Zurich, Switzerland (Wirz-Justice et al., 2019), 2.9% in Toronto, Canada (Levitt et al., 2000), and 2.4% in North Wales (Michalak et al., 2001) (Blazer et al., 1998). Lack of sunlight is thought to be a primary cause of SAD (Levitan, 2007).

Light therapy with bright white light (commonly 10,000 lux provided by a light therapy lamp for half an hour in the morning) has been found to reduce the symptoms of SAD. Lux is a measure of illuminance defined as one lumen evenly distributed over one square meter. Lumen is a measure of the amount of visible light emitted by a bulb. For reference, typically rooms have an illuminance of 90-180 lux (Cajochen et al., 2000).

According to the phase-shifting hypothesis, one mechanism through which light therapy benefits seasonal affective disorder is by shifting the circadian rhythm, i.e. the day-night cycle of the body. Light affects the circadian rhythm in two ways (Lewy et al., 2006):

- The information that light is entering the eye is transported to the suprachiasmatic nucleus (SCN) of the hypothalamus (Figure 1). The SCN contains a group of neurons and is found above the optic chiasm. The SCN is the body’s master pacemaker or biological clock, regulating the circadian rhythm of many processes in the body.
- Light entering the eye suppresses the secretion of the sleep-hormone melatonin.

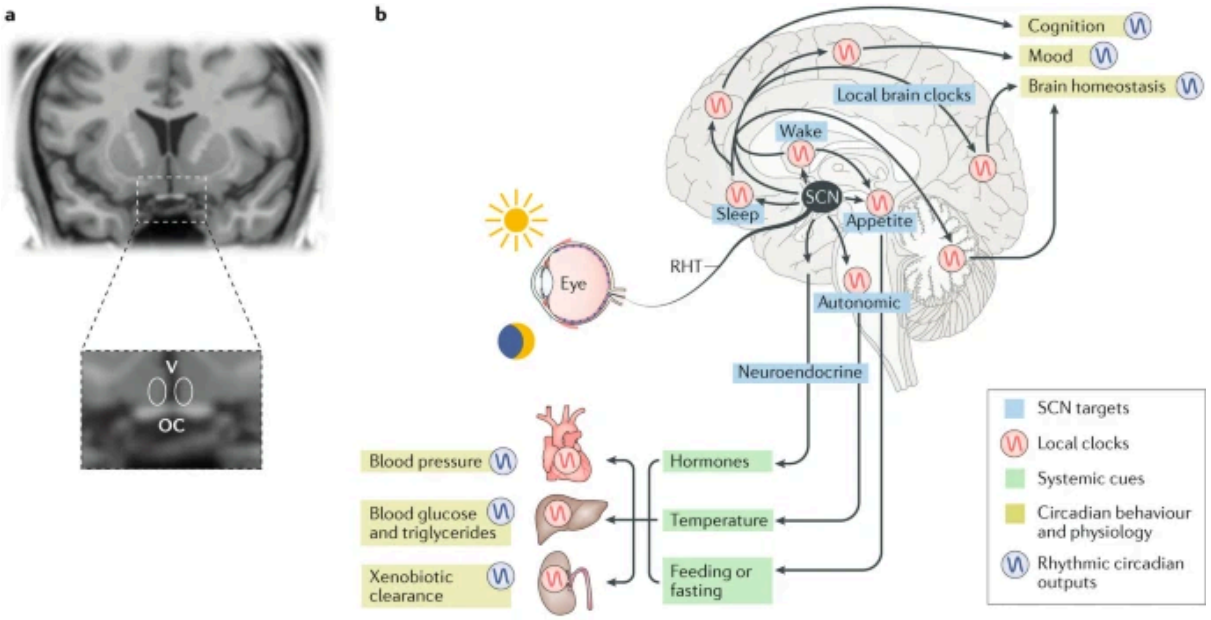


Figure 1. The suprachiasmatic nucleus (SCN). a) A coronal MRI scan of a human brain with the SCN shown in the magnified cutout. b) Functions and projections of the SCN.

From Hastings, M.H., Maywood, E.S. & Brancaccio, M. Generation of circadian rhythms in the suprachiasmatic nucleus. *Nat Rev Neurosci* 19, 453–469 (2018). https://doi.org/10.1038/s41583-018-0026-z

Light in the morning advances the circadian rhythm, while light in the evening delays it. Light in the morning has been found to be generally more effective against SAD than light in the evening. The phase-shifting hypothesis explains this by saying that the missing morning light in winter delays the circadian rhythm, which causes depressive symptoms in some people. A small minority of SAD patients might benefit from evening light instead of morning light, because their circadian rhythm is advanced by the missing evening light in winter (Huang et al., 2024).

Huang et al. (2024) argues that additional mechanisms might be at play, because: 1) light therapy seems to improve non-seasonal depression too, 2) higher illuminance is necessary for the antidepressant effect (5000 lux of white light) than for the phase-shifting effect (120 lux of white light), 3) phase shifting does not consistently predict improvement. Other sources report antidepressant effects for less than 5000 lux of white light, but still

substantially higher than 120 lux (such as 2500 lux) (J. L. Anderson et al., 2009). Huang et al. (2024) identified only two human neuroimaging studies on light therapy. They are not with SAD patients but nevertheless valuable for understanding how light therapy might affect mood. Fisher et al. (2014) investigated the effect of light therapy in the brain's threat circuit in healthy individuals. They found that three weeks of morning light therapy reduced activity in the PFC and the amygdala while viewing pictures of fearful and angry faces. In addition, light therapy strengthened the connection between these two areas and between different areas within the PFC. The latter was in part moderated by the 5-HTTLPR (serotonin-transporter-linked promoter region) genotype. The results of this study suggest that the antidepressant effect of light therapy is due to effects on serotonin and the threat circuit of the brain.

Ma et al. (2020) investigated the effects of two weeks of morning light therapy on the salience circuit of the brain in individuals with subclinical sleep disturbances. The salience circuit consists mainly of the insula, anterior cingulate cortex, and thalamus. This study found that light therapy improved sleep and reduced the functional connectivity of the right anterior insula. Other studies (M. C. Chen et al., 2014; Nofzinger et al., 2004; Wang et al., 2016) have found that when the insula is hyperactive, people experience negative emotions before sleep and have difficulty falling asleep. Taken together, these studies suggest that the insula might play a role in light therapy.

Other studies have investigated the non-visual effects of light on the brain in other contexts. Non-visual effects, e.g. on alertness, outlast visual effects and can therefore be differentiated. They found that non-visual effects of light included an increased activation in the occipito-parietal network, decreased activation in the hypothalamus (Perrin et al., 2004), and increased dynamic responses in the posterior thalamus and subcortical regions supporting attentional effects (Vandewalle et al., 2006). Blue light (450 to 480 nm) is considered the most physiologically active wavelength. A lower intensity of blue light compared to white light is necessary to induce shifts in the circadian rhythm. Melanopsin, a

photopigment in the retina responsible for the circadian effects of light, is most sensitive to blue light. This can also be seen in the brain: compared with green light exposure, blue light exposure led to changes in the left intraparietal sulcus, supramarginal gyrus, right insula, left middle frontal gyrus, and left thalamus (Vandewalle et al., 2007).

1.3.2 Creatine supplementation for cognition

Creatine has a crucial role in providing cells with energy by maintaining the level of adenosine triphosphate (ATP), which is the energy currency of the cell (Figure 2). Creatine is synthesized from methionine, arginine, and glycine in the brain, liver, kidneys, and pancreas (Rawson & Venezia, 2011). It is present in meat and to a small extent in dairy products and can also be taken as a supplement synthesised out of sarcosine and cyanamide in the laboratory (Smith & Tan, 2006).

Due to their high and fluctuating energy demands, creatine is particularly important for muscle and brain cells. It is well known that creatine supplementation increases muscle creatine and improves muscle performance. It is much less understood to what extent brain creatine and brain performance are affected by creatine supplementation. Creatine can enter the brain through the blood-brain barrier via a membrane protein called the creatine transporter. The creatine transporter is relatively rare in the blood-brain barrier and is further downregulated by incoming creatine, which might limit creatine from entering the brain, especially in the long term (Candow et al., 2023).

A clear case where supplemented creatine nevertheless passes the blood-brain barrier to a relevant extent and drastically improves brain performance is in individuals with cerebral creatine deficiency syndromes. Cerebral creatine deficiency syndromes are rare and come in three forms: 1) AGAT enzyme deficiency, 2) GAMT enzyme deficiency, 3) creatine transporter deficiency. The first two affect creatine synthesis and the third affects, as the name says, creatine transport. All three cause intellectual disability. When found early in life,

the first two are successfully treated with creatine supplementation (J. F. Clark & Cecil, 2015). Unfortunately, there is no reliable treatment for the last one yet but treatments are being studied (J. Li & Xu, 2023).

Studies measuring the effect of creatine supplementation on human brain creatine levels have been few, small, and heterogeneous. Overall, they seem to indicate that supplementing creatine can be successful at elevating brain creatine levels, but to a lesser extent than muscle creatine levels. In the following, I will focus on studies in healthy adults, as they are the population in Study 2. The main cognitive tasks in Study 2 were the Backwards Digit Span, which is a verbal working memory task, and Raven's Advanced Progressive Matrices, which is an abstract reasoning task. The brain areas of interest for the Backwards Digit Span include the auditory phonological loop (inferior parietal lobule and inferior frontal operculum), the salience network (anterior insula and dorsal anterior cingulate cortex), and the central executive network, which includes the lateral prefrontal cortex (R. Li et al., 2012; Papagno et al., 2017). Abstract reasoning tasks have been linked to the multiple demand network (insula, the cerebellum, and broad fronto-parietal areas) and the re-evaluation network (frontal pole and frontal orbital cortex, superior and middle frontal gyri, angular gyrus, bilateral occipital regions, and cerebellum) (Zurrin et al., 2024). The available studies described below only partly cover these brain areas but they nevertheless help answer the more general question of if and to what extent supplemented creatine reaches the brain in healthy adults.

Dechent et al. (1999) tested the effect of creatine supplementation on brain creatine in the thalamus, the cerebellum, grey matter, and white matter "by means of quantitative localized proton magnetic resonance spectroscopy in vivo (2.0 T, stimulated echo acquisition mode sequence; repetition time = 6,000 ms, echo time = 20 ms, middle interval = 10 ms, automated spectral evaluation)". They found a trend for a single 20g dose of creatine in 6 healthy young adults increasing brain total creatine by 7.7% in the thalamus, 3.1% in grey matter, and 3.1% in white matter. After 28 days of this dose, brain total creatine increased by

14.6% (p = .01) in the thalamus, 5.4% (p = .03) in the cerebellum, .4.7% (p = .05) in grey matter, and 11.5% (p = .02) in white matter. Lyoo et al. (2003) found that creatine supplementation significantly increased brain creatine (Cr) and marginally increased brain phosphocreatine by 3.9% in 15 healthy men (10 in creatine group, 5 in placebo group). In this study, creatine resonance was measured using proton 1H-MRS and phosphocreatine was measured by phosphorus 31P-MRS. No effect on brain creatine was found in a study by Wilkinson et al. (2006) with 20g/day creatine supplementation for 5 days in 18 healthy men (12 in creatine group, 6 in placebo group) using single-voxel proton magnetic resonance spectroscopy of the deep frontal cerebral white matter. Pan and Takahashi (2007) used high-field MR (31)P and (1)H spectroscopic imaging to study the effect of 20g creatine supplementation for 7 days on various measures of cerebral energetics in 12 healthy adults and concluded that “[creatine] supplementation appears to improve high-energy phosphate turnover in healthy brain and can result in either a decrease or an increase in high-energy phosphate concentrations”. Turner et al. (2015) used MRI to investigate the effect of 20g creatine supplementation for 7 days on brain creatine in 15 healthy adults in a placebo-controlled cross-over design with a 5-week washout period. They found a creatine increase of 5.9% in the brain area studied, a mixed sensorimotor grey and white matter voxel. Solis et al. (2017) studied the effect of supplementing 0.3g/day creatine per kg bodyweight (e.g. 21g for a 70kg individual) for 7 days on brain creatine. This study used phosphorus magnetic resonance spectroscopy (31P-MRS) and had different populations: 14 vegetarian adults, 17 omnivore adults, 18 elderly individuals and 15 children. Creatine supplementation did not significantly change brain phosphocreatine in any of these groups (-0.7% to +3.9%). This study also tested changes in muscle phosphocreatine with the same method and found it significantly increased after creatine supplementation.

In addition to the aforementioned studies on healthy adults, there have also been a few studies on the effect of creatine supplementation on brain creatine in different populations. A study in healthy children found no effect (Merege-Filho et al., 2017) in the studied areas (left

dorsolateral prefrontal cortex, left hippocampus, and occipital lobe). In sleep-deprived participants (Gordji-Nejad et al., 2024) and in several populations of patients with depression, creatine supplementation has been found to significantly increase brain creatine levels (Hellem et al., 2015; Kondo et al., 2011, 2016). This is consistent with other studies that suggest that impaired bioenergetics might be a factor in sleep-deprivation (Harper et al., 2013; Weigend et al., 2019) and in depression (Gardner & Boles, 2011; Klinedinst & Regenold, 2015). The mechanisms and potential effects of creatine monohydrate on measures of brain function are summarised in Figure 2.

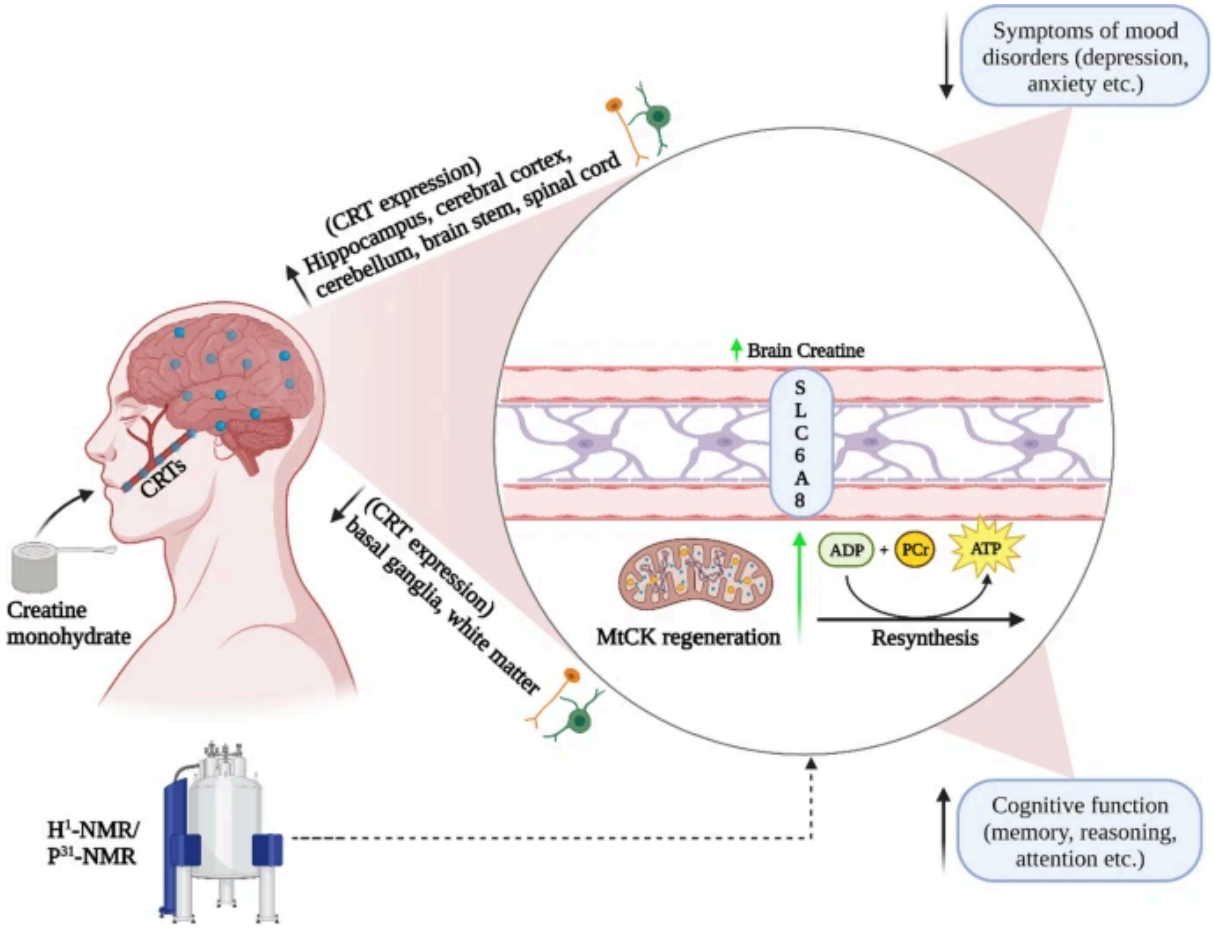


Figure 2. Mechanisms and potential effects of creatine monohydrate on measures of brain function. ADP, adenosine diphosphate; ATP, adenosine triphosphate; CRT, creatine transporter, MtCK, mitochondrial creatine kinase; NMR, nuclear magnetic resonance; PCr, phosphocreatine.

From Candow, D. G., Forbes, S. C., Ostojic, S. M., Konstantinos Prokopidis, Stock, M. S., Harmon, K. K., & Faulkner, P. (2023). Correction to: “Heads Up” for Creatine Supplementation and its Potential Applications for Brain Health and Function. *Sports Medicine, 54*(1), 235–236. https://doi.org/10.1007/s40279-023-01888-z

1.3.3. Psychotherapeutic exercises for anxiety

Anxiety disorders are characterised by “excessive fear and anxiety and related behavioral disturbances” (American Psychiatric Association, 2013). Examples for anxiety disorders are panic disorder, generalised anxiety disorder, and social anxiety disorder. Panic disorder is characterised by sudden, repeated episodes of intense fear called panic attacks, often accompanied by physical symptoms such as heart palpitations, sweating, and shortness of breath. Generalised anxiety disorder is marked by persistent and excessive worry about multiple areas of life, causing restlessness, difficulty concentrating, and sleep disturbances. Social anxiety disorder involves an overwhelming fear of social situations due to concerns about being judged or scrutinised by others, which can lead to avoidance of these situations. Cognitive behavioural therapy (CBT) helps individuals identify and challenge unhelpful thought patterns, learn coping strategies, and gradually face anxiety-provoking situations, thereby reducing symptoms and improving daily functioning (Hofmann, 2011). Mindfulness-based therapies focus on cultivating a non-judgemental awareness of the present moment, helping individuals observe anxious thoughts and emotions without being controlled by them (Williams & Penman, 2011).

The brain’s fear circuit includes limbic structures such as the amygdala, hippocampus, striatum, anterior cingulate cortex, and insula. A review (Brooks & Stein, 2015) concluded that the fear circuit is overactive in anxiety disorders and that cognitive behavioural therapy (CBT) improves symptoms of anxiety by improving the control the prefrontal cortex (PFC) has on this circuit.

Two studies investigated the neural effects of CBT in panic disorder. Kircher et al. (2013) found that CBT improved panic symptoms by reducing the activity in the left inferior frontal gyrus in the PFC and increasing the connectivity between regions of the fear circuit (amygdalae, insulae, anterior cingulate cortex) and the inferior frontal gyrus. Lueken et al.

(2013) found that after CBT, the amygdala, hippocampus, and right pregenual anterior cingulate cortex were less active when presented with a safety signal. The greatest symptom improvements were observed when the connectivity was increased between the amygdala and the anterior cingulate cortex.

Several studies suggest that CBT improves generalised anxiety disorder by activating the PFC and improving its control over the amygdala, insula and hippocampus when processing emotional stimuli (Ball et al., 2014; Fonzo et al., 2014; Maslowsky et al., 2010; McClure et al., 2007). Similarly, several studies suggest that CBT improves social anxiety by activating the occipito-temporal PFC, which improves control of the limbic circuit (Doehrmann et al., 2013; P. R. Goldin et al., 2013, 2014; Klumpp et al., 2013). In two studies, Mansson et al. found that internet-delivered CBT improved social anxiety by reducing activation of the amygdala and increasing connectivity between the amygdala, the anterior cingulate cortex, and the PFC (Månsson et al., 2013, 2015).

Studies on the neural effects of mindfulness-based therapies have found many overlapping neural circuits between mindfulness and CBT. Goldin et al. (2021) prompted participants to react, accept, or reappraise negative self-beliefs before and after 12 weeks of group CBT, a group-based mindfulness intervention called mindfulness-based stress reduction, or a waitlist control. Reacting means an unguided response, accepting is a mechanism commonly associated with mindfulness, while reappraising is a mechanism commonly associated with CBT. Interestingly, Goldin et al. (2021) found that CBT and mindfulness reduced symptoms of social anxiety via the same brain circuits related to reappraisal and acceptance. Both CBT and mindfulness increased activation in the dorsomedial prefrontal cortex (DMPFC) and dorsal anterior cingulate cortex (dACC) when reacting. CBT and mindfulness increased activation in DMPFC, dACC, left dorsolateral prefrontal cortex (DLPFC), left ventrolateral prefrontal cortex (VLPFC), left supramarginal gyrus, left posterior superior temporal gyrus, and thalamus during reappraisal. CBT and mindfulness increased activation in the left DLPFC, DMPFC, dACC, and thalamus during acceptance. Results by

Hölzel et al. (2013) indicated that mindfulness improved generalised anxiety disorder by decreasing amygdala activation, increasing ventrolateral PFC activation and increasing amygdala-PFC functional connectivity. Goldin et al. (2013) found that mindfulness-based therapy was more effective than aerobic exercise at improving social anxiety. The improvement was associated with increased activation of the prefrontal cortex and attention-related parietal cortical regions.

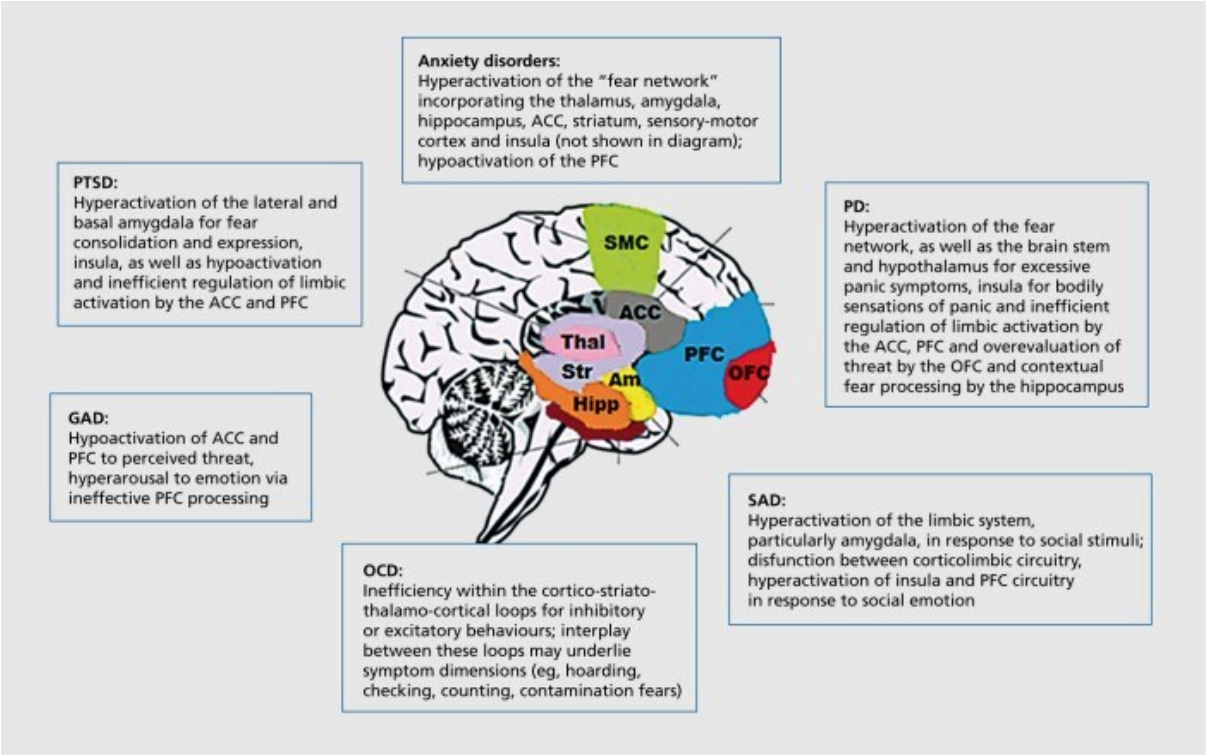


Figure 3. Neural correlates of anxiety and related disorders.

From Brooks, S. J., & Stein, D. J. (2015). A systematic review of the neural bases of psychotherapy for anxiety and related disorders. *Anxiety, 17*(3), 261–279. https://doi.org/10.31887/dcns.2015.17.3/sbrooks

1.4 Motivations

1.4.1 Motivation for Study 1

The standard of care for seasonal affective disorder (SAD) is light therapy with a standard light therapy lamp providing 10,000 lux for half an hour in the morning or alternatively, 5000

lux for one hour or 2500 lux for two hours (I. M. Anderson et al., 2008; Bauer et al., 2013; Gelenberg et al., 2010; Kurlansik & Ibay, 2012; Mårtensson et al., 2015; Praschak-Rieder & Willeit, 2003; Ravindran et al., 2009; Terman & Terman, 2005).

Treatment with a standard light therapy lamp has been found to reduce the symptoms of SAD in many studies. However, many patients do not experience complete remission (Golden et al., 2005; Mårtensson et al., 2015; Pjrek et al., 2020). Patients feel better during the spring and summer months, so we hypothesise that the fact that they do not fully remit with the standard treatment is due to ways in which it differs from summer conditions. The standard treatment differs from summer conditions in various ways. During summer, bright light exposure is longer, many hours instead of half an hour. Summer light is broad and often covers the whole visual field, while standard light therapy lamps only cover a smaller part. While standard light therapy lamps provide broadly similar illuminance to summer conditions if the instructed distance to the lamp is kept, in practice, treatments with the most common commercially available light therapy lamps likely provide less than 10,000 lux, because to receive 10,000 lux, patients' eyes have to be very close (e.g. 10-20 cm) to the lamp, which severely restricts activities that are possible during treatment and as a consequence might not be followed by patients (D. H. W. Li et al., 2010; Matour et al., 2017).

Light rooms in clinics are a different form of light therapy. These rooms are evenly and brightly lit. Studies on light rooms have found positive neurological and psychological effects in SAD patients and other patient populations (Canazei et al., 2017; Kripke et al., 1983, 1992; Rastad et al., 2008, 2011, 2017; Stain-Malmgren et al., 1998; Thalén et al., 1995; Van Someren et al., 1997; Wirz-Justice et al., 1999). However, visiting a clinic light room is not practical for many patients. The clinic light rooms reported in the literature provided patients with 1100 to 4300 lux for a maximum of 3 hours. The set-up of the new treatment tested in Study 1 was based on interviews with SAD patients who have experimented with lighting solutions similar to clinic-based light rooms but in their own home (e.g. Chapman (2023)). These patients reported finding their experimental set-ups more effective than a standard

light therapy lamp. To our knowledge, this is the first study to include both a light room condition and a standard light therapy lamp condition, and the first study to test a home-based light room. The reason why home-based light rooms have not been studied before might be that until fairly recently, they would have been very energy-intensive and expensive. Today, improvements in LED technology have overcome these barriers.

The aim of Study 1 was to determine the feasibility, acceptance, and first estimates of the potential effect size of a treatment we called Bright, whole-ROom, All-Day (BROAD) light therapy. As the name implies, this treatment consisted of brightly illuminating a whole room in patients' homes for a large portion of the day (6 hours). There were two versions of BROAD light therapy, with warm white light and with cold white light. The control condition was the standard light therapy lamp treatment for 30 minutes in the morning.

1.4.2 Motivation for Study 2

Cognition is essential for functioning and it is therefore valuable to improve cognitive performance (see 1.1). A supplement that has gained attention in recent years as possibly being helpful in this regard is creatine. It is cheap and has a good safety profile (Bender et al., 2008; de Souza e Silva et al., 2019; Kreider et al., 2017; Kutz & Gunter, 2003). Creatine is a well-established, well-studied supplement used for over 30 years by athletes to improve strength (Branch, 2003; Butts et al., 2018). Creatine improves strength by facilitating the recycling of ATP and thus giving muscle cells access to more energy (see 1.3 for more details).

Creatine is crucial to the functioning of the brain. This can most clearly be seen in individuals with a rare genetic condition called creatine deficiency syndrome, in which the synthesis or transport of creatine in the brain is impaired. This causes intellectual disability which can be prevented and reversed with creatine supplementation (J. F. Clark & Cecil, 2015). In individuals with creatine deficiency syndrome, supplemented creatine is clearly able to pass

the blood-brain barrier. Studies on the effect of creatine supplementation on healthy individuals have been more conflicting, but overall it seems that also for these individuals, supplemented creatine can reach the brain (Lyoo et al., 2003; Merege-Filho et al., 2017; Solis et al., 2017; Turner et al., 2015; Wilkinson et al., 2006). For more information, see 1.3.

In contrast to the effect of creatine supplementation on strength, the literature on creatine supplementation on cognitive performance is not well developed yet. A relatively low number of small and heterogeneous studies have studied the effect of creatine on cognitive performance, with conflicting results. They use very different cognitive tasks and study different populations, like healthy young adults, older adults, men, women, children, people with depression or other conditions, or sleep-deprived people. Some found large effects, some small effects, some found no effect. In their systematic review of the literature, albeit of only six studies, Avgerinos et al. (2018) concluded that creatine supplementation might improve short-term memory and “intelligence/reasoning”.

Larger replications of methodologically strong and promising studies are needed. One of these studies was by Rae et al. (2003). Rae et al. (2003) gave 45 participants 5g of creatine for six weeks and 5g of placebo for six weeks in a cross-over design, i.e. each participant received both supplements, in random order. Cognitive performance was measured by the Backward Digit Span (BDS), which is a test of working memory, and Raven’s Advanced Progressive Matrices (RAPM), which is a test of abstract reasoning. Some consider it a measure of general intelligence *g*, although this has been debated (Gignac, 2015; Raven, 2008). Rae et al. (2003) found highly significant and large effects of creatine supplementation on cognition: a 1.5 digits longer Backwards Digit Span and four more matrices solved in RAPM, corresponding to about one standard deviation in their study (see calculations in appendix).

The participants in Rae et al. (2003) were all vegetarian. There are reasons why individuals on a vegetarian diet might benefit more from creatine supplementation than individuals on an

omnivore diet. As mentioned above, creatine is found primarily in meat, so a vegetarian diet contains almost no creatine. Studies have found that muscle creatine is lower in vegetarians than in omnivores (Solis et al., 2017). One study compared the effect of creatine supplementation on memory performance in vegetarian and omnivore participants and found an effect only for vegetarian participants (Benton & Donohoe, 2011).

However, this is just one exploratory result of one study, so it only provides very preliminary evidence. Furthermore, the amount of creatine consumed in a typical omnivore diet is much lower than the typical supplemented amount of creatine. Typically, 5 g of creatine per day is supplemented, which would correspond to 1 kg of meat (Brosnan & Brosnan, 2016). Studies on brain creatine levels found no difference between vegetarians and omnivores (Solis et al., 2017). It might be that a higher amount of creatine than consumed through diet is needed to affect the brain, possibly because of the blood-brain barrier.

The aim of Study 2 was to replicate Rae et al. (2003) and to expand their study by adding 1) a group of omnivore participants to analyse the effect of diet and 2) eight exploratory cognitive tasks on task switching, attention, verbal fluency, and memory.

1.4.3 Motivation for Study 3

Anxiety disorders and subclinical anxiety are common and significantly impair quality of life (see 1.1). A number of psychotherapeutic approaches have been found to be effective at treating anxiety (Y.-F. Chen et al., 2017; D. A. Clark, 2013; Conrad & Roth, 2007; Kyeong et al., 2017; Maddock & Blair, 2023). Mindfulness is the act of paying attention to the present moment, noticing thoughts, feelings, and surroundings without judging them or trying to change them. Cognitive restructuring, by contrast, tackles unhelpful thought patterns by guiding individuals to challenge their thoughts and replace unhelpful thoughts with more helpful ones. Other techniques, such as gratitude practice, guided imagery to a happy place, and positive expressive writing, focus patients' attention on positive thoughts. Relaxation

techniques, such as progressive muscle relaxation or diaphragmatic breathing, focus on reducing physical tension.

Psychotherapy is inaccessible to many individuals for many reasons, such as cost and availability of therapists (World Health Organization, 2022) (see 1.2). Smartphones have become common even in low-income countries and there is almost no cost of providing app-based psychotherapeutic exercises to more people once an app is developed. This makes app-based psychotherapeutic exercises highly scalable and accessible.

Studies on app-based psychotherapeutic exercises and digital psychotherapeutic interventions more generally have emerged recently, with promising preliminary results (Firth, Torous, Nicholas, Carney, Pratap, et al., 2017; Firth, Torous, Nicholas, Carney, Rosenbaum, et al., 2017).

To our knowledge, no studies have tested a large number (twelve) of app-based psychotherapeutic exercises under standardised conditions, i.e. on the same platform, in the same setting, with the same duration and population, thus enabling a direct comparison of the exercises. This comparison is interesting to better understand the emerging treatment with app-based psychotherapeutic exercises, but also to understand how the effects of psychotherapeutic exercises compare to each other more generally, as a comparison of this nature to our knowledge has also not been done for psychotherapeutic exercises delivered in other ways and the comparison might generalise to other forms of delivery. Integrating psychotherapeutic exercises across different schools of psychotherapy has become increasingly popular, therefore it might be of interest to therapists of different schools to know more about the effects of exercises of their own and other schools.

The aim of Study 3 was to test the immediate effects of twelve psychotherapeutic exercises based on cognitive behavioural therapy, mindfulness, and acceptance and commitment therapy, on anxiety. Immediate effects might be relevant for e.g. engagement, compliance, breaking negative feedback loops, and during crises. The techniques tested included several

forms of mindfulness, several forms of cognitive restructuring, progressive muscle relaxation, diaphragmatic breathing, positive expressive writing, gratitude practice, and guided imagery. We compared the exercises together to two different controls (measurement-only and, in an exploratory fashion, reading control). In an exploratory fashion, we compared the effects of the exercises to each other.

1.5 RCTs, CONSORT, Open Science, and frequentist, Bayesian, and robust statistics

Randomised controlled trials (RCTs), Consolidated Standards of Reporting Trials (CONSORT), Open Science, and different statistical approaches (frequentist, Bayesian and robust) are all methods that help science be unbiased, transparent, interpretable and reproducible. These methods are present in all three studies of this dissertation and thus constitute a methodological red thread.

An RCT is, as the name implies, an experiment where participants are *randomly* assigned to a treatment group or a *control* group. Random assignment prevents self-selection and, if the sample is large enough, makes it unlikely that the two groups differ substantially by chance. The control group experiences measurements and time passing between them just like the treatment group. This ensures that the difference measured before and after the intervention is not simply due to passing time or due to the measurements themselves. In addition to time and measurements, a control group can have other things in common with the treatment group, such as effort, attention, and superficial similarities between the control and treatment interventions. A placebo control is designed to elicit similar treatment expectations as the treatment, so that we can rule out that the difference between the two groups is due to differences in expectation (placebo effect). To minimise effects from staff or participant expectation, to the extent that this is possible, staff and participants should be blind about which group a participant is in. This is easier in some cases, such as when giving

participants different supplements that look the same, and more difficult in other cases, such as with psychotherapy. For psychotherapy in particular, there is debate about what constitutes a good control intervention (Gaab, 2023).

The CONSORT guideline is a checklist that guides researchers in how to report RCTs in a manner that is comprehensive, accurate, and transparent. This allows others to better understand the quality and nature of the study and also to replicate it. The CONSORT guideline was first published in 1996 and last revised in 2010. It encourages, among other things, reporting in detail how the experiment was recruited for and performed, including how exactly randomisation and blinding were done, and in which ways control and treatment interventions were (dis)similar. The main CONSORT guideline is for parallel two-group RCTs. Extensions of this guideline have been created for other forms of RCTs and for specific topics. CONSORT is complemented by guidelines for other forms of studies, such as for observational studies (STROBE, Strengthening the Reporting of Observational Studies in Epidemiology) and systematic reviews and meta-analyses (PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses).

Open Science principles aim to make science more transparent, accessible, and reproducible (Center for Open Science, 2025). They advocate for preregistration, registered reports, open data, open code, and open access papers. Preregistration means publishing the research plan and hypotheses before data collection starts. This prevents so-called HARKing: Hypothesising After Results are Known (Shrout & Rodgers, 2018). Preregistering encourages researchers to be transparent about which aspects of their research were exploratory (i.e. had no prior hypothesis), which were confirmatory (i.e. had a hypothesis), and to what extent the data match the hypotheses. Retrospectively changing a hypothesis to match the data leads to the mistaken impression that the hypothesis is better supported than it actually is and contributes to the replication crisis (i.e. the problem that many findings cannot be successfully replicated) (Shrout & Rodgers, 2018). This misreporting does not necessarily have to be intentional. A researcher might have had many hypotheses and

misremember that they always gave more weight to the hypothesis which now seems most plausible in light of the new evidence. A registered report is similar to a preregistration but takes it one step further: The research plan and hypotheses are not only published but also peer-reviewed and the paper is accepted by a journal, before collecting data (Nosek & Lakens, 2014). This has two advantages: 1) the methods can be improved by the peer review before it is too late, and 2) the journal decides whether to accept or reject a paper based on the quality of the methods and relevance of the question, rather than what the results happen to be. This counteracts another component of the replication crisis; fluke positive results getting published while valuable null results are not published. Open data means the data are available to the public so that analyses can be verified (Center for Open Science, 2025). Open code, i.e. making the code used in the statistical analysis publicly available, also contributes to this goal (Center for Open Science, 2025). Finally, open access to research papers means that the public has access to the results and details of the study without having to pay (or be part of an institution that pays) for this access (Center for Open Science, 2025). To publish open access, the researcher or their institution usually has to pay an open access fee. This means that publishing open access is not affordable to everyone. Publishers have started to respond to this by lowering or waiving their open access fees for researchers in low-income countries (Saloojee & Pettifor, 2024).

Frequentist statistics, also called NHST (Null Hypothesis Significance Testing) answers the question “If the null hypothesis is true, and we run the same experiment a large number of times: How often would we see a difference in the data that is as large or larger than the difference that we found?”. This probability is the p-value (Field, 2018). In behavioural sciences such as psychology, a p-value below 0.05 is conventionally used to classify a result as statistically significant (Field, 2018). This number originated in 1925 from Fisher, who used it as a convenient heuristic but recommended treating it as a vague threshold, not a cut-off, and not using it blindly to reject hypotheses (Field, 2018). The p-value is closely related to the concept of confidence intervals (CIs). By convention, a 95% CI is used, e.g., a

95% CI of the mean of each group. The 95% CI is constructed so that, if the same experiment is conducted a large number of times, the population mean is within this interval in 95% of these experiments (Field, 2018). A common misconception is that a difference between two groups is statistically significant at $p < .05$ only if their 95% CIs do not overlap. For this heuristic, the 83.4% CI would be more appropriate (Knol et al., 2011). $P = .05$ corresponds to a substantial overlap of the 95% CIs (of about ¼ of the length of the intervals) (Field, 2018). Because a p-value only describes the statistical significance of a result and not its practical significance, a measure of effect size needs to be reported. While there are commonly used heuristics for what constitutes a small/medium/large effect, what effect size is practically significant needs to be determined in context.

Standard frequentist statistics assume a normal distribution and are not reliable if the data deviate too much from it, i.e. if there are significant outliers (Field, 2018). An early solution to this problem was nonparametric statistics, which do not assume normality. Nonparametric statistics convert data into ranks. This approach neutralises outliers at the cost of losing information on how large the difference between two data points is. Wilcox (2011) and Field (2018) argue that the more modern approach of robust statistics is superior to nonparametric statistics.

Robust methods retain the magnitude information present in the original scores, rather than collapsing everything into ranks (Wilcox, 2011; Wilcox & Rousselet, 2023). Instead of using the mean, robust statistics uses robust estimators such as trimmed means or M-estimators. A trimmed mean is computed by discarding a certain percentage of the smallest and largest values before taking the average of the remaining data. An M-estimator starts with a special “loss” function (instead of the usual squared deviations) that grows more slowly for large outliers. One then iteratively adjusts the estimate so that the sum of these loss values is minimized. Robust approaches can include bootstrapping (i.e. resampling with replacement) to generate confidence intervals and p-values that do not depend on assumptions of normality or homoscedasticity. Unlike fully nonparametric approaches, these methods can

accommodate unequal variances and skewed shapes without discarding the relative distances among data points. This often gives robust techniques both higher power (i.e., fewer missed effects) and more reliable control of Type I error rates. In contrast to nonparametric statistics, robust statistics often need to be done with a computer.

Bayesian statistics answers the questions “How likely is my hypothesis to be true?” and “How should I update my beliefs about different hypotheses based on new data?”. In contrast to frequentist statistics, which only specifies the null hypothesis, Bayesian statistics needs one or more specified alternative hypotheses. To answer the question “How likely is my hypothesis to be true?” using Bayesian statistics, one specifies one’s original belief about the hypothesis with a so-called prior distribution. When new data are collected, it is compared to the null hypothesis and the alternative hypothesis. Depending on how likely the new data are under each hypothesis, one updates one’s belief about which hypothesis is more likely to be true. The ratio of support the new data give to the null vs the alternative hypothesis is called the Bayes factor. A Bayes factor of 1 means the data are equally likely under each hypothesis. A Bayes factor lower than 1 means the data are more likely under the null hypothesis; a Bayes factor greater than 1 means the data are more likely under the alternative hypothesis. For example, a Bayes factor of 3 means the data are 3 times more likely under the alternative hypothesis than under the null hypothesis. The prior odds (ratio of the two prior distributions) are multiplied by the Bayes factor to arrive at the posterior odds (ratio of the two posterior distributions). The posterior distribution indicates the updated belief about a hypothesis (Field, 2018).

1.6 Aims of this dissertation

The aim of this dissertation is to investigate interventions for mental health (in particular, seasonal affective disorder and anxiety) and cognition. The studies I performed to accomplish this have in common that they were randomised, controlled, and double-blind or partially blind where full double-blind was not possible. Studies 1 and 3 had a parallel design

while Study 2 had a cross-over design. I analysed the data using descriptive statistics and exploratory standard frequentist statistics (Study 1, which was a feasibility study), Bayesian statistics (Study 2), and standard and robust frequentist statistics (Studies 2 and 3). Following Open Science principles, all studies were preregistered and the data and papers are freely available to the public.

Specifically, the research question of Study 1 was: Is broad, whole room, all day light therapy at home a feasible intervention against seasonal affective disorder? The research question of Study 2 was: Does supplementing creatine enhance cognition in healthy individuals? For Study 3, it was: To what extent do different app-based psychotherapeutic exercises reduce state anxiety?

2 METHODS

2.1 Study design

The studies this thesis is based on are randomised, controlled, and double-blind or partially blind where full double-blind was not possible (see the section on blinding). Studies 1 and 3 had a parallel design while Study 2 had a cross-over design. In Study 1, which aimed to find a more effective version of light therapy against seasonal affective disorder than the standard of care, the experimental group received what we called BROAD light therapy. The control group received the standard of care (a standard light box treatment for 30 minutes in the morning). In Study 2, which tested the cognitive effects of creatine, the experimental group received creatine and the control group received maltodextrin. In Study 3, which tested the immediate effects of app-based psychotherapeutic exercises on anxiety, each experimental group received a different psychotherapeutic exercise. There were two control groups: one group received the instruction to do what they would usually do and the other group read a text on anxiety. The papers adhered to the CONSORT reporting guidelines.

2.2 Participants

Participants in Study 1 were adults with a diagnosis of “major depressive disorder (recurrent episode) with a winter seasonal pattern”, also known as winter depression, diagnosed by psychologists in a video interview. Participants in Study 2 were mostly healthy adults. Participants in Study 3 were adults with a high level of state anxiety.

2.3 Interventions

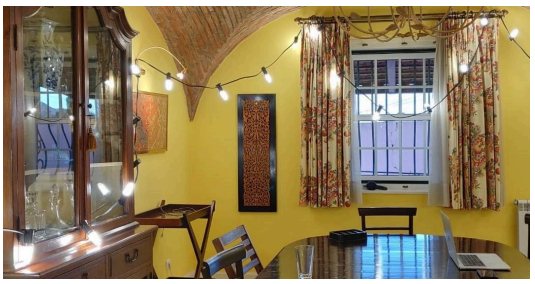
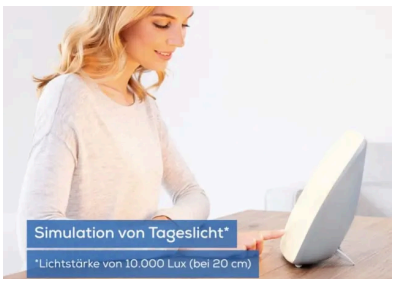
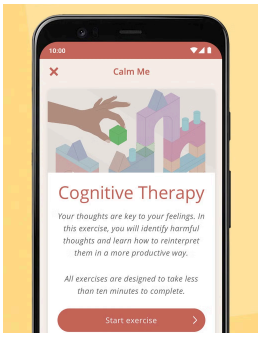
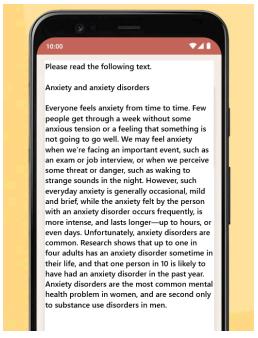
Study	Treatment(s) of interest	Control(s)
1	BROAD light therapy, 6h/day, 4 weeks 	Standard light therapy lamp, 0.5h/day, 4 weeks  From https://www.beurer.com/de/p/60811/
2	Creatine supplementation, 5g/day, 6 weeks 	Maltodextrin supplementation, 5g/day, 6 weeks 
3	One of twelve app-based psychotherapeutic exercises, 7 minutes  From: https://anxiety-relief-by-mind-ease.en.softonic.com	Measurement-only control or reading control (text about anxiety), 7 minutes  Adapted from: https://anxiety-relief-by-mind-ease.en.softonic.com

Figure 4. Interventions in the three studies.

2.3.1 Study 1

Study 1 included a new treatment that we called Bright, whole-ROom, All-Day (BROAD) light therapy and, as an active control, a standard light therapy lamp. Both treatments were instructed to be performed five days a week for four weeks. The instructed duration for BROAD light therapy was 6 hours per day and for the control group, 30 minutes per day.

BROAD light therapy consisted of publicly available, bright, white LED bulbs and socket cords. Participants installed the socket cords and LED bulbs with hooks in their own home. The installation was checked via photos and suggestions for correction were made where necessary. There were two versions of BROAD light therapy: the cold white version and the warm white version. The cold white version consisted of 40 bulbs with 1800 lumens each (72,000 lumens in total), with a colour rendering index of 80-90 and a colour temperature of 6000 K, while the warm white version consisted of 70 bulbs of 1460 lumens each (102,240 lumens in total), with a colour rendering index of 80-90 and a colour temperature of 4000 K. The warm white version had more lumens to ensure the two versions had the same amount of blue light. Blue light is thought to be the physiologically active component of light. Participants measured the lux at eye level at their most typical position in the room with a lux meter app on their smartphones. In the cold light condition, the average illuminance at eye level while seated was 1433 lux (ranging from 550 to 4061 lux). In the warm light condition, it was 1829 lux (ranging from 500 to 6800 lux). Paper lanterns to diffuse the light were optional and rarely chosen by participants. The two versions were tested to explore whether participants would react differently to them. For example, participants might prefer a warmer colour temperature.

The standard light therapy lamp was the TL 41 by Beurer (Ulm, Germany) or similar lamps, based on availability. It delivered 10,000 lux at a distance of 20 cm, with a rapid drop in illuminance beyond that range. Its colour rendering index was 80 and its colour temperature

was 6500 K. Participants were instructed to keep a distance of 20 cm to the lamp and only briefly and occasionally look directly into it.

The similarity between the intervention groups consisted of receiving a plausible and active treatment, as well as the same reminders, questionnaires, and set-up checks using photos.

2.3.2 Study 2

Study 2 included two supplements in powder form: creatine monohydrate and a placebo supplement. Five grams per day of each supplement was consumed for six weeks.

The creatine monohydrate was by the company Alzchem (Trostberg, Germany). The placebo supplement was maltodextrin by the company Nutricia (Frankfurt am Main, Germany). After testing 40 participants, we realised that differences in solubility were masked better when the supplements were mixed into yoghurt or food of comparable consistency. From this point, participants were therefore asked to mix the supplements into such food.

The two supplements were similar in that they were both white, tasteless powders. Creatine was less soluble in water than maltodextrin. The containers for the two supplements were identical for each participant, except for a label specifying the order in which to take them. At the end of their last testing, participants were asked to guess their supplement order.

2.3.3 Study 3

Study 3 included twelve psychotherapeutic exercises of the Mind Ease app as treatment conditions, as well as a measurement-only control and a reading control.

The app's exercises drew primarily from third-wave cognitive behavioural therapy (CBT), encompassing classic CBT, mindfulness, and acceptance and commitment therapy (ACT) (Hayes & Hofmann, 2017). The exercises "Silver lining" and "cognitive distortion list"

employed the standard CBT technique of cognitive restructuring (Hofmann, 2011). "Anxiety-excitement reappraisal", an unusual cognitive restructuring method, was inspired by a method in the book “DARE” (McDonagh, 2015) and integrated exposure and emotional reappraisal from classic CBT, and briefly, committed action from ACT (Hayes et al., 2016; Hofmann, 2011). The exercises "positive expressive writing" and "gratitude practice" were based on positive psychology (Seligman et al., 2005). "Guided imagery" employed the relaxation technique of positive guided imagery, used across several therapies including CBT (Hofmann, 2011). "Diaphragmatic breathing" and "progressive muscle relaxation" constituted two physiology-oriented relaxation exercises (Conrad & Roth, 2007; Zaccaro et al., 2018). The largely mindfulness-based exercises were "mindful breathing", "body scan", "Leaves on a Stream", and "dropping anchor" (Hayes et al., 2016; Kabat-Zinn, 2009). See the appendix for detailed descriptions of each exercise.

The measurement-only control group underwent the same measurements as the treatment groups but was told to continue with their normal activities for 7 minutes in between the measurements until a bell rang. The reading control group underwent the same measurements as the treatment groups. In between the two measurements, this group read an informative text on anxiety (Rector et al., 2005). The reading control was more similar to the treatment groups, in that it addressed the topic of anxiety and included reading. Control and treatment interventions were expected to be of a similar duration of about 7 minutes. The actual durations were slightly higher: M = 9.8 minutes (SD = 4.1) for the exercises, M = 12.6 minutes (SD = 3.5) for the measurement-only control, and M = 8.9 minutes (SD = 4.1) for the reading control. Before they took part, participants knew the study focused on online methods to reduce negative feelings, but they were not told about the use of control groups. Thus, those in the measurement-only group probably realised they were in a control group. Meanwhile, the reading control group was less clearly a control group, as they were asked to read a lengthy text on anxiety (see appendix), which might have seemed like a psychoeducational activity. Conducting the study entirely online and automatically, with no

staff present, likely lowered the pressure to please the experimenter. Participants were aware that they would be paid regardless of their answers, so there was no financial motive to misrepresent themselves.

Originally, there was an additional group that was assigned to a psychotherapeutic exercise not randomly but based on a machine learning algorithm used by the Mind Ease app. However, the algorithm was copied incorrectly from the app to the study website, so the data from this group (N = 82) were excluded.

2.4 Outcomes

2.4.1 Study 1

In Study 1, we focused on assessing the feasibility and acceptance of the new treatment by evaluating participant feedback, side effects, recruitment, and adherence.

The preregistered primary outcome was the level of symptom severity after four weeks of treatment, as measured by the Hamilton Depression Rating Scale-Seasonal Affective Disorders 29-items Version, self-report version (SIGH-SAD-SR, or SIGH-SAD for short) (Terman, Williams, White, Gould, 2008). We used this outcome in Study 1 for preliminary descriptive insights into the effectiveness of the new treatment.

The secondary outcomes were symptom severity after two weeks as measured by the SIGH-SAD, and the percent of remitted participants after two and four weeks.

2.4.2 Study 2

In Study 2, the preregistered confirmatory outcomes were the Backwards Digit Span (BDS), which assessed working memory, and standardised 10-minute versions of Raven's Advanced Progressive Matrices (RAPM), which assessed abstract reasoning (Rae et al., 2003; Wechsler, 1955). For both tasks, the score was the sum of correct answers.

The BDS test consisted of giving participants a series of digits, which they had to remember and say in reverse order. The test started with a series of two digits and became increasingly longer. There were always two series of the same length. The test was stopped when participants failed twice at the same length.

The RAPM tests were the same standardised 10-minute versions as used by Rae et al. (2003). At each timepoint, a different version was administered, verified to be equally difficult by Rae et al. (2003). Each test consisted of 20 increasingly difficult tasks, which participants had a total of 10 minutes to complete. Each task consisted of pictures arranged in a 3x3 matrix, with one picture missing. Participants had to choose the right picture to complete the matrix out of a selection of 8 pictures.

The full list of cognitive tests, including exploratory tests, was:

- The Auditory Verbal Learning Test (AVLT, in German: VLMT, Lux et al. (2001)), part 1: immediate verbal recall
- The Brief-Visuospatial-Memory Test – Revised (BVMT-R), a test of visuospatial memory (Benedict et al., 1996)
- The D2 Test of Attention (Brickenkamp, 2002), a test of sustained attention
- The Trail-Making-Test A (TMT-A), a test of visual attention (Reitan, 1958)
- The Trail-Making-Test B (TMT-B), a test of task switching (Reitan, 1958)
- The Stroop test (in German: Farb-Wort-Interferenz Test, a test of inhibitory control (Bäumler & Stroop, 1985)
- Forward Digit Span
- Backward Digit Span
- The Block-Tapping-Test, a test of visuospatial working memory (Schellig, 1997)
- Raven's Advanced Progressive Matrices (RAPM)
- The Auditory Verbal Learning Test (AVLT, in German: VLMT, Lux et al. (2001)), part 2: delayed verbal recall and recognition

- Regensburger Wortflüssigkeitstest, a test of verbal fluency (Aschenbrenner et al., 2000)

The tasks were administered in the order shown above. In addition, at baseline, the crystallized verbal intelligence test MWT-B (Mehrfachwahl-Wortschatz-Intelligenztest, Lehrl, (2005)) and a demographic questionnaire were administered at the end of the testing. After completing each six weeks of supplementation participants were asked to report side effects in a free text box. At all three timepoints, participants filled out a questionnaire on their daily form, including how many hours they had slept the night before, when they had last consumed caffeine that day, etc.

2.4.3 Study 3

In Study 3, the preregistered confirmatory outcome was the average of the responses to three sliders native to the Mind Ease app, with which participants were asked to indicate before and after the intervention how they felt at that moment (“How I’m feeling right now”). The sliders ranged from “very tense” to “very relaxed” (in equal intervals: “very tense”, “quite tense”, “somewhat tense”, neither tense nor relaxed”, “somewhat relaxed”, “quite relaxed”, “very relaxed”), and respectively for “very worried” to “very calm”, and “very bad” to “very good”. “Neither…nor” corresponded to the midpoint of the scales, with each of the other labels covering one sixth of the scales. Only the label corresponding to the current position of the continuous slider was visible. The most positive possible answer was scored as 0, i.e. no anxiety, and the most negative possible answer was scored as 100, i.e. high anxiety—these numbers were not visible to participants.

We chose the slider questions because they were quick to complete, already included in the app, and helped us capture real user experiences as closely as possible. In a separate, pre-registered feedback survey (N = 40), participants reported high levels of satisfaction with clarity, ease of use, and content validity (see appendix). The psychometric properties of this scale were determined in a pre-study (N = 199). Like the main study, the participants in the

psychometric pre-study and the feedback survey were adults in the US recruited on the platform Positly (details on methods and participants are in the appendix).

The first goal of the psychometric pre-study was to determine the internal consistency of the scale as measured by Cronbach's α. Cronbach's α was .94 in the pre-study, with inter-item correlations between .81 and .90. In the main study, Cronbach's α was .70 pre-intervention and .90 post-intervention. Such high internal consistency is anticipated when measuring related aspects of state anxiety, and it aligns with findings from other state anxiety scales. (Guillén-Riquelme & Buela-Casal, 2011; Marteau & Bekker, 1992).

The second goal of the psychometric pre-study was to determine the convergent validity of the three-slider scale with the state subscale of the State-Trait Anxiety Inventory (STAI). The STAI state subscale has 20 self-report items (e.g. “I feel at ease”), each rated on a 4-point Likert scale from “not at all” to “very much so”. The three-slider scale showed a strong, highly significant correlation with the STAI state subscale, $r(197) = .872$, $p < .001$, when all data were included. After excluding one obvious outlier based on the scatterplot (STAI state subscale score of 76 and three-slider scale score of 27), the correlation remained very similar, $r(196) = .893$, $p < .001$. Linear regression without the outlier produced the equation STAI state subscale score = three-slider scale score * 0.55 + 20.66. We used this formula to present the main study findings in STAI state subscale terms, making the results more familiar to readers. Details on these results are in the appendix.

2.5 Randomisation and blinding

In Study 1, participants were randomly assigned using Excel to one of the three treatment groups (standard light therapy lamp, warm BROAD light therapy, cold BROAD light therapy) in three strata based on level of symptom severity (SIGH-SAD score above 29, between 29 and 20, or below 20) and in a randomly varying block size of 4 or 8 participants. I sent the participant code and baseline SIGH-SAD score to the senior author (Dr. Jan Brauner) via

email and he replied with the assigned group. The senior author had no direct contact with participants, nor any knowledge about them apart from the two pieces of information that I provided. There was an equal chance of being assigned to the BROAD light therapy group or to the control group (standard light therapy lamp). Within the BROAD light therapy group, there was an equal chance of being assigned to the warm white light group or to the cold white light group. Participants were blind to the extent that this was possible. The treatments were visibly different, but participants were not told if they were in a control group or not. All participants received the same standardised reminders to fill out the survey after two and four weeks of treatment. Participants filled out the SIGH-SAD-SR remotely by themselves with no staff interaction.

In Study 2, participant codes were simply randomly assigned with equal chance to one of the two supplement orders (creatine and then placebo or placebo and then creatine) by the pharmacy of the University Hospital Heidelberg using Excel. The pharmacy provided the testers with the supplements, each container labelled with the participant code and “A” for the first supplement and “B” for the second supplement. Testers gave or sent the supplement to the participants. This order was unknown to participants and the study staff until the end of participants’ participation and revealed using the SNOSE method (sequentially numbered, sealed envelopes).

In Study 3, participants were simply randomly assigned to one of the two control groups or one of the twelve treatment groups automatically using GuidedTrack. The control groups were designed to each be 1.5 times the size of a treatment group. Participants were not told whether they were in a treatment or a control group. To make the measurement-only control group less likely to think that they were a control group, all groups were told that the study was about measuring short-term changes in mood. Treatment expectation was likely lower for the control groups than the treatment groups (see 4.2 limitations). There was no interaction between staff and participants.

2.6 Statistical analysis

2.6.1 Study 1

Study 1 was a feasibility study. Feasibility was assessed qualitatively by evaluating success at recruiting, participants setting up BROAD light therapy, descriptive symptom reduction, participant feedback, and side effects. The study was not powered for statistical between-group comparisons. In line with the CONSORT guideline on pilot and feasibility trials (Eldridge et al., 2016), the findings were reported primarily in a narrative and descriptive fashion. Continuous outcomes were reported as mean and standard deviation and categorical outcomes were reported as count and percentage. In an exploratory manner, we also give descriptive results of the improvement in SAD symptoms after four weeks adjusted for confounders (see appendix for details).

An exploratory Pearson correlation test was performed with the variables $\log_2$(lux) at eye level and symptom reduction at four weeks. The log of lux was used instead of lux because the scatterplot (lux, symptom reduction) showed a log-like curve. This is consistent with psychophysical principles (e.g., Weber–Fechner law) that describe a logarithmic rather than linear relationship between stimulus intensity (in this case, brightness) and perceived or physiological response (Bhatia, 2001). Further exploratory Pearson correlation tests were performed with measures of adherence and symptom reduction at four weeks.

2.6.2 Study 2

Study 2 involved preregistered hypotheses about two cognitive tasks: the Backward Digit Span and Raven's Advanced Progressive Matrices. The analyses regarding these two tasks were therefore confirmatory in nature. The other eight cognitive tasks were exploratory.

We had hypothesised that vegetarian/vegan participants would benefit more from creatine supplementation than omnivore participants. We tested this with a 2x2x2 mixed measures

ANOVA with diet (vegetarian/vegan vs omnivore) as a between-subjects independent variable, supplement (creatine vs placebo) and supplement order (creatine first or placebo first) as within-subjects independent variables, and cognitive performance as the dependent variable. Because including diet as a factor made no relevant difference to the results, we merged the two diet groups for all other analyses.

For each task, a 2x2 mixed measures ANOVA was performed with supplement (creatine vs placebo) and supplement order (creatine first or placebo first) as independent variables and cognitive performance as the dependent variable. We had preregistered using a t-test for this analysis, however, this was a mistake. The t-test is only valid if the two supplement order groups are of exactly equal size. The criterion was not fulfilled, so the right test was the mixed measures ANOVA. The Greenhouse-Geisser correction was used in all these ANOVAs, but it did not alter any of the values.

We checked how robust the results of the ANOVAs mentioned above were to outliers and assumptions of normality by running different robust versions of these ANOVAs with 5% and 20% winsorising, 20% trimming, and bootstrapping combined with 20% trimming (using the WRS2 R package, sppi functions).

In addition, we performed Bayesian analyses to assess the relative likelihood that creatine has a small effect and that it has a large effect like in Rae et al. (2003), compared to the null hypothesis that it has no effect. This analysis was performed on the estimated marginal means (EMMs) of the creatine scores and placebo scores. We operationalised the null and alternative hypotheses using two different methods: 1) point models, in which the probability mass centers on the effect sizes of interest and forms Cauchy distributions around it, and 2) half-normal models centred on zero, which treat the effect sizes of interest as overestimates of the real effect size. Half-normal models give a higher probability to effect sizes that are smaller, and give a low probability to effect sizes twice as large as the effect size of interest.

Half-normal models are commonly used when replicating studies. See the appendix for more details on the Bayesian analyses.

Separate exploratory analyses were performed for participants with a high and low baseline performance, and for the first supplementation and second supplementation (see appendix).

We included participant data irrespective of how well a participant had adhered to the supplement regime. One participant took the supplements in the wrong order. This participant was analysed with the real, not the prescribed, order.

2.6.3 Study 3

Study 3 involved preregistered hypotheses about the effect of each psychotherapeutic exercise separately compared to the measurement-only control and about the effect of all exercises combined compared to the measurement-only control. In an exploratory fashion, we performed the same analyses with the reading control instead of the measurement-only control. Also in an exploratory fashion, we compared the effects of the exercises to each other.

As preregistered, these comparisons were performed with a 2x2 mixed ANOVA with group as the between-subjects independent variable, time (pre vs post) as the within-subjects independent variable, and level of anxiety as the dependent variable. The Greenhouse-Geisser correction was used in all these ANOVAs, but it did not alter any of the values. In addition, Cohen's d was calculated for an independent samples t-test with group as the independent variable and change in anxiety (pre minus post) as the dependent variable. This t-test is equivalent to the interaction term in the 2x2 mixed ANOVA. Reporting Cohen's d provides a more comprehensive picture and makes the effect easier to interpret. Furthermore, we analysed the same comparisons using linear regression. This approach is not equivalent and provides a different perspective on the same questions.

We checked how robust the results of the ANOVAs mentioned above were to outliers and assumptions of normality by running different robust versions of these ANOVAs with 5% and 20% winsorising, 20% trimming, and bootstrapping combined with 20% trimming (using the WRS2 R package, sppi functions).

We excluded participants based on low baseline anxiety (as preregistered, those with a three-slider scale score below 50, corresponding to a STAI state subscale score of 48). In addition, we excluded participants if they took more than 30 minutes or less than 3 minutes. This criterion was not preregistered, but it makes sense because less than 3 minutes is not enough time to complete the exercise adequately. Spending more than 30 minutes (one participant spent over 7000 minutes) meant participants likely interrupted the study with other activities and even if they did not, longer durations increase noise too much, as state anxiety changes over time. The analyses without this time criterion produced very similar results and are found in the appendix. No participant received an intervention different from the assigned one.

3 STUDY SYNOPSES

Table 1. Overview of studies included in the present dissertation.

No.	Reference	Open Science
1	**Sandkühler, J. F.**, Brochhagen, S., Rohde, P., Muscheidt, R. C., Grömer, T. W., Müller, H., & Brauner, J. (2022). 100,000 lumens to treat seasonal affective disorder: A proof of concept RCT of Bright, whole-ROom, All-Day (BROAD) light therapy. ***Depression and Anxiety.*** https://doi.org/10.1002/da.23281	Preregistered, open data, open materials, open access
2	**Sandkühler, J. F.**, Kersting, X., Faust, A., Königs, E. K., Altman, G., Ettinger, U., Lux, S., Philipsen, A., Müller, H., & Brauner, J. (2023). The effect of creatine supplementation on cognitive performance—a randomised controlled study. ***BMC Medicine***. https://doi.org/10.1186/s12916-023-03146-5	Preregistered, open data, open code, open materials, open access
3	**Sandkühler, J. F.***, Kahl, F.*, Sadurska, M. Z., Brietbart, P., Greenberg, S., Brauner, J. (2025). The Immediate Impact of App-Based Psychotherapeutic Exercises on Anxiety: An RCT. ***Depression and Anxiety.*** https://doi.org/10.1155/da/5586831	Preregistered, open data, open code, open access

3.1 Study 1: Light therapy for seasonal affective disorder

Sandkühler, J. F., Brochhagen, S., Rohde, P., Muscheidt, R. C., Grömer, T. W., Müller, H., & Brauner, J. (2022). 100,000 lumens to treat seasonal affective disorder: A proof of concept RCT of Bright, whole-ROom, All-Day (BROAD) light therapy. *Depression and Anxiety, 39*(12), 760–769. https://doi.org/10.1002/da.23281

Light therapy with a standard light therapy lamp is a first-line treatment for seasonal affective disorder (SAD). This treatment improves symptoms to some extent, but in many cases, does not lead to full remission. Some patients have experimented with lighting solutions and reported benefiting more from lighting that is more powerful, broader, and has a longer duration than traditional light therapy lamps (see 1.4, motivation of Study 1).

The aim of Study 1 was to test the feasibility and obtain first estimates of the effectiveness of a treatment we called Bright, whole-ROom, All-Day (BROAD) light therapy. This treatment consisted of patients spending six hours a day, five days a week, for four weeks, in a very brightly illuminated room in their own homes. There were two versions of BROAD light therapy, with warm white light and with cold white light. The control group received the standard of care, 30 minutes in the morning every day, five days a week, for four weeks with a standard light therapy lamp. For more details on the interventions, see 2.3 Interventions, Study 1. Participants were randomly assigned to one of the two BROAD light therapy groups or to the control conditions so that the control group and the combined BROAD light therapy group were of the same size (see 2.5). Recruitment, side effects, and adherence were evaluated to determine the feasibility of the new treatment and of larger future studies. The Hamilton Depression Rating Scale-Seasonal Affective Disorders 29-items, self-report version (SIGH-SAD-SR) was used to assess the level of SAD symptoms before the start of the treatment, as well after two and four weeks of treatment.

Sixty-two patients took part in the study, none dropped out. Photos of the set-ups showed that participants were able to successfully set up BROAD light therapy in their homes as instructed, in some cases with slight corrections. The illuminance at eye level, measured in lux, was lower in the BROAD light therapy group compared to the standard light therapy lamp if the short distance to the eyes was kept for the latter as instructed. However, the fact that BROAD light therapy did not restrict participants to a certain position made a longer duration possible and led to more lux overall throughout the day. Adherence to the instructed treatment durations was high. Feedback was overall positive for all three groups. Eight participants in the BROAD light therapy group and four participants in the standard light therapy lamp group experienced side effects. These side effects included eye ache, headache, and agitation, were not severe and were known from other studies on light therapy (Botanov & Ilardi, 2013).

Because this was a feasibility trial not powered for between-group differences, only descriptive results are reported for symptom improvement, as recommended by the CONSORT guideline on pilot and feasibility trials (Eldridge et al., 2016). Both groups experienced high levels of symptom improvement (see Table 2 and Figure 5), similar to those found in previous studies (Pjrek et al., 2020). For more details on the outcome measures, see 2.4 Outcomes, Study 1.

Table 2. Secondary endpoints.

	BROAD light therapy	Standard light therapy lamp
Symptom improvement after 4 weeks	11.58 (8.39), [8.45; 14.72]	13.24 (9.64), [9.76; 16.72]
Symptom improvement after 4 weeks, adjusted (arbitrary units)	12.56 (7.66), [9.70; 15.42]	12.92 (9.19), [9.61; 16.23]
Subscale A symptom improvement after 4 weeks	7.22 (5.32), [5.23; 9.20]	8.21 (6.55), [5.85; 10.57]
Subscale B symptom improvement after 4 weeks	4.37 (4.49), [2.69; 6.04]	5.03 (4.32), [3.47; 6.59]
Symptom improvement after 2 weeks	8.92 (8.75), [5.65; 12.18]	10.8 (8.16), [7.80; 13.79]
Remission after 2 weeks	3 (10%)	3 (10%)
Remission after 4 weeks	7 (22%)	6 (20%)

Mean (standard deviation) [95% confidence intervals of the mean improvement] in SIGH-SAD points given for unadjusted symptom improvement and N (%) given for remission. The adjusted symptom improvement does not correspond to SIGH-SAD points, but only serves as a unitless measure of similarity. The adjustment accounts for differences between the groups in baseline symptom severity, timepoint of treatment, and delay between baseline assessment and start of treatment.

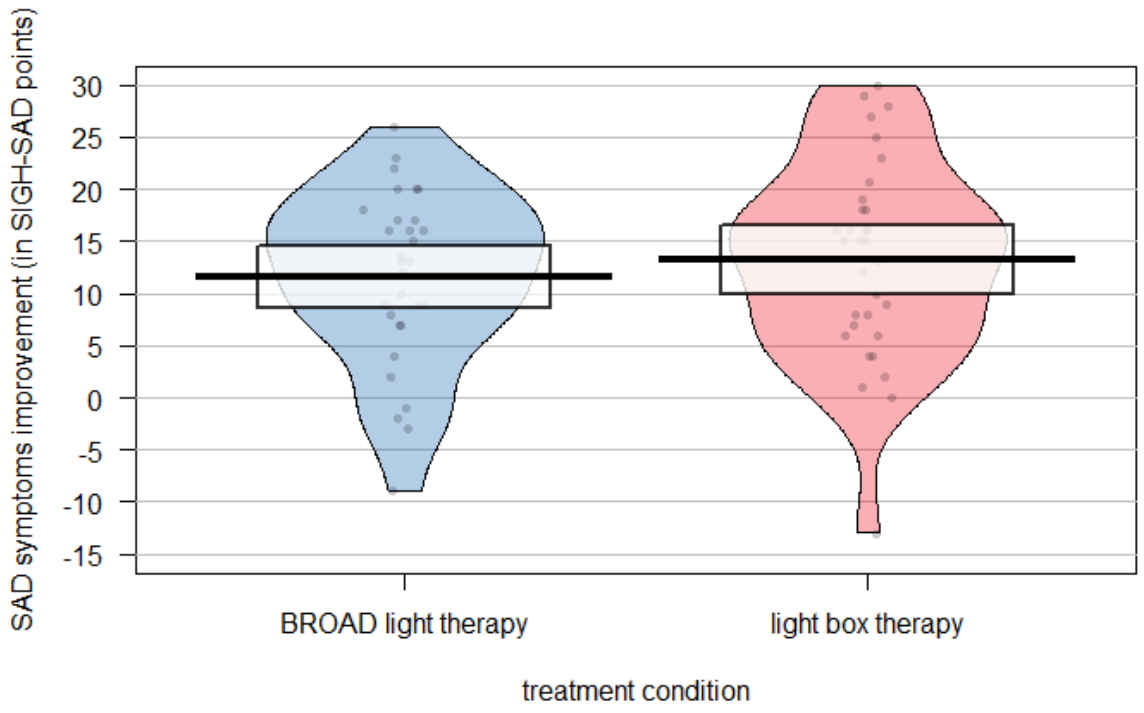


Figure 5. Key descriptive finding of Study 1. Improvement in SIGH-SAD score after four weeks of treatment. The plot shows the means, 95% confidence intervals of the means and distribution for both the standard light therapy lamp group ("light box therapy") and the BROAD light therapy group.

The BROAD light therapy group exhibited great variation in illuminance at eye level, from 500 lux to 6800 lux. This variation was due to details in how participants set up the lights and properties of the room. An exploratory correlation between symptom improvement after four weeks and the illuminance in the BROAD light therapy group found that participants who received more lux had greater symptom improvement, $r(28) = .43$, $p = .032$. The median illuminance in the BROAD light therapy group was 1400 lux. BROAD light therapy participants with a lower illuminance than this reduced their symptoms (as measured by the SIGH-SAD) by M = 8.9 (SD = 8.5), i.e. less than the control group, which reduced their symptoms by M = 13.24 (SD = 9.64). In contrast, participants with a higher illuminance than this experienced a symptom reduction of M = 16.7 (SD = 6.5), i.e. a greater reduction than the control group. Every doubling of illuminance (lux) was associated with a four SIGH-SAD points greater reduction in symptoms.

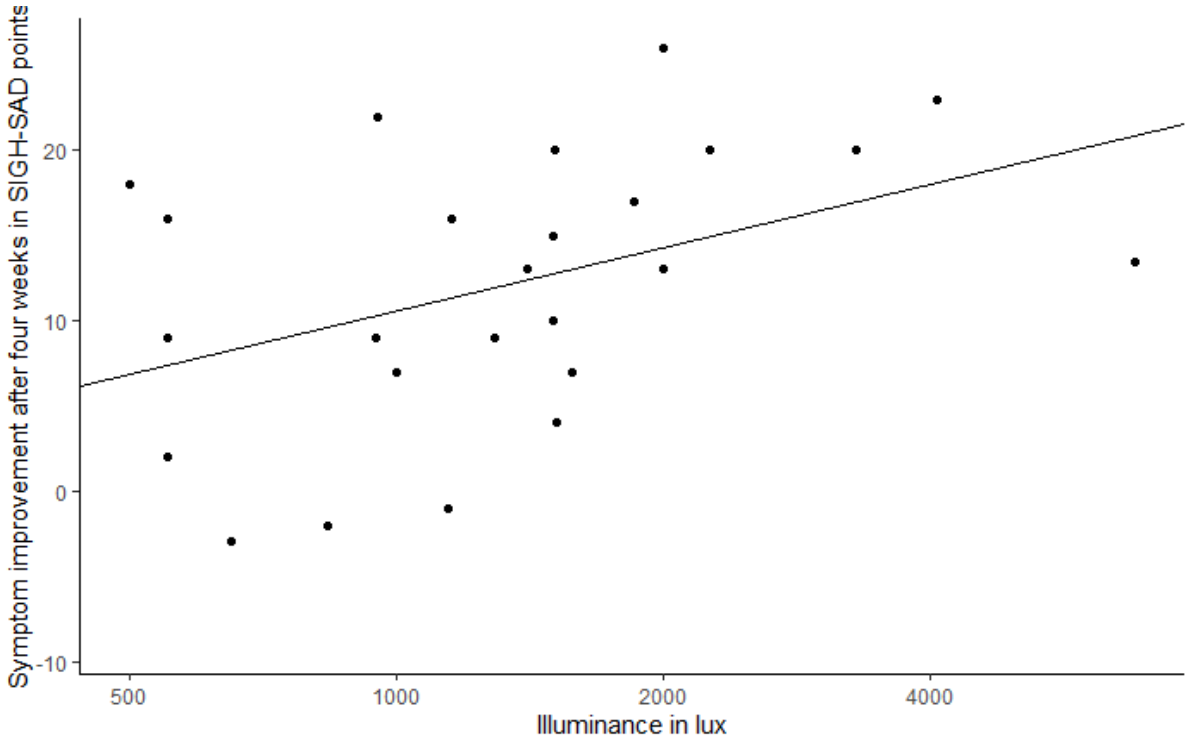


Figure 6. Key exploratory finding of Study 1. Correlation between symptom improvement after four weeks and log(lux) in the BROAD light therapy group. Participants who received more lux had better outcomes. Scatter plot with regression line.

In conclusion, Study 1 found BROAD light therapy to be feasible. Both the BROAD light therapy group and the standard light therapy lamp group experienced large improvements in SAD symptoms, in line with previous studies The exploratory correlation showed that optimising the BROAD light therapy set-up to maximise illuminance at eye level might improve how effective this treatment is.

3.2 Study 2: Creatine supplementation for cognition

Sandkühler, J. F., Kersting, X., Faust, A., Königs, E. K., Altman, G., Ettinger, U., Lux, S., Philipsen, A., Müller, H., & Brauner, J. (2023). The effect of creatine supplementation on cognitive performance—a randomised controlled study. *BMC Medicine*. https://doi.org/10.1186/s12916-023-03146-5

Study 2 is the largest study to date examining how creatine supplementation affects cognitive performance. Creatine is a substance that aids in the recycling of adenosine triphosphate (ATP), a key energy source in both muscle and brain tissue. It is a safe, widely studied supplement for strength training. Previous research has indicated that taking creatine may raise its levels in the brain and possibly enhance cognitive function. However, findings on cognitive outcomes have been inconsistent, which could be linked to differences in study populations, supplementation methods, and the types of cognitive tests employed. Rae et al. (2003) found a large effect (greater than one standard deviation) of creatine supplementation on cognitive performance for the Backward Digit Span (BDS) and Raven's Advanced Progressive Matrices (RAPM).

The aim of Study 2 was to replicate Rae et al. (2003) and to expand their study by adding a group of omnivore participants to analyse the effect of diet and by adding eight exploratory cognitive tasks on inhibitory control, reaction time, memory, reasoning, task switching, and attention. Like Rae et al. (2003), our study was a cross-over study that was randomised, placebo-controlled, and double-blind, with 5g of each supplement being taken every day for 6 weeks. The study included a total of 123 participants, about half of them omnivores and half vegetarians. Participants reported side effects, mainly digestional discomfort, significantly more often for creatine supplementation compared to placebo supplementation, $p = .002$, RR = 4.25 (see 4.2., Limitations).

The results of Study 2 were inconclusive. The mixed ANOVA (and equivalent independent samples t-test) found that the effect of creatine supplementation on cognitive performance for BDS was in the expected direction and bordered on significance, $p = .067$, $\eta^2_P = .028$, $d = 0.17$. The same analysis for RAPM found that the effect of creatine supplementation was in the expected direction but further from statistical significance, $p = .327$, $\eta^2_P = .008$, $d = 0.09$.

After creatine supplementation, in the BDS test participants remembered 0.2 more digits than after placebo supplementation. Because there were two series for each length, this equated to 0.41 more points in the creatine condition (Figure 7). After creatine supplementation, in the RAPM test participants were able to correctly solve 0.23 more matrices than after placebo supplementation (Figure 7).

For BDS, the size of the effect of creatine supplementation was $\eta^2_P = .028$, which means that 2.8% of the variance that was not explained already by other factors was explained by the creatine supplementation. For RAPM, the size of the effect of creatine supplementation was $\eta^2_P = .008$, which means that 0.8 % of the variance that was not explained already by other factors was explained by the creatine supplementation. If these tests had been IQ tests, this effect size would correspond to a 2.5 IQ point increase under creatine for BDS and 1 IQ point increase under creatine for RAPM.

Robustness checks using different robust ANOVAs (5% winsorised, 20% winsorised, 20% trimmed, bootstrapped and 20% trimmed) found that the RAPM finding was robust (p-values between .354 and .672), but the finding for BDS was less so. For BDS, the 5% winsorised ANOVA had a p-value of .009, while the bootstrapped and 20% trimmed ANOVA had a p-value of .370. The other ANOVAs had p-values in between these two.

Bayesian analyses were conducted to better understand how to update one's belief about different hypotheses based on the data. According to van Doom et al. (2021), Bayes factors (BF_{10}) between ⅓ and 3 mean that there is not enough data to decisively favour one hypothesis over the other. Within this interval, a BF_{10} below 1 gives weak support for the null

hypothesis, whereas a BF_{10} above 1 weakly supports for the alternative hypothesis. Once a BF_{10} exceeds 3 or falls below ⅓, the evidence is considered moderate and when it exceeds 10 or falls under 1/10 the evidence is regarded as strong. Bayesian analyses found weak to moderate evidence for a small effect of creatine on cognitive performance in BDS, weak evidence for a small effect of creatine on cognitive performance in RAPM, and for both tasks, strong evidence against large effects like those found in Rae et al. (2003).

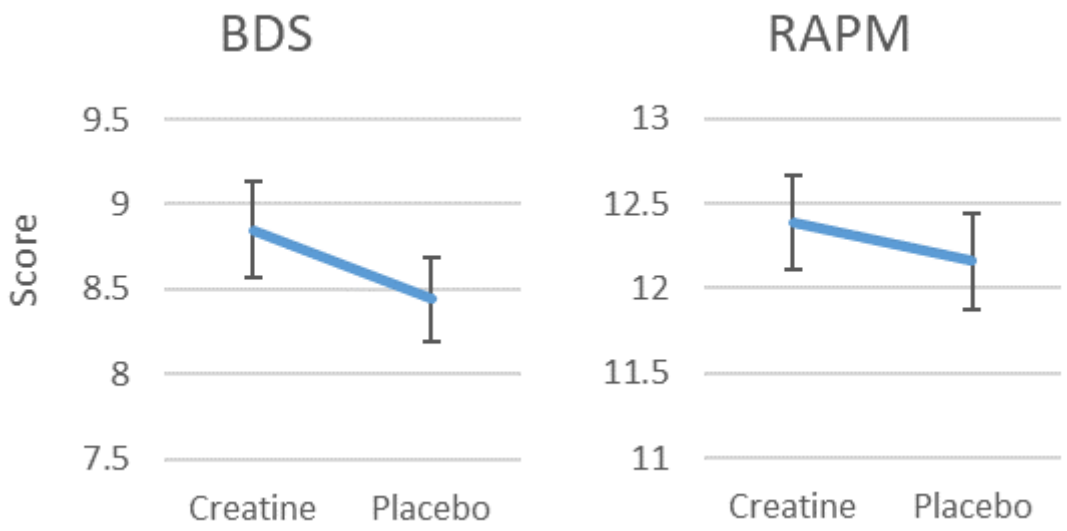


Figure 7. Key findings of Study 2. a) Estimated marginal means for the Backward Digit Span (BDS) score. b) Estimated marginal means for Raven's Advanced Progressive Matrices (RAPM) score. Error bars represent standard errors.

Vegetarian and vegan participants did not benefit more from creatine than omnivore participants. In fact, the effect was in the opposite direction, although not significant (interaction between diet, supplement, and supplement order for RAPM p = .606, for BDS p = .559; interaction between diet and supplement for RAPM p = .392, for BDS p = .808), with similar results for the robust ANOVAs. Bayesian analyses of diet as a moderator found strong support in favour of the null hypothesis over the effect size in Benton and Donohoe (2011) (d = 0.36) and strong to weak support in favour of the null hypothesis over smaller effect sizes (see appendix). The robust ANOVAs and Bayesian analyses reported above

were performed without diet as a factor, because eliminating this factor did not affect the supplement effect and simplified the analyses. Similarly, there was no effect of diet for the exploratory tasks.

Exploratory analyses found that there was no creatine effect for the 8 exploratory cognitive tasks; the findings were what would be expected by chance variation. There was no effect of age or gender for any of the confirmatory or exploratory analyses.

The arguably more promising result for BDS compared to the other cognitive tasks might indicate that creatine is more helpful for working memory than for other types of cognitive tasks. Alternatively, it might be that BDS was a harder task than the others. It is definitely harder than the Forward Digit Span, which was the same task except participants said the digits in the order presented to them instead of in reverse order. It also seems likely that BDS was harder than other short-term/working memory tasks like the AVLT, in which participants had to remember as many words as possible out of a list of 15 words, or the BVMT-R, in which participants had to remember and draw as many as possible of six simple abstract figures. It is harder to compare the difficulty of the BDS to non-memory tasks.

In conclusion, the effects for the two confirmatory cognitive tasks were in the expected direction but not statistically significant. The effect bordered significance for BDS (p = .067) but not RAPM (p = .327). The effect was robust for RAPM but less so for BDS. Bayesian analyses found weak to moderate evidence for a small creatine effect for these tasks, and strong evidence against a large effect like that found by Rae et al. (2003). Creatine was associated with side effects in some participants, which might make supplementation unadvisable or unappealing for them. Larger studies are needed to be more certain about whether or not there is a small effect of creatine on cognitive performance and how to reduce the likelihood of side effects.

3.3 Study 3: App-based psychotherapeutic exercises for anxiety

---

Sandkühler, J. F.*, Kahl, F.*, Sadurska, M. Z., Brietbart, P., Greenberg, S., Brauner, J. (2025). The Immediate Impact of App-Based Psychotherapeutic Exercises on Anxiety: An RCT. ***Depression and Anxiety.*** https://doi.org/10.1155/da/5586831

Anxiety disorders and subclinical anxiety are common and significantly impair quality of life. Psychotherapy is inaccessible to many individuals. App-based psychotherapeutic exercises are a solution that is accessible and highly scalable (see 1.4, Study 3).

The aim of Study 3 was to test the immediate effects of twelve psychotherapeutic exercises of the Mind Ease app on anxiety. Immediate effects are relevant for e.g. engagement, compliance, breaking negative feedback loops, and during crises. We compared the exercises together and separately to two different controls (measurement-only and, in an exploratory fashion, reading control). Also in an exploratory fashion, we compared the effects of the exercises to each other. Participants (N = 1092) were assigned randomly to one of the exercises or one of the control groups. Double-blinding was implemented as much as possible (see section 2.5): Participants were not told whether they were in a control group but treatment expectation was likely lower in the control groups (see section 4.2). Staff did not interact with participants. State anxiety was measured with a custom scale validated against the state subscale of the State-Trait Anxiety Inventory.

The confirmatory analysis comparing all the psychotherapeutic exercises together to the measurement-control group found a large and significant effect ($p < .001$, $\eta^2_P = .059$, d [95% CI] = 0.8 [0.6; 1.0]), as did the confirmatory analysis comparing each exercise individually to the measurement-control group (all $p = < .001$, $\eta^2_P = .06$ to $.37$, $d = 0.5$ to 1.5). Exploratory

analyses with the reading control instead of the measurement-only control were very similar (all exercises together $p < .001$, $\eta^2_P = .063$, d [95% CI] = 0.8 [0.6; 1.0], exercises individually $p = .002$ to $< .001$, $\eta^2_P = .06$ to $.37$, $d = 0.5$ to 1.5). For all exercises, statistical significance was reached for the pre-registered threshold of $p < .05$, as well for conservative Bonferroni-corrected thresholds (correcting for twelve exercises, $p < .004$ and in all but one borderline case when correcting for twelve exercises and two control groups, $p < .002$). Skew values for anxiety scores (pre and post) ranged from -0.4 to 0.4, with kurtosis between -1.1 and 1.6. Winsorised, bootstrap, and trimmed ANOVAs produced comparable findings when used as robustness checks. Exploratory analyses indicated that higher baseline anxiety was linked to a stronger treatment effect, $r(1052) = .178$, $p < .001$. For the pre and post-anxiety scores of the merged exercise groups and the merged control groups, see Figure 8. For the improvement in anxiety for each group separately, see Figure 9.

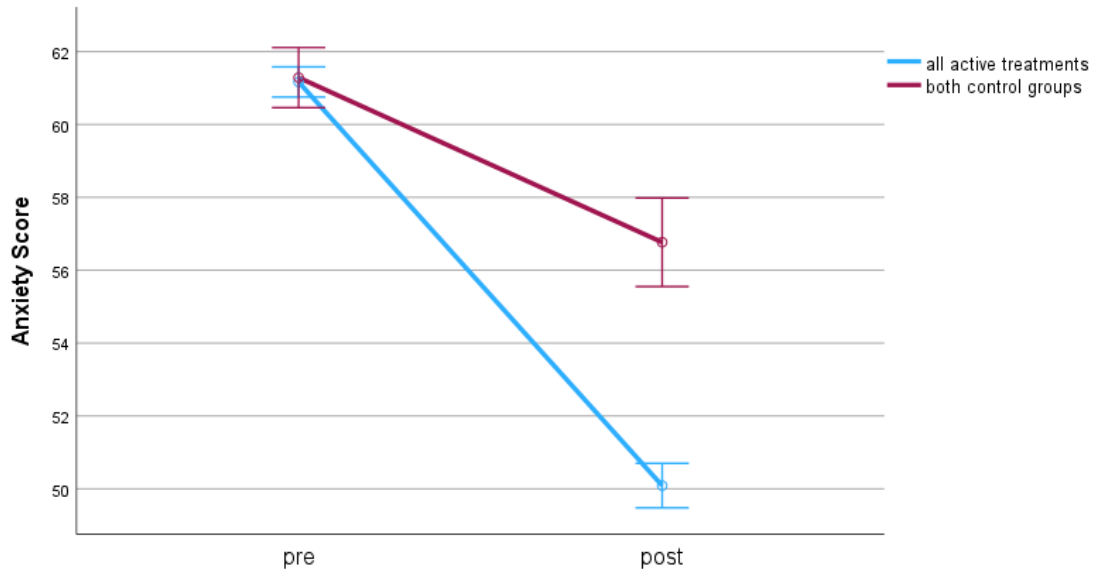


Figure 8. Key finding of Study 3 (merging groups). Pre and post-intervention means of the anxiety scores for all exercises together and the two control groups together. The anxiety score is the STAI state subscale equivalent. The error bars represent 95% confidence intervals.

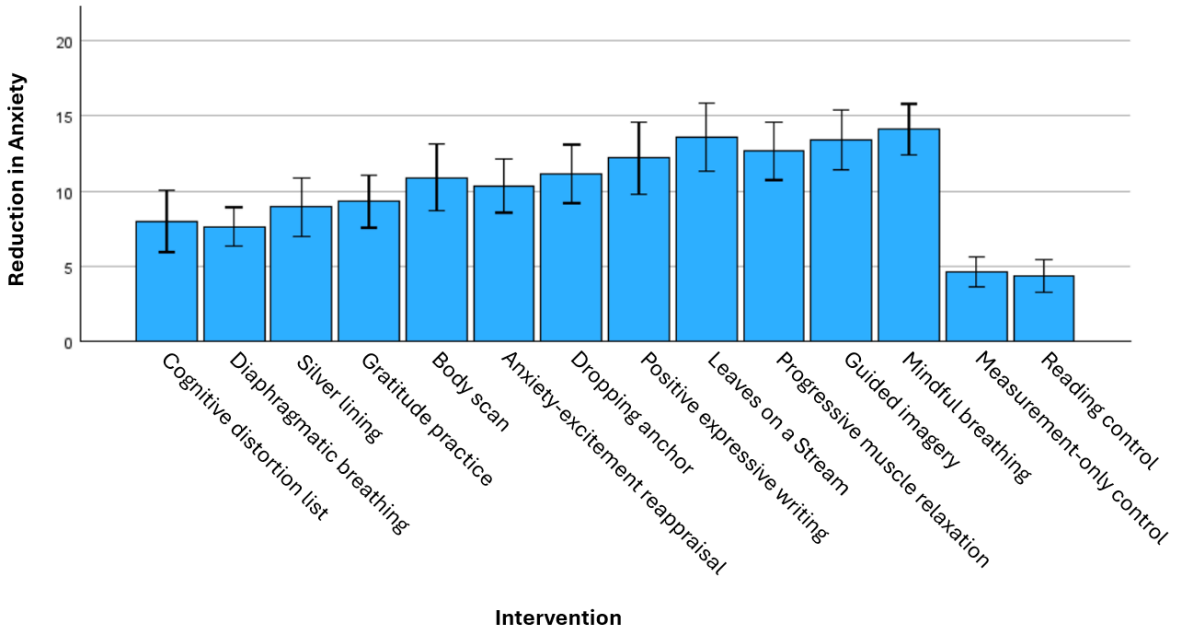


Figure 9. Key finding of Study 3 (individual treatment groups). Reduction in anxiety, given in STAI state subscale equivalent, for the twelve exercises and two controls. Error bars represent 95% confidence intervals. For scatter plots and box plots of these data, see the appendix.

Exploratory analyses found substantial differences between psychotherapeutic exercises. These differences were statistically significant at p = .05 much more often than would be expected by chance—28 out of 66 tests were significant at p = .05, while only 5%, i.e. three, would be expected to by chance. Adjusting the significance threshold for multiple testing to p = .0008 using the conservative Bonferroni method would classify nine comparisons as statistically significant.

Differences between exercises were substantial even within the same category (Table 3). For example, the mindfulness exercise “mindful breathing” was significantly more effective than the mindfulness exercise “body scan”, p = .025, d = -.39. Exercises that were primarily based on mindfulness (d [95% CI] = 0.9 [0.6; 1.2], 1.0 [0.7; 1.3], 1.2 [0.9; 1.5], and 1.5 [1.2; 1.9]) tended to be more effective than exercises that were primarily based on cognitive restructuring (d [95% CI] =0.5 [0.2; 0.8], 0.7 [0.3; 1.0], and 0.9 [0.6; 1.2]). Interestingly, the “mindful breathing” exercise was more effective than the “diaphragmatic breathing” exercise, p < .001, d = -1.00, suggesting that the mindfulness component made the exercise much

more effective. More generally, it is interesting that diaphragmatic breathing, which is a well-known relaxation exercise designed to have immediate positive effects, was less effective at reducing immediate anxiety than many exercises that are not relaxation exercises—all mindfulness exercises and even one of the cognitive restructuring exercises ("anxiety-excitement reappraisal"). Similarly, it might seem surprising that the "gratitude practice" exercise, which directs attention to life's positive aspects, was less effective at reducing immediate anxiety than mindfulness exercises, which by their nature are not about focusing on positive aspects, but just about observing (e.g. thoughts, the breath, bodily sensations) nonjudgmentally. Given that our participants started with high baseline anxiety, this likely included paying attention to thoughts and bodily sensations related to anxiety. By contrast, the positive-focused relaxation exercise "guided imagery" and the relaxation exercise "progressive muscle relaxation" were among the most effective exercises.

Table 3. Comparing exercises to each other.

	Cog. distort.	Diaphr. breath.	Silver lining	Grat. pract.	Body scan	Anx.-excit.	Drop. anchor	Expr. writ.	LoaS	PMR	Guided imag.	d	M
Cog. distort.	/											**0.5**	3.4
Diaphr. breath.	.05, .766	/										**0.6**	3.0
Silver lining	.11, .506	-.19, .269	/									**0.7**	4.3
Grat. pract.	-.17, .324	-.26, .123	-.05, .766	/								**0.8**	4.7
Body scan	-.33, .057	-.43, .013	-.23, .185	-.19, .265	/							**0.9**	6.3
Anx.-excit.	-.30, .084	-.42, .015	-.18, .283	-.14, .408	.07, .700	/						**0.9**	5.7
Drop. anchor	-.38, .027	-.50, .003	-.27, .111	-.23, .166	-.03, .877	-.10, .554	/					**1.0**	6.5
Expr. writ.	-.47, .009	-.61, .001	-.37, .037	-.34, .055	-.14, .439	-.22, .223	-.12, .501	/				**1.1**	7.5
LoaS	-.62, .0003	-.75, <.0001	-.52, .002	-.49, .003	-.28, .094	-.37, .026	-.27, .103	-.15, .393	/			**1.2**	9.0
PMR	-.57, .001	-.74, <.0001	-.46, .007	-.44, .011	-.20, .241	-.30, .083	-.19, .273	-.05, .764	.11, .522	/		**1.3**	8.0
Guided imag.	-.64, .0002	-.79, <.0001	-.53, .002	-.51, .002	-.28, .099	-.37, .024	-.27, .107	-.14, .437	.02, .894	-.09, .587	/	**1.3**	8.7
Mindful breath.	-.78, <.0001	-1.00, <.0001	-.67, <.0001	-.66, .0001	-.39, .025	-.51, .003	-.38, .012	-.23, .192	-.06, .719	-.19, .254	-.09, .586	**1.5**	9.5

Cohen's d and two-sided p-value of Welch t-test for comparison between two exercises is given (unadjusted for multiple testing). Data with the time criterion is used. Light green: $p < .05$. Full green: $p < .0008$ (adjusted significance threshold). The last two columns show the effect size d and the mean M of each exercise when compared to the measurement-only control. Abbreviations used for the names of the exercises: Cog. distort. = Identifying cognitive distortions, Diaphr. breath. = Diaphragmatic breathing, Grat. pract. = Gratitude practice, Anx. - excit. = Anxiety-excitement reappraisal, Drop. anchor = Dropping anchor, Expr. writ. = Positive expressive writing, LoaS = Leaves on a Stream, PMR = Progressive muscle relaxation, Guided imag. = Guided imagery, Mindful breath. = Mindful breathing.

To conclude, Study 3 found that all twelve app-based psychotherapeutic exercises were more effective than the control conditions at reducing immediate anxiety, with medium to large effects. Differences between exercises were substantial, including within categories. Mindfulness exercises tended to be more effective than cognitive restructuring exercises. Perhaps surprisingly, some relaxation and positive-focused exercises were less effective at reducing immediate anxiety than exercises not focused on relaxation or positivity, such as mindfulness and even cognitive restructuring exercises.

4 DISCUSSION

The studies presented in this dissertation contribute to improving mental health and cognition by evaluating three distinct interventions: Bright, whole-ROom, All-Day (BROAD) light therapy for seasonal affective disorder (SAD), creatine supplementation for cognitive enhancement, and app-based psychotherapeutic exercises for anxiety. Together, they represent a diverse approach to addressing mental health and cognitive performance through accessible interventions. This discussion highlights the implications, limitations, and potential for future research of each study, considers their combined contribution to mental health and cognitive enhancement, and discusses the similar research methods used to study the interventions.

4.1 Integration

4.1.1 Improving mental health and cognition

The results of Study 1 show that BROAD light therapy, which uses whole-room, all-day light exposure, could be a feasible alternative to the standard 30-minute light box therapy against seasonal affective disorder. While numerical symptom improvement was similar across both treatments, BROAD therapy offers a more flexible approach for patients. This is particularly relevant for those who find adherence to daily light box sessions challenging. The association between higher illuminance and greater symptom reduction in the BROAD light therapy group tentatively suggests that if BROAD light therapy is set up the right way, it could be more effective than the standard 30-minute light box therapy. However, these findings were exploratory and need to be confirmed by follow-up studies.

In Study 2, the effects of creatine supplementation on cognitive performance were modest but noteworthy, particularly for the Backward Digit Span (BDS), a verbal working memory

task. Although the effect size for creatine supplementation was small, I would argue that the potential for cumulative benefits across large populations or extended timeframes can be relevant. Creatine is already widely used for physical performance enhancement, and its safety profile makes it a viable candidate for cognitive supplementation. However, some individuals experience (non-serious and temporary) side effects, which might make creatine supplementation unadvisable or at least unappealing to these people. Furthermore, a lack of significant effects on Raven's Advanced Progressive Matrices (RAPM) and other exploratory cognitive tasks suggests that creatine's cognitive benefits may be limited to specific domains. The study did not find stronger effects for vegetarians, in contrast to our hypothesis that they would benefit more due to lower baseline creatine levels. Larger studies are needed to confirm cognitive effects of creatine, perhaps focusing on populations who are sleep-deprived, older, or cognitively impaired, where the potential benefits might be more pronounced. While the latter two are speculation, based on the premise that it is easier to improve lower performance, a recent study provided exciting evidence that even a single dose of creatine might improve cognitive performance when sleep-deprived (Gordji-Nejad et al., 2024).

Study 3 demonstrated that psychotherapeutic exercises delivered through an app can provide immediate and substantial relief from anxiety. Immediate anxiety reduction is particularly important during emotional crises and for engagement and adherence. The large effect sizes across all exercises indicate that brief, app-based interventions can be powerful tools for mental health management. These results are especially promising given the accessibility of mobile apps, which can be used at scale to address subclinical anxiety or supplement traditional therapy for clinical anxiety disorders. Effect sizes differed substantially between exercises, including between exercises of the same category. Mindfulness exercises tended to have a larger positive effect than cognitive restructuring exercises. Typical relaxation exercises were not in all cases more effective than exercises that included exposure to anxiety-related thoughts and bodily sensations. The psychotherapeutic

techniques used in the Mind Ease app—derived from cognitive behavioural therapy (CBT), mindfulness, and acceptance and commitment therapy (ACT)—are well-established in providing long-term benefits. However, the long-term effectiveness of these specific app-based exercises remains to be tested. Future trials should assess whether consistent use of the app over longer periods leads to sustained improvements in anxiety, and whether these exercises can effectively prevent the development of anxiety disorders and possibly also substance use disorders.

The studies in this thesis found effects of various sizes. Study 1 found a statistically significant exploratory correlation, indicating that every doubling of illuminance (lux) was associated with a four SIGH-SAD points greater reduction in symptoms ($r(28) = .43$, $p = .032$). This effect size might be considered moderate or large based on different heuristics suggested in the literature (Gignac & Szodorai, 2016). Both groups (BROAD light therapy and standard light therapy lamp) showed large and highly statistically significant improvements, $d = 1.4$ in each case. However, this is the difference before vs. after treatment without comparing it to a passive control group. Study 1 used the standard light therapy lamp condition as an active control. The effects of light therapy with a standard light therapy lamp are well studied, so the numerically similar effects found in the new treatment and the active control are unlikely to be due just to the passing of time. The effect size is in line with those in previous studies (Pjrek et al., 2020). By contrast, the effect sizes in Study 2 (the effects of creatine on cognitive performance) were small ($d = 0.17$ and $d = 0.09$) and more uncertain. The effects were in the expected direction for the two confirmatory cognitive tasks, but only one bordered significance ($p = .067$ and $p = .327$). There was no indication of an effect of creatine on the exploratory cognitive tasks. The small effect sizes found for the two confirmatory cognitive tasks are in stark contrast to the large effects of about $d = 1$ found by the study we were replicating and expanding (Rae et al., 2003). Study 3 found that the app-based psychotherapeutic exercises had highly statistically significant, moderate to large

effects on immediate anxiety (d = 0.5 to 1.5). These effect sizes are broadly in line with previous studies (see the appendix for a comprehensive discussion).

A number of factors might explain the similarities and differences in effect sizes, such as 1) how established the effects the studies built on were, 2) the breadth and nature of the target population, 3) the level of functioning, and 4) how ambitious the studies' aims were. Light therapy against SAD and psychotherapeutic exercises have well-established effects. The studies in this dissertation tested variations of these interventions. These interventions were new to some extent but also had much in common with interventions known to be effective. This can explain why these two studies found moderate to large effects. By contrast, the literature on creatine supplementation to improve cognition in healthy adults is not well established. There are well-established benefits of creatine supplementation, but these are regarding muscle performance, including in healthy individuals, and cognitive performance in individuals with creatine deficiency syndromes. These scenarios are too different from the scenario of Study 2—creatine supplementation to improve cognitive performance in healthy adults—to expect effects based on these previous findings. More similar scenarios to Study 2 have been tested before, but there were only a few studies, usually small, with heterogeneous methods and results. Study 2 had a much broader target population (healthy adults) than Study 1 (adults with SAD) and Study 3 (adults with high state anxiety). More targeted interventions might be more likely to have larger effects. In addition, the participants in Studies 1 and 3 started at a relatively low level of functioning regarding the outcome variables (SAD symptoms and state anxiety), while the participants in Study 2 started at a relatively high level of functioning regarding the outcome variables (cognitive performance). It might be easier to improve functioning when it is lower. In other words: Study 2 had a more ambitious aim (substantially improving cognitive performance in the general population using a little-researched method) compared to the more modest aims of Studies 1 and 3 (more incremental improvements in our knowledge of well-established interventions to improve low functioning in terms of SAD symptoms and state anxiety in targeted populations). This made

Study 2 less likely to succeed, but with more far-reaching consequences if it did than Studies 1 and 3.

The interventions studied present complementary tools for different facets of mental health and cognition. For example, depression and anxiety are highly comorbid and share many similar psychological mechanisms, such as cognitive inflexibility and unhelpful beliefs. The app-based psychotherapeutic exercises of Study 3 are based on approaches (mostly mindfulness and cognitive behavioural therapy) that are backed by many studies as being effective for anxiety *and* depression, including the recurrent type of depression that is seasonal affective disorder, as well as many other mental health conditions. This means that the population of Study 1, patients with seasonal affective disorder, would likely benefit from these app-based psychotherapeutic exercises, too. Conversely, bright white light has been found to benefit not just patients with seasonal affective disorder, but also individuals with other forms of depression, and some studies have suggested it might have a (modest) anxiolytic effect as well (Terman & Terman, 2005; Youngstedt & Kripke, 2007). In addition, individuals seeking to improve their cognitive performance (the topic of Study 2) are likely to benefit from the interventions of Study 1 (bright white light) and Study 3 (app-based psychotherapeutic exercises). Due to the substantial effect mental health has on cognitive performance, this is, in particular, true for populations for which these interventions constitute mental health treatments (individuals with seasonal affective disorder, anxiety, etc)—but not just them. Bright white light has been found to increase alertness and executive function in individuals with seasonal affective disorder and also in healthy individuals (Huang et al., 2024). Psychotherapeutic exercises, in particular mindfulness exercises, have been found to improve cognitive performance also in healthy individuals (Zeidan et al., 2010). Conversely, some studies suggest that creatine supplementation might reduce symptoms of depression, although this evidence is still very preliminary (Bakian et al., 2020).

4.1.2 Accessibility

Any discussion of how to improve mental health and cognition must also consider how easily these interventions can reach the people who can benefit from them. The interventions explored in this dissertation each offer ways to improve accessibility and scalability in mental health care and cognitive enhancement.

One advantage of BROAD light therapy, the novel approach to treating SAD tested in Study 1, is its accessibility. Individuals can move around and engage in other activities while receiving treatment. This flexibility could make the treatment more accessible than the light box treatment to some people. The materials required for BROAD therapy—bright white bulbs and socket cords—are publicly available and easily ordered online. Accessibility is further supported by the fact that the therapy can be administered at home or in the office, eliminating the need to travel to a clinic for treatment. This not only reduces logistical barriers, such as transportation time and costs, but also offers psychological comfort for patients who might feel more at ease in their own environment rather than in a clinical setting. While BROAD therapy is more expensive than a standard light therapy lamp, it remains cheaper than clinic-based light therapy or ongoing psychotherapy sessions. For patients who value the added flexibility and convenience, this additional cost could be justified. If future research confirms that maximising illuminance at eye level leads to greater treatment effects for BROAD light therapy than standard light box therapy, as suggested by the exploratory analyses, the benefits of BROAD therapy would likely outweigh its costs for many patients, and insurance coverage could make the therapy even more accessible.

Creatine supplementation offers an accessible approach to cognitive enhancement if a positive effect is confirmed by future research. As a widely available, inexpensive over-the-counter supplement, creatine is already commonly used by athletes and fitness enthusiasts. It requires no special equipment, no time-consuming process, and little effort to implement.

Finally, the app-based exercises against anxiety tested in this dissertation are a very accessible mental health treatment. In contrast to traditional psychotherapy, patients do not need to wait for a free spot, which can take a long time. They do not need to travel to the therapist nor face the stigma that unfortunately still often surrounds visiting a therapist. If psychotherapy is not covered by insurance, paying for it can be very expensive for the patient. Self-help apps are cheap, because once developed, it costs almost nothing to deliver them to an additional person. This also makes them a very scalable solution. I would like to emphasise that in many cases self-help apps are unlikely to be sufficient. However, they might fill the gap for people who would not seek or receive any treatment otherwise, supplement psychotherapy, prevent mental health from deteriorating to the point of needing psychotherapy, and in some cases replace it.

The interventions in the three studies do not only have in common that they all focus on accessible interventions, but they also overlap in the factors contributing to making these interventions accessible. For example, the interventions in all three studies are relatively low-cost. Creatine monohydrate in powder form commonly costs less than 20 €/kg on amazon.de. Using the common supplementation regime (also used in Study 2) of 5 g/day, the costs for creatine supplementation are 10 Cents per day, i.e. 3 € per month. Study 2 had the cheapest intervention. The psychotherapeutic exercises tested in Study 3 were all part of the Mind Ease app. These exercises can be accessed in the app for free, although for full functionality of the app users have to pay 10 USD (9 €) per month, or less if they choose yearly (50 USD) or lifetime (150 USD) access. The one-time cost of the materials used for BROAD light therapy per participant in Study 1 was 350 €. In addition, there are energy costs of this treatment. We compensated participants in Study 1 for these expenses with 35 € for 4 weeks of treatment. This makes the intervention in Study 1 the most expensive of the three studies. It might be possible to adapt the treatment to conserve some of its potential benefits over standard light therapy (broad lighting, allowing more movement and longer treatment durations) while reducing the number of bulbs and therefore the costs of

equipment and energy. However, even in its current state, BROAD light therapy is cheap compared to e.g. psychotherapy or the costs of having seasonal affective disorder, and would be worth the increased costs compared to standard light therapy lamps *if* the right setup is confirmed as more effective by future studies.

Another common factor that makes all three studied interventions accessible is that they require minimal effort. BROAD light therapy (Study 1) requires material to be ordered online, set up at home, and after this the lighting only needs to be switched on and off every day. Creatine supplementation (Study 2) only requires consuming a teaspoon of creatine powder mixed in e.g. a glass of water every day. App-based psychotherapeutic exercises (Study 3) require only 5-10 minutes per day.

All three interventions are low-risk. The interventions with the lowest risk of the three studies were the psychotherapeutic exercises in Study 3. No adverse effects were reported. The interventions in Studies 1 and 2 had some side effects: eye ache, headache, and agitation for light therapy; digestional problems and weight gain for creatine supplementation. These side effects were not severe and affected only a minority of participants, however, for affected individuals this might make the interventions unsuitable. It seems plausible that side effects might be reduced by adjusting the illuminance and dispersion of the light and mixing creatine in warm water, respectively, although this would need to be tested.

Availability and scalability are important accessibility factors. All three interventions are highly available and scalable, mainly because they can be acquired and implemented by individuals for themselves without requiring trained staff. The most easily available and most scalable intervention of the three are the psychotherapeutic exercises (Study 3). They only require a smartphone, which has become increasingly common even in low-income countries. No additional infrastructure is needed to scale up the intervention. The interventions in Studies 1 and 2 are commercially available by a large number of companies. The materials can be ordered online or bought in stores. The materials seem to be available

through both of these options, not just in Germany but also in many other countries, including low-income countries. Only individuals in remote areas might have difficulty obtaining these interventions. Scaling these interventions up is not as trivial as with the psychotherapeutic exercises, but seems very feasible, as there seems to be no particular reason why companies could not increase the production of these materials.

The interventions studied in this dissertation all demonstrate the potential to make mental health and cognitive enhancement more accessible and scalable. BROAD light therapy offers a flexible, home-based solution to SAD, while creatine supplementation might provide a low-cost, effortless option for (slight) cognitive enhancement. The self-help app for anxiety overcomes many of the traditional barriers to mental health care, offering a cost-effective and highly scalable solution. Together, these interventions represent advancements in making mental health treatments and cognitive enhancements available to a broader population.

4.1.3 RCTs, CONSORT, Open Science, and frequentist, Bayesian, and robust statistics

This section discusses how the three studies adhered to best practices in clinical research, including the randomised controlled trial (RCT) design, CONSORT guidelines, Open Science principles, and statistical methods.

All three studies had a randomised design. This ensured that differences between the groups were not due to self-selection and that known or unknown confounding variables were likely to be balanced between the groups. For known confounding variables, such as the baseline level of depression/anxiety/cognitive performance, I checked if this balance between groups had been achieved, with overall very satisfactory results. Studies 2 and 3 randomised participants individually. Because Study 1 had a smaller sample size, participants were randomised in three strata based on baseline symptom severity and in

blocks of varying sizes. This made it more likely that the groups would have an equal number of participants and that the baseline symptom severity would be similar across groups, while still keeping the assignment unpredictable. This was not necessary for Studies 2 and 3, because larger sample sizes make substantial, random differences between groups unlikely.

All three studies included control groups. The control group in Study 1 was active and consisted of the standard of care in light therapy. In Study 2, the control substance was maltodextrin, which is superficially similar to the treatment substance (creatine). Study 3 had two control groups: An “activity as usual”/measurement-only control group and a reading control group, which allowed us to compare the psychotherapy exercises to a real-world alternative and to an attention-matched alternative, just in case these two would impact anxiety in different ways (they did not). These control groups ruled out effects of time, measurement, attention, and in Study 1 and 2 to a large extent, expectation. Expectation was likely similar between treatment groups but lower for the control group than the treatment groups in Study 3. Given the transparent nature of psychotherapeutic exercises, it is difficult to create a control group that is superficially similar (involves reading text etc.), elicits high treatment expectations, and at the same time is not actually a psychotherapeutic intervention. Research staff was blind and participants were blind to the extent that this was possible, ruling out for the most part effects of researcher or participant expectation. While participants were not told whether they were receiving the intervention of interest or the control intervention, some might have guessed correctly due to side effects or apparent differences between the interventions (see 4.2).

Studies 2 and 3 were confirmatory and followed the respective CONSORT reporting guideline for confirmatory trials, while Study 1, a feasibility study, adhered to the CONSORT guidelines for pilot and feasibility trials. The CONSORT guidelines provide a structured framework for reporting trials, including a detailed account of study design, randomisation,

blinding, and participant flow. This ensured all the relevant information was reported so that the studies can be understood in depth, assessed, and replicated.

All three studies followed Open Science principles. Each study was preregistered and the data were made publicly available, ensuring transparency in the research process. Preregistration involves publicly declaring the study design and analysis plan before data collection begins. Deviations from preregistered plans were clearly documented and justified. For example, we did not preregister a time criterion in Study 3, but it became apparent that some participants took so little time they could not have done the exercise properly, and others took so much time that they most likely did other activities in between. So, I ran and reported two analyses: one with and one without these participants. The results were the same.

The studies combined frequentist, Bayesian, and robust statistical methods to ensure comprehensive and reliable analyses. Frequentist statistics were the primary analysis method in all three studies. In frequentist analysis, p-values are used to assess how likely the observed data would be under the null hypothesis. This approach is widely understood and provides a standard way to assess evidence. In Study 1, the Pearson correlation and its p-value were used to explore the relationships between illuminance in BROAD light therapy and clinical outcomes. In Studies 2 and 3, frequentist methods such as mixed ANOVAs were employed to examine the effects of creatine supplementation and app-based psychotherapeutic exercises, respectively. In addition to frequentist analysis, Study 2 included Bayesian statistical methods. To ensure the reliability of findings and account for potential violations of assumptions, robust (frequentist) statistical methods were also employed in Studies 2 and 3. These methods, including trimming, winsorizing, and bootstrapping, provide more resilient estimates by minimising the influence of outliers and deviations from normality.

Bayesian analysis helps interpret the data in a more nuanced way depending on which effect sizes are of interest. When even small effect sizes are of interest and the effect of an intervention is large and highly significant, such as in Study 3, a Bayesian analysis might be superfluous. In Study 2 however, the effect of the intervention was small and we were interested in better understanding how much to update towards or away from two different hypotheses: a (more realistic) small effect, and the large effect of the study we were replicating and expanding on. The combination of frequentist, Bayesian, and robust methods strengthened the understanding of the findings, as each approach offers unique advantages.

4.2 Limitations

The studies in this thesis are not without limitations. Some limitations are common among the three studies: 1) The specific samples limit generalisation. 2) Differences between the treatments and control conditions limited the extent of blinding. 3) While my samples were larger than is common for these types of studies, the samples were nevertheless not large enough to answer all the questions that would have been interesting to answer.

4.2.1 Study 1

Study 1 had several limitations. The study was not powered to detect differences between BROAD light therapy and the control group, the standard light therapy lamp. Originally, it was planned to recruit participants over several winters to reach a sufficiently high sample size to test differences between groups. However, results of the first winter, i.e., the exploratory correlation between illuminance at eye level in the BROAD light therapy group and improvement in SAD symptoms, suggested that differences in how participants had set up BROAD light therapy in their homes had a substantial impact. The illuminance had varied greatly between participants and had been too low in many cases. The intervention for future winters would have needed to be adjusted so that all participants in the BROAD light therapy group would have a higher illuminance, and the data from the first winter could not have

been merged with the following winters. We considered the feasibility and exploratory correlation results of the first winter to be sufficiently interesting in their own right to be published separately.

Another limitation was that the packaging of the standard light therapy lamp, which advertised the therapeutic properties of the lamp, likely increased the placebo effect of this treatment. By contrast, the packaging for the BROAD light therapy group did not include any such advertising. A further disadvantage for BROAD light therapy might have been aesthetic. I consider the standard light therapy lamp more aesthetically pleasing than the BROAD light therapy set-up, and suspect that correcting this disadvantage might be relevant for mood. Lastly, Study 1 did not include an inactive control group, so it did not prove that the improvement in SAD symptoms in both groups was due to the treatments. Nevertheless, the standard of care (light therapy lamp) employed as the control condition in this study is well-researched and our effect sizes are in line with previous studies.

4.2.2 Study 2

Study 2 had several limitations. Creatine levels in the brain and the blood were not measured and adherence was self-reported. After they completed their last test, we asked participants to guess their supplement order. Participants guessed correctly more often than not, bordering on significance (p = .08). Participants who guessed correctly based their guesses on solubility, negative side effects, and positive side effects. However, it is worth noting that almost all participants, even those who guessed correctly, reported being unsure about their guess. Maintaining complete blinding for participants who experience side effects is challenging. The side effects of creatine supplementation are widely recognised and not serious (Bender et al., 2008; de Souza e Silva et al., 2019; Kreider et al., 2017; Kutz & Gunter, 2003). However, they might make creatine supplementation unadvisable or unappealing for individuals who experience them. It is possible that supplementing creatine

in a different way, e.g. mixed in warm water, might have reduced the likelihood of side effects (see below).

To address solubility differences, we asked participants to stir their supplements into yoghurt. Using yoghurt instead of water might have lowered creatine absorption because yoghurt usually has less water and is cold, both of which reduce creatine solubility (Harris et al., 2002). Even so, yoghurt remains fairly fluid and is about 70–90% water (Bodner-Montville et al., 2006). It also contains carbohydrates (4–18 g/100 g) (Bodner-Montville et al., 2006), which can help the body hold on to creatine (Hultman et al., 1996; Steenge et al., 2000). In addition, yoghurt's pH is lower than water (4 vs. 6.5–9.5) (Hennighausen, 2001; Ott et al., 2000), which may further support absorption (Jäger et al., 2011). Overall, we believe mixing creatine with yoghurt probably allows absorption on par with cold water, though not as well as warm water. For future research, we suggest using cellulose as the placebo and a cellulose–creatine mix as the treatment, since these look very similar when dissolved in water. The alternative—capsules—would mean taking many pills every day, which would likely reduce adherence and greatly increase costs.

The COVID-19 pandemic started in the middle of the study, which meant that testing had to be changed from in-person to over video. The blocktapping (Corsi) task had to be set up digitally. The number of participants was lower for this task, because the digital version of this task was not immediately available. More generally, the pandemic might have introduced noise in the study. On the other hand, the standard deviations in our study were similar to those of Rae et al. (2003), so maybe this effect was not pronounced.

We could not analyse about four percent of the data because it lacked participant and timepoint labels. Although this study had a larger sample size than others, an even bigger sample would be needed to detect smaller but still important effects.

4.2.3 Study 3

Study 3 had several limitations. Only immediate effects on anxiety were assessed, so it remains unclear how the effects of the different app-based psychotherapeutic exercises develop over time. We believe these exercises likely produce lasting benefits, particularly when performed regularly, since they use methods already known to have long-term positive effects. However, more investigation is needed to determine how well these specific exercises work and how their impact changes and compares over time.

Treatment expectations were likely higher in the exercise groups than in the control groups (measurement-only control and reading control). This is a problem for the comparisons between control groups and exercise groups, but not between exercise groups. It is not straightforward to determine an appropriate control group for psychotherapeutic exercises and this question is debated in the literature (Gaab, 2023). A control intervention that looks like a psychotherapeutic exercise and elicits high treatment expectations might be in fact considered psychotherapeutic.

Participants in this study were recruited through the Positly platform, which uses Amazon Mechanical Turk along with extra quality checks to reduce inattentive users and spammers (see the Participants section in the paper for details). Because these participants may have been less personally interested in the exercises, they might have put in less effort than typical app users—suggesting that real-world effects might be stronger than what we observed. Limiting recruitment to the United States and relying on volunteers who are comfortable with Positly and willing to join the study may also affect how widely these findings apply.

The original plan was to have a substantially higher sample size to be able to detect even small effects. Because recruitment was slower and more expensive than expected, we adjusted the sample size in the preregistration to N = 1126, i.e. twice the size collected at

that point. This sample size was not powered to detect small differences between exercises, but was still powered to detect medium-sized differences.

4.3 Future research

In addition to addressing the limitations described in section 4.2, there are a number of promising directions for future research on improving mental health and cognition.

One possibly promising direction for future research might be to study how to personalise treatment. For example, for treatment against seasonal affective disorder, different individuals might have different optimal times of treatment (Lewy et al., 2006), levels of illuminance at eye level, colour temperatures, and other details of the set-up of the light. Some people might respond better to light therapy and some to other treatments such as psychotherapy or medication. For other individuals, less talked about factors might play a role, e.g. temperature. For apps against anxiety, the effect of an intervention might depend on the situation, type of anxiety, culture, and other user characteristics.

Another helpful avenue for future research might be to better understand side effects of the interventions presented here. Light therapy can cause headache, eye ache, and agitation in some people. It would be good to know to what extent this can be avoided without sacrificing the treatment effect, e.g. by diffusing the light better or by downregulating the illuminance for sensitive patients. Creatine supplementation causes digestive discomfort in some people. It would be good to know to what extent this can be avoided by e.g. using warm water to dissolve the creatine or dividing the creatine dose into smaller doses spread throughout the day.

While studies in this area are already generally randomised, controlled, and mostly CONSORT-compliant, they are often not preregistered and rarely employ Bayesian or robust statistics. I strongly recommend future research be preregistered to improve the transparency of the research process and thus the credibility of the findings. I recommend

future research to use additional statistical approaches, in particular robust statistics including bootstrapping and trimming, to give a more complete picture and to show to what extent findings are robust to deviations from normality.

4.4 Conclusions

The studies included in this dissertation contribute to our understanding of how to improve mental health and cognition in an accessible manner. Three interventions—a new version of light therapy against seasonal affective disorder, creatine supplementation to improve cognition, and a self-help app against anxiety—were tested and found to be overall promising and worth further investigation.

In the overall quest to find scalable interventions for mental health and cognition, this thesis represents an important extension of previous research by 1) finding new ways to make light therapy more accessible and potentially more effective, 2) testing the effect of creatine on many different tests of cognition and with the largest sample size found in the literature to date, 3) providing the first official test of the exercises in the Mind Ease app against anxiety and testing an unusually large number of exercises in the same standardised fashion.

REFERENCES

Alhola, P., & Polo-Kantola, P. (2007). Sleep deprivation: Impact on cognitive performance. *Neuropsychiatric Disease and Treatment*, *3*, 553–567.

Ali, A., Ambler, G., Strydom, A., Rai, D., Cooper, C., McManus, S., Weich, S., Meltzer, H., Dein, S., & Hassiotis, A. (2013). The relationship between happiness and intelligent quotient: the contribution of socio-economic and clinical factors. *Psychological Medicine*, *43*(6), 1303–1312.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)*. American Psychiatric Pub.

American Psychological Association. (2018a). *Cognition*. *APA Dictionary of Psychology*. https://dictionary.apa.org/cognition

American Psychological Association. (2018b). *Mental health*. *APA Dictionary of Psychology*. https://dictionary.apa.org/mental-health

Anderson, I. M., Ferrier, I. N., Baldwin, R. C., Cowen, P. J., Howard, L., Lewis, G., Matthews, K., McAllister-Williams, R. H., Peveler, R. C., Scott, J., & Tylee, A. (2008). Evidence-based guidelines for treating depressive disorders with antidepressants: a revision of the 2000 British Association for Psychopharmacology guidelines. *Journal of Psychopharmacology* , *22*(4), 343–396.

Anderson, J. L., Glod, C. A., Dai, J., Cao, Y., & Lockley, S. W. (2009). Lux vs. wavelength in light treatment of Seasonal Affective Disorder. *Acta Psychiatrica Scandinavica*, *120*(3), 203–212.

Aschenbrenner, S., Tucha, O., & Lange, K. W. (2000). *Regensburger Wortflüssigkeits-Test: RWT*. Hogrefe, Verlag für Psychologie.

Avgerinos, K. I., Spyrou, N., Bougioukas, K. I., & Kapogiannis, D. (2018). Effects of creatine supplementation on cognitive function of healthy individuals: A systematic review of randomized controlled trials. *Experimental Gerontology*, *108*, 166–173.

Bakian, A. V., Huber, R. S., Scholl, L., Renshaw, P. F., & Kondo, D. (2020). Dietary creatine intake and depression risk among U.S. adults. *Translational Psychiatry*, *10*(1), 52.

Ball, T. M., Stein, M. B., & Paulus, M. P. (2014). Toward the application of functional neuroimaging to individualized treatment for anxiety and depression: Review: Neuroimaging and individualized treatment. *Depression and Anxiety*, *31*(11), 920–933.

Batty, G. D., Mortensen, E. L., & Osler, M. (2005). Childhood IQ in relation to later psychiatric disorder: evidence from a Danish birth cohort study. *The British Journal of Psychiatry: The Journal of Mental Science*, *187*(2), 180–181.

Bauer, M., Pfennig, A., Severus, E., Whybrow, P. C., Angst, J., Moeller, H.-J., & on Unipolar Depressive Disorders, Š. B. of T. T. F. (2013). World Federation of Societies of Biological Psychiatry (WFSBP) guidelines for biological treatment of unipolar depressive disorders, part 1: update 2013 on the acute and continuation treatment of unipolar depressive disorders. *The World Journal of Biological Psychiatry: The Official Journal of the World Federation of Societies of Biological Psychiatry*, *14*(5), 334–385.

Bäumler, G., & Stroop, J. R. (1985). *Farbe-Wort-Interferenztest nach JR Stroop (FWIT)*. Hogrefe, Verlag für Psychologie.

Bender, A., Samtleben, W., Elstner, M., & Klopstock, T. (2008). Long-term creatine supplementation is safe in aged patients with Parkinson disease. *Nutrition Research* , *28*(3), 172–178.

Benedict, R. H. B., Schretlen, D., Groninger, L., Dobraski, M., & Shpritz, B. (1996). Revision of the Brief Visuospatial Memory Test: Studies of normal performance, reliability, and validity. *Psychological Assessment*, *8*(2), 145.

Benton, D., & Donohoe, R. (2011). The influence of creatine supplementation on the cognitive functioning of vegetarians and omnivores. *The British Journal of Nutrition*, *105*(7), 1100–1105.

Bertua, C., Anderson, N., & Salgado, J. F. (2005). The predictive validity of cognitive ability tests: A UK meta-analysis. *Journal of Occupational and Organizational Psychology*, *78*(3), 387–409.

Bhatia, V. B. (2001). *Textbook of astronomy and astrophysics with elements of cosmology*. Narosa Publishing House.

Blazer, D. G., Kessler, R. C., & Swartz, M. S. (1998). Epidemiology of recurrent major and minor depression with a seasonal pattern. The National Comorbidity Survey. *The British Journal of Psychiatry: The Journal of Mental Science*, *172*, 164–167.

Bodner-Montville, J., Ahuja, J. K. C., Ingwersen, L. A., Haggerty, E. S., Enns, C. W., & Perloff, B. P. (2006). USDA Food and Nutrient Database for Dietary Studies: Released on the web. *Journal of Food Composition and Analysis: An Official Publication of the United Nations University, International Network of Food Data Systems*, *19*, S100–S107.

Bostrom, N., & Sandberg, A. (2009). Cognitive enhancement: methods, ethics, regulatory challenges. *Science and Engineering Ethics*, *15*(3), 311–341.

Botanov, Y., & Ilardi, S. S. (2013). The acute side effects of bright light therapy: a placebo-controlled investigation. *PloS One*, *8*(9), e75893.

Branch, J. D. (2003). Effect of creatine supplementation on body composition and performance: a meta-analysis. *International Journal of Sport Nutrition and Exercise Metabolism*, *13*(2), 198–226.

Breit, M., Brunner, M., Molenaar, D., & Preckel, F. (2022). Differentiation hypotheses of intelligence: A systematic review of the empirical evidence and an agenda for future research. *Psychological Bulletin*, *148*(7-8), 518–554.

Brickenkamp, R. (2002). Test d2: Aufmerksamkeits-Belastungs-Test Hogrefe. *Testzentrale: Göttingen, Germany*.

Brooks, S. J., & Stein, D. J. (2015). A systematic review of the neural bases of psychotherapy for anxiety and related disorders. *Dialogues in Clinical Neuroscience*, *17*(3), 261–279.

Brosnan, M. E., & Brosnan, J. T. (2016). The role of dietary creatine. *Amino Acids*, *48*(8), 1785–1791.

Butts, J., Jacobs, B., & Silvis, M. (2018). Creatine Use in Sports. *Sports Health*, *10*(1), 31–34.

Cajochen, C., Zeitzer, J. M., Czeisler, C. A., & Dijk, D. J. (2000). Dose-response relationship for light intensity and ocular and electroencephalographic correlates of human alertness. *Behavioural Brain Research*, *115*(1), 75–83.

Canazei, M., Pohl, W., Bauernhofer, K., Papousek, I., Lackner, H. K., Bliem, H. R., Marksteiner, J., & Weiss, E. M. (2017). Psychophysiological Effects of a Single, Short, and Moderately Bright Room Light Exposure on Mildly Depressed Geriatric Inpatients: A Pilot Study. *Gerontology*, *63*(4), 308–317.

Candow, D. G., Forbes, S. C., Ostojic, S. M., Prokopidis, K., Stock, M. S., Harmon, K. K., & Faulkner, P. (2023). “heads up” for creatine supplementation and its potential applications for brain health and function. *Sports Medicine (Auckland, N.Z.)*, *53*(Suppl 1), 49–65.

Center for Open Science. (2025). *What is open science?* https://www.cos.io/open-science

Chang, Y. K., Labban, J. D., Gapin, J. I., & Etnier, J. L. (2012). The effects of acute exercise on cognitive performance: a meta-analysis. *Brain Research*, *1453*, 87–101.

Chapman, D. (2023). *YOU NEED MORE LUX*. Meaningness. https://meaningness.com/sad-light-led-lux

Chen, M. C., Chang, C., Glover, G. H., & Gotlib, I. H. (2014). Increased insula coactivation with salience networks in insomnia. *Biological Psychology*, *97*, 1–8.

Chen, Y.-F., Huang, X.-Y., Chien, C.-H., & Cheng, J.-F. (2017). The effectiveness of diaphragmatic breathing relaxation training for reducing anxiety. *Perspectives in Psychiatric Care*, *53*(4), 329–336.

Chisholm, D., Sweeny, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., & Saxena, S. (2016). Scaling-up treatment of depression and anxiety: a global return on investment analysis. *The Lancet. Psychiatry*, *3*(5), 415–424.

Clark, D. A. (2013). Cognitive Restructuring. In *The Wiley Handbook of Cognitive Behavioral Therapy* (pp. 1–22). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118528563.wbcbt02

Clark, J. F., & Cecil, K. M. (2015). Diagnostic methods and recommendations for the cerebral

creatine deficiency syndromes. *Pediatric Research*, *77*(3), 398–405.

Conrad, A., & Roth, W. T. (2007). Muscle relaxation therapy for anxiety disorders: it works but how? *Journal of Anxiety Disorders*, *21*(3), 243–264.

Daniels, N. (2000). Normal functioning and the treatment-enhancement distinction. *Cambridge Quarterly of Healthcare Ethics: CQ: The International Journal of Healthcare Ethics Committees*, *9*(3), 309–322.

Dechent, P., Pouwels, P. J., Wilken, B., Hanefeld, F., & Frahm, J. (1999). Increase of total creatine in human brain after oral supplementation of creatine-monohydrate. *The American Journal of Physiology*, *277*(3), R698–R704.

de Souza e Silva, A., Pertille, A., Reis Barbosa, C. G., Aparecida de Oliveira Silva, J., de Jesus, D. V., Ribeiro, A. G. S. V., Baganha, R. J., & de Oliveira, J. J. (2019). Effects of Creatine Supplementation on Renal Function: A Systematic Review and Meta-Analysis. *Journal of Renal Nutrition: The Official Journal of the Council on Renal Nutrition of the National Kidney Foundation*, *29*(6), 480–489.

Doehrmann, O., Ghosh, S. S., Polli, F. E., Reynolds, G. O., Horn, F., Keshavan, A., Triantafyllou, C., Saygin, Z. M., Whitfield-Gabrieli, S., Hofmann, S. G., Pollack, M., & Gabrieli, J. D. (2013). Predicting treatment response in social anxiety disorder from functional magnetic resonance imaging. *JAMA Psychiatry (Chicago, Ill.)*, *70*(1), 87–97.

Du, B., Tandoc, M. C., Mack, M. L., & Siegel, J. A. (2020). Indoor CO2 concentrations and cognitive function: A critical review. *Indoor Air*, *30*(6), 1067–1082.

Eldridge, S. M., Chan, C. L., Campbell, M. J., Bond, C. M., Hopewell, S., Thabane, L., Lancaster, G. A., & PAFS consensus group. (2016). CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *Pilot and Feasibility Studies*, *2*, 64.

Etnier, J. L., & Chang, Y.-K. (2019). Exercise, cognitive function, and the brain: Advancing our understanding of complex relationships. *Journal of Sport and Health Science*, *8*(4), 299–300.

Fernando, S. D., Rodrigo, C., & Rajapakse, S. (2010). The “hidden” burden of malaria: cognitive impairment following infection. *Malaria Journal*, *9*(1), 366.

Field, A. P. (2018). *Discovering Statistics Using IBM SPSS Statistics, 5th Edition*. Sage.

Firth, J., Torous, J., Nicholas, J., Carney, R., Pratap, A., Rosenbaum, S., & Sarris, J. (2017). The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. *World Psychiatry: Official Journal of the World Psychiatric Association* , *16*(3), 287–298.

Firth, J., Torous, J., Nicholas, J., Carney, R., Rosenbaum, S., & Sarris, J. (2017). Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. *Journal of Affective Disorders*, *218*, 15–22.

Fisher, P. M., Madsen, M. K., Mc Mahon, B., Holst, K. K., Andersen, S. B., Laursen, H. R., Hasholt, L. F., Siebner, H. R., & Knudsen, G. M. (2014). Three-week bright-light intervention has dose-related effects on threat-related corticolimbic reactivity and functional coupling. *Biological Psychiatry*, *76*(4), 332–339.

Fonzo, G. A., Ramsawh, H. J., Flagan, T. M., Sullivan, S. G., Simmons, A. N., Paulus, M. P., & Stein, M. B. (2014). Cognitive-behavioral therapy for generalized anxiety disorder is associated with attenuation of limbic activation to threat-related facial emotions. *Journal of Affective Disorders*, *169*, 76–85.

Gaab, J. (2023). Why psychotherapy is an open-label placebo and open-label placebos are psychotherapy. In L. Colloca, J. Noel, P. D. Franklin, & C. Seneviratne (Eds.), *Placebo Effects Through the Lens of Translational Research* (pp. 240–C4.4P66). Oxford University PressNew York.

Gardner, A., & Boles, R. G. (2011). Beyond the serotonin hypothesis: Mitochondria, inflammation and neurodegeneration in major depression and affective spectrum disorders. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *35*(3), 730–743.

Gelenberg, A. J., Freeman, M. P., Markowitz, J. C., Rosenbaum, J. F., Thase, M. E., Trivedi, M. H., & Van Rhoads, R. S. (2010). Practice guideline for the treatment of patients with major depressive disorder (3rd). *American Journal of Psychiatry*, *157*(4), 1–45.

Gignac, G. E. (2015). Raven's is not a pure measure of general intelligence: Implications for

g factor theory and the brief measurement of g. *Intelligence*, *52*, 71–79.

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, *102*, 74–78.

Global Council on Brain Health. (2019). Real Deal Brain Health Supplements: GCBH Recommendations Vitamins, Minerals, Other Dietary Supplements. *AARP*. https://doi.org/10.26419/pia.00094.001

Golden, R. N., Gaynes, B. N., Ekstrom, R. D., Hamer, R. M., Jacobsen, F. M., Suppes, T., Wisner, K. L., & Nemeroff, C. B. (2005). The efficacy of light therapy in the treatment of mood disorders: a review and meta-analysis of the evidence. *The American Journal of Psychiatry*, *162*(4), 656–662.

Goldin, P. R., Thurston, M., Allende, S., Moodie, C., Dixon, M. L., Heimberg, R. G., & Gross, J. J. (2021). Evaluation of cognitive behavioral therapy vs mindfulness meditation in brain changes during reappraisal and acceptance among patients with social anxiety disorder: A randomized clinical trial: A randomized clinical trial. *JAMA Psychiatry (Chicago, Ill.)*, *78*(10), 1134–1142.

Goldin, P. R., Ziv, M., Jazaieri, H., Hahn, K., Heimberg, R., & Gross, J. (2013). Impact of cognitive behavioral therapy for social anxiety disorder on the neural dynamics of cognitive reappraisal of negative self-beliefs: randomized clinical trial: Randomized clinical trial. *JAMA Psychiatry (Chicago, Ill.)*, *70*(10), 1048–1056.

Goldin, P. R., Ziv, M., Jazaieri, H., Weeks, J., Heimberg, R. G., & Gross, J. J. (2014). Impact of cognitive-behavioral therapy for social anxiety disorder on the neural bases of emotional reactivity to and regulation of social evaluation. *Behaviour Research and Therapy*, *62*, 97–106.

Goldin, P., Ziv, M., Jazaieri, H., Hahn, K., & Gross, J. J. (2013). MBSR vs aerobic exercise in social anxiety: fMRI of emotion regulation of negative self-beliefs. *Social Cognitive and Affective Neuroscience*, *8*(1), 65–72.

Gordji-Nejad, A., Matusch, A., Kleedörfer, S., Jayeshkumar Patel, H., Drzezga, A., Elmenhorst, D., Binkofski, F., & Bauer, A. (2024). Single dose creatine improves

cognitive performance and induces changes in cerebral high energy phosphates during sleep deprivation. *Scientific Reports*, *14*(1), 4937.

Gottfredson, L. S., & Deary, I. J. (2004). Intelligence predicts health and longevity, but why? *Current Directions in Psychological Science*, *13*(1), 1–4.

Guillén-Riquelme, A., & Buela-Casal, G. (2011). Psychometric revision and differential item functioning in the State Trait Anxiety Inventory (STAI). *Psicothema*, *23*(3), 510–515.

Harper, D. G., Plante, D. T., Jensen, J. E., Ravichandran, C., Buxton, O. M., Benson, K. L., O'Connor, S. P., Renshaw, P. F., & Winkelman, J. W. (2013). Energetic and cell membrane metabolic products in patients with primary insomnia: a 31-phosphorus magnetic resonance spectroscopy study at 4 tesla. *Sleep*, *36*(4), 493–500.

Harris, R. C., Nevill, M., Harris, D. B., Fallowfield, J. L., Bogdanis, G. C., & Wise, J. A. (2002). Absorption of creatine supplied as a drink, in meat or in solid form. *Journal of Sports Sciences*, *20*(2), 147–151.

Hayes, S. C., & Hofmann, S. G. (2017). The third wave of cognitive behavioral therapy and the rise of process-based care. *World Psychiatry: Official Journal of the World Psychiatric Association* , *16*(3), 245–246.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2016). *Acceptance and Commitment Therapy, Second Edition: The Process and Practice of Mindful Change*. Guilford Publications.

Hellem, T. L., Sung, Y.-H., Shi, X.-F., Pett, M. A., Latendresse, G., Morgan, J., Huber, R. S., Kuykendall, D., Lundberg, K. J., & Renshaw, P. F. (2015). Creatine as a novel treatment for depression in females using methamphetamine: A pilot study. *Journal of Dual Diagnosis*, *11*(3-4), 189–202.

Hennighausen, R. H. (2001). Die neue Trinkwasserverordnung - TrinkwV 2001. *Gesundheitswesen* , *63*(11), 724–730.

Hofmann, S. G. (2011). *An Introduction to Modern CBT: Psychological Solutions to Mental Health Problems*. John Wiley & Sons.

Hölzel, B. K., Hoge, E. A., Greve, D. N., Gard, T., Creswell, J. D., Brown, K. W., Barrett, L. F., Schwartz, C., Vaitl, D., & Lazar, S. W. (2013). Neural mechanisms of symptom

improvements in generalized anxiety disorder following mindfulness training. *NeuroImage. Clinical*, *2*, 448–458.

Huang, X., Tao, Q., & Ren, C. (2024). A comprehensive overview of the neural mechanisms of light therapy. *Neuroscience Bulletin*, *40*(3), 350–362.

Huhn, M., Tardy, M., Spineli, L. M., Kissling, W., Förstl, H., Pitschel-Walz, G., Leucht, C., Samara, M., Dold, M., Davis, J. M., & Leucht, S. (2014). Efficacy of pharmacotherapy and psychotherapy for adult psychiatric disorders: a systematic overview of meta-analyses: A systematic overview of meta-analyses. *JAMA Psychiatry (Chicago, Ill.)*, *71*(6), 706–715.

Hultman, E., Söderlund, K., Timmons, J. A., Cederblad, G., & Greenhaff, P. L. (1996). Muscle creatine loading in men. *Journal of Applied Physiology*, *81*(1), 232–237.

Jäger, R., Purpura, M., Shao, A., Inoue, T., & Kreider, R. B. (2011). Analysis of the efficacy, safety, and regulatory status of novel forms of creatine. *Amino Acids*, *40*(5), 1369–1383.

Jokela, M. (2022). Why is cognitive ability associated with psychological distress and wellbeing? Exploring psychological, biological, and social mechanisms. *Personality and Individual Differences*, *192*(111592), 111592.

Kabat-Zinn, J. (2009). *Wherever You Go, There You Are: Mindfulness Meditation in Everyday Life*. Hachette UK.

Khandaker, G. M., Stochl, J., Zammit, S., Goodyer, I., Lewis, G., & Jones, P. B. (2018). Childhood inflammatory markers and intelligence as predictors of subsequent persistent depressive symptoms: a longitudinal cohort study. *Psychological Medicine*, *48*(9), 1514–1522.

Killgore, W. D. S., & Weber, M. (2014). Sleep Deprivation and Cognitive Performance. In *Sleep Deprivation and Disease* (pp. 209–229). Springer New York.

Kircher, T., Arolt, V., Jansen, A., Pyka, M., Reinhardt, I., Kellermann, T., Konrad, C., Lueken, U., Gloster, A. T., Gerlach, A. L., Ströhle, A., Wittmann, A., Pfleiderer, B., Wittchen, H.-U., & Straube, B. (2013). Effect of cognitive-behavioral therapy on neural correlates of fear conditioning in panic disorder. *Biological Psychiatry*, *73*(1), 93–101.

Klinedinst, N. J., & Regenold, W. T. (2015). A mitochondrial bioenergetic basis of depression. *Journal of Bioenergetics and Biomembranes*, *47*(1-2), 155–171.

Klumpp, H., Fitzgerald, D. A., & Phan, K. L. (2013). Neural predictors and mechanisms of cognitive behavioral therapy on threat processing in social anxiety disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *45*, 83–91.

Knol, M. J., Pestman, W. R., & Grobbee, D. E. (2011). The (mis)use of overlap of confidence intervals to assess effect modification. *European Journal of Epidemiology*, *26*(4), 253–254.

Kondo, D. G., Forrest, L. N., Shi, X., Sung, Y.-H., Hellem, T. L., Huber, R. S., & Renshaw, P. F. (2016). Creatine target engagement with brain bioenergetics: a dose-ranging phosphorus-31 magnetic resonance spectroscopy study of adolescent females with SSRI-resistant depression. *Amino Acids*, *48*(8), 1941–1954.

Kondo, D. G., Sung, Y.-H., Hellem, T. L., Fiedler, K. K., Shi, X., Jeong, E.-K., & Renshaw, P. F. (2011). Open-label adjunctive creatine for female adolescents with SSRI-resistant major depressive disorder: a 31-phosphorus magnetic resonance spectroscopy study. *Journal of Affective Disorders*, *135*(1-3), 354–361.

Kreider, R. B., Kalman, D. S., Antonio, J., Ziegenfuss, T. N., Wildman, R., Collins, R., Candow, D. G., Kleiner, S. M., Almada, A. L., & Lopez, H. L. (2017). International Society of Sports Nutrition position stand: safety and efficacy of creatine supplementation in exercise, sport, and medicine. *Journal of the International Society of Sports Nutrition*, *14*, 18.

Kripke, D. F., Mullaney, D. J., Klauber, M. R., Risch, S. C., & Gillin, J. C. (1992). Controlled trial of bright light for nonseasonal major depressive disorders. *Biological Psychiatry*, *31*(2), 119–134.

Kripke, D. F., Risch, S. C., & Janowsky, D. (1983). Bright white light alleviates depression. *Psychiatry Research*, *10*(2), 105–112.

Kurlansik, S. L., & Ibay, A. D. (2012). Seasonal affective disorder. *American Family Physician*, *86*(11), 1037–1041.

Kutz, M. R., & Gunter, M. J. (2003). Creatine monohydrate supplementation on body weight and percent body fat. *Journal of Strength and Conditioning Research / National Strength & Conditioning Association*, *17*(4), 817–821.

Kyeong, S., Kim, J., Kim, D. J., Kim, H. E., & Kim, J.-J. (2017). Effects of gratitude meditation on neural network functional connectivity and brain-heart coupling. *Scientific Reports*, *7*(1), 5058.

Lehrl, S. (2005). *Manual zum MWT-B*. Spitta-Verlag.

Leucht, S., Hierl, S., Kissling, W., Dold, M., & Davis, J. M. (2012). Putting the efficacy of psychiatric and general medicine medication into perspective: review of meta-analyses. *The British Journal of Psychiatry: The Journal of Mental Science*, *200*(2), 97–106.

Levitan, R. D. (2007). The chronobiology and neurobiology of winter seasonal affective disorder. *Dialogues in Clinical Neuroscience*, *9*(3), 315–324.

Levitt, A. J., Boyle, M. H., Joffe, R. T., & Baumal, Z. (2000). Estimated prevalence of the seasonal subtype of major depression in a Canadian community sample. *Canadian Journal of Psychiatry. Revue Canadienne de Psychiatrie*, *45*(7), 650–654.

Lewy, A. J., Lefler, B. J., Emens, J. S., & Bauer, V. K. (2006). The circadian basis of winter depression. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(19), 7414–7419.

Li, D. H. W., Cheung, G. H. W., Cheung, K. L., & Lam, T. N. T. (2010). Determination of vertical daylight illuminance under non-overcast sky conditions. *Building and Environment*, *45*(2), 498–508.

Li, J., & Xu, S. (2023). Diagnosis and treatment of X-linked creatine transporter deficiency: Case report and literature review. *Brain Sciences*, *13*(10), 1382.

Li, R., Qin, W., Zhang, Y., Jiang, T., & Yu, C. (2012). The neuronal correlates of digits backward are revealed by voxel-based morphometry and resting-state functional connectivity analyses. *PloS One*, *7*(2), e31877.

Lueken, U., Straube, B., Konrad, C., Wittchen, H.-U., Ströhle, A., Wittmann, A., Pfleiderer, B., Uhlmann, C., Arolt, V., Jansen, A., & Kircher, T. (2013). Neural substrates of

treatment response to cognitive-behavioral therapy in panic disorder with agoraphobia. *The American Journal of Psychiatry*, *170*(11), 1345–1355.

Lux, S., Helmstaedter, C., & Lendt, M. (2001). *Verbaler Lern- und Merkfähigkeitstest: VLMT ; Manual*. Beltz-Test.

Lyoo, I. K., Kong, S. W., Sung, S. M., Hirashima, F., Parow, A., Hennen, J., Cohen, B. M., & Renshaw, P. F. (2003). Multinuclear magnetic resonance spectroscopy of high-energy phosphate metabolites in human brain following oral supplementation of creatine-monohydrate. *Psychiatry Research*, *123*(2), 87–100.

Maddock, A., & Blair, C. (2023). How do mindfulness-based programmes improve anxiety, depression and psychological distress? A systematic review. *Current Psychology (New Brunswick, N.J.)*, *42*(12), 10200–10222.

Ma, J., Kim, M., Kim, J., Hong, G., Namgung, E., Park, S., Lim, S. M., Lyoo, I. K., & Yoon, S. (2020). Decreased functional connectivity within the salience network after two-week morning bright light exposure in individuals with sleep disturbances: a preliminary randomized controlled trial. *Sleep Medicine*, *74*, 66–72.

Månsson, K. N. T., Carlbring, P., Frick, A., Engman, J., Olsson, C.-J., Bodlund, O., Furmark, T., & Andersson, G. (2013). Altered neural correlates of affective processing after internet-delivered cognitive behavior therapy for social anxiety disorder. *Psychiatry Research*, *214*(3), 229–237.

Månsson, K. N. T., Frick, A., Boraxbekk, C.-J., Marquand, A. F., Williams, S. C. R., Carlbring, P., Andersson, G., & Furmark, T. (2015). Predicting long-term outcome of Internet-delivered cognitive behavior therapy for social anxiety disorder using fMRI and support vector machine learning. *Translational Psychiatry*, *5*(3), e530.

Marteau, T. M., & Bekker, H. (1992). The development of a six‐item short‐form of the state scale of the Spielberger State—Trait Anxiety Inventory (STAI). *The British Journal of Clinical Psychology*, *31*(3), 301–306.

Mårtensson, B., Pettersson, A., Berglund, L., & Ekselius, L. (2015). Bright white light therapy in depression: A critical review of the evidence. *Journal of Affective Disorders*, *182*, 1–7.

Maslowsky, J., Mogg, K., Bradley, B. P., McClure-Tone, E., Ernst, M., Pine, D. S., & Monk, C. S. (2010). A preliminary investigation of neural correlates of treatment in adolescents with generalized anxiety disorder. *Journal of Child and Adolescent Psychopharmacology*, *20*(2), 105–111.

Matour, S., Mahdavinejad, M., & Fayaz, R. (2017). Horizontal and Vertical Illuminance Correlations in a Case Study for Shaded South Facing Surfaces. *International Journal of Energy and Environmental Engineering*, *11*(12), 1046–1051.

McClure, E. B., Adler, A., Monk, C. S., Cameron, J., Smith, S., Nelson, E. E., Leibenluft, E., Ernst, M., & Pine, D. S. (2007). fMRI predictors of treatment outcome in pediatric anxiety disorders. *Psychopharmacology*, *191*(1), 97–105.

McDonagh, B. (2015). *Dare: The New Way to End Anxiety and Stop Panic Attacks*. BMD Publishing.

Merege-Filho, C. A. A., Otaduy, M. C. G., de Sá-Pinto, A. L., de Oliveira, M. O., de Souza Gonçalves, L., Hayashi, A. P. T., Roschel, H., Pereira, R. M. R., Silva, C. A., Brucki, S. M. D., da Costa Leite, C., & Gualano, B. (2017). Does brain creatine content rely on exogenous creatine in healthy youth? A proof-of-principle study. *Applied Physiology, Nutrition, and Metabolism = Physiologie Appliquee, Nutrition et Metabolisme*, *42*(2), 128–134.

Michalak, E. E., Wilkinson, C., Dowrick, C., & Wilkinson, G. (2001). Seasonal affective disorder: prevalence, detection and current treatment in North Wales. *The British Journal of Psychiatry: The Journal of Mental Science*, *179*, 31–34.

Mikkelsen, K., Stojanovska, L., Polenakovic, M., Bosevski, M., & Apostolopoulos, V. (2017). Exercise and mental health. *Maturitas*, *106*, 48–56.

Ng, T. W. H., Eby, L. T., Sorensen, K. L., & Feldman, D. C. (2005). Predictors of objective and subjective career success: A meta‐analysis. *Personnel Psychology*, *58*(2), 367–408.

Nofzinger, E. A., Buysse, D. J., Germain, A., Price, J. C., Miewald, J. M., & Kupfer, D. J. (2004). Functional neuroimaging evidence for hyperarousal in insomnia. *The American*

Journal of Psychiatry, *161*(11), 2126–2128.

Nosek, B. A., & Lakens, D. (2014). Registered reports: A method to increase the credibility of published results. *Social Psychology*, *45*(3), 137–141.

Ott, A., Hugi, A., Baumgartner, M., & Chaintreau, A. (2000). Sensory investigation of yogurt flavor perception: mutual influence of volatiles and acidity. *Journal of Agricultural and Food Chemistry*, *48*(2), 441–450.

Pan, J. W., & Takahashi, K. (2007). Cerebral energetic effects of creatine supplementation in humans. *American Journal of Physiology. Regulatory, Integrative and Comparative Physiology*, *292*(4), R1745–R1750.

Papagno, C., Comi, A., Riva, M., Bizzi, A., Vernice, M., Casarotti, A., Fava, E., & Bello, L. (2017). Mapping the brain network of the phonological loop. *Human Brain Mapping*, *38*(6), 3011–3024.

Parry, D. A., Oeppen, R. S., Amin, M. S. A., & Brennan, P. A. (2018). Could exercise improve mental health and cognitive skills for surgeons and other healthcare professionals? *The British Journal of Oral & Maxillofacial Surgery*, *56*(5), 367–370.

Perrin, F., Peigneux, P., Fuchs, S., Verhaeghe, S., Laureys, S., Middleton, B., Degueldre, C., Del Fiore, G., Vandewalle, G., Balteau, E., Poirrier, R., Moreau, V., Luxen, A., Maquet, P., & Dijk, D.-J. (2004). Nonvisual responses to light exposure in the human brain during the circadian night. *Current Biology: CB*, *14*(20), 1842–1846.

Pierre-Louis, A.-M., Qamruddin, J., Espinosa, I., & Challa, S. (2010). The Malaria Control Success Story. *World Bank*. https://documents1.worldbank.org/curated/fr/304221468001788072/930107812_20140825105005822/additional/634310PUB0Yes0061512B09780821387450.pdf

Pjrek, E., Friedrich, M.-E., Cambioli, L., Dold, M., Jäger, F., Komorowski, A., Lanzenberger, R., Kasper, S., & Winkler, D. (2020). The Efficacy of Light Therapy in the Treatment of Seasonal Affective Disorder: A Meta-Analysis of Randomized Controlled Trials. *Psychotherapy and Psychosomatics*, *89*(1), 17–24.

Praschak-Rieder, N., & Willeit, M. (2003). Treatment of seasonal affective disorders.

Dialogues in Clinical Neuroscience, *5*(4), 389–398.

Qian, M., Ba, D. W., Watkins, W. E., Mstat, V. G., Yan, Y. Q., Li, M., & Pei, Z. (2005). The effects of iodine on intelligence in children: a meta- analysis of studies conducted in China. *Asia Pac J Clin Nutr*, *14*(1), 32–42.

Rae, C., Digney, A. L., McEwan, S. R., & Bates, T. C. (2003). Oral creatine monohydrate supplementation improves brain performance: a double-blind, placebo-controlled, cross-over trial. *Proceedings. Biological Sciences / The Royal Society*, *270*(1529), 2147–2150.

Rastad, C., Ulfberg, J., & Lindberg, P. (2008). Light room therapy effective in mild forms of seasonal affective disorder—A randomised controlled study. *Journal of Affective Disorders*, *108*(3), 291–296.

Rastad, C., Ulfberg, J., & Lindberg, P. (2011). Improvement in Fatigue, Sleepiness, and Health-Related Quality of Life with Bright Light Treatment in Persons with Seasonal Affective Disorder and Subsyndromal SAD. *Depression Research and Treatment*, *2011*, 543906.

Rastad, C., Wetterberg, L., & Martin, C. (2017). Patients' Experience of Winter Depression and Light Room Treatment. *Psychiatry Journal*, *2017*, 6867957.

Raven, J. (2008). *The Raven Progressive Matrices Tests: Their Theoretical Basis and Measurement Model*.

Ravindran, A. V., Lam, R. W., Filteau, M. J., Lespérance, F., Kennedy, S. H., Parikh, S. V., & Patten, S. B. (2009). Canadian Network for Mood and Anxiety Treatments (CANMAT) Clinical guidelines for the management of major depressive disorder in adults.: V. Complementary and alternative medicine treatments. *Journal of Affective Disorders*, *117*, S54–S64.

Rawson, E. S., & Venezia, A. C. (2011). Use of creatine in the elderly and evidence for effects on cognitive function in young and old. *Amino Acids*, *40*(5), 1349–1362.

Rector, N. A., Bourdeau, D., Kitchen, K., Joseph-Massiah, L., & Laposa, J. M. (2005). *Anxiety disorders: An information guide*. camh.

Reitan, R. M. (1958). Validity of the Trail Making Test as an Indicator of Organic Brain Damage. *Perceptual and Motor Skills*, *8*(3), 271–276.

Rosen, L. N., Targum, S. D., Terman, M., Bryant, M. J., Hoffman, H., Kasper, S. F., Hamovit, J. R., Docherty, J. P., Welch, B., & Rosenthal, N. E. (1990). Prevalence of seasonal affective disorder at four latitudes. *Psychiatry Research*, *31*(2), 131–144.

Roshanaei-Moghaddam, B., Pauly, M. C., Atkins, D. C., Baldwin, S. A., Stein, M. B., & Roy-Byrne, P. (2011). Relative effects of CBT and pharmacotherapy in depression versus anxiety: is medication somewhat better for depression, and CBT somewhat better for anxiety? *Depression and Anxiety*, *28*(7), 560–567.

Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence*, *53*, 118–137.

Saloojee, H., & Pettifor, J. M. (2024). Maximizing access and minimizing barriers to research in low- and middle-income countries: Open Access and health equity. *Calcified Tissue International*, *114*(2), 83–85.

Schellig, D. (1997). *Block-tapping-test*. Swets Test Services Frankfurt.

Seligman, M. E. P., Steen, T. A., Park, N., & Peterson, C. (2005). Positive psychology progress: empirical validation of interventions. *The American Psychologist*, *60*(5), 410–421.

Shrout, P. E., & Rodgers, J. L. (2018). Psychology, science, and knowledge construction: Broadening perspectives from the replication crisis. *Annual Review of Psychology*, *69*(Volume 69, 2018), 487–510.

Smith, A. L., & Tan, P. (2006). Creatine synthesis: An undergraduate organic chemistry laboratory experiment. *Journal of Chemical Education*, *83*(11), 1654.

Solis, M. Y., Artioli, G. G., Otaduy, M. C. G., Leite, C. da C., Arruda, W., Veiga, R. R., & Gualano, B. (2017). Effect of age, diet, and tissue type on PCr response to creatine supplementation. *Journal of Applied Physiology*, *123*(2), 407–414.

Soong, W. T., Chao, K. Y., Jang, C. S., & Wang, J. D. (1999). Long-term effect of increased lead absorption on intelligence of children. *Archives of Environmental Health*, *54*(4),

297–301.

Stain-Malmgren, R., Kjellman, B. F., & Aberg-Wistedt, A. (1998). Platelet serotonergic functions and light therapy in seasonal affective disorder. *Psychiatry Research*, *78*(3), 163–172.

Steenge, G. R., Simpson, E. J., & Greenhaff, P. L. (2000). Protein- and carbohydrate-induced augmentation of whole body creatine retention in humans. *Journal of Applied Physiology*, *89*(3), 1165–1171.

Sternberg, R. J., Grigorenko, E., & Bundy, D. A. (2001). The predictive value of IQ. *Merrill-Palmer Quarterly (Wayne State University. Press)*, *47*(1), 1–41.

Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence*, *35*(5), 401–426.

Terman, M., & Terman, J. S. (2005). Light therapy for seasonal and nonseasonal depression: efficacy, protocol, safety, and side effects. *CNS Spectrums*, *10*(8), 647–663; quiz 672.

Terman, Williams, White, Gould. (2008). *Selbstbewertungs-Fragebogen (SIGH-SAD-SA 2008)* (A. B. A. V. Wirz, Trans.). Center for Environmental Therapeutics. https://cet.org/wp-content/uploads/2018/01/SIGH-SAD-SA-DE.pdf

Thalén, B. E., Kjellman, B. F., Mørkrid, L., & Wetterberg, L. (1995). Melatonin in light treatment of patients with seasonal and nonseasonal depression. *Acta Psychiatrica Scandinavica*, *92*(4), 274–284.

Turner, C. E., Russell, B. R., & Gant, N. (2015). Comparative quantification of dietary supplemented neural creatine concentrations with (1)H-MRS peak fitting and basis spectrum methods. *Magnetic Resonance Imaging*, *33*(9), 1163–1167.

Vandewalle, G., Balteau, E., Phillips, C., Degueldre, C., Moreau, V., Sterpenich, V., Albouy, G., Darsaud, A., Desseilles, M., Dang-Vu, T. T., Peigneux, P., Luxen, A., Dijk, D.-J., & Maquet, P. (2006). Daytime light exposure dynamically enhances brain responses. *Current Biology: CB*, *16*(16), 1616–1621.

Vandewalle, G., Gais, S., Schabus, M., Balteau, E., Carrier, J., Darsaud, A., Sterpenich, V., Albouy, G., Dijk, D. J., & Maquet, P. (2007). Wavelength-dependent modulation of brain

responses to a working memory task by daytime light exposure. *Cerebral Cortex (New York, N.Y.: 1991)*, *17*(12), 2788–2795.

van Doorn, J., van den Bergh, D., Böhm, U., Dablander, F., Derks, K., Draws, T., Etz, A., Evans, N. J., Gronau, Q. F., Haaf, J. M., Hinne, M., Kucharský, Š., Ly, A., Marsman, M., Matzke, D., Gupta, A. R. K. N., Sarafoglou, A., Stefan, A., Voelkel, J. G., & Wagenmakers, E.-J. (2021). The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin & Review*, *28*(3), 813–826.

Van Someren, E. J., Kessler, A., Mirmiran, M., & Swaab, D. F. (1997). Indirect bright light improves circadian rest-activity rhythm disturbances in demented patients. *Biological Psychiatry*, *41*(9), 955–963.

Votruba, N., Thornicroft, G., & FundaMentalSDG Steering Group. (2016). Sustainable development goals and mental health: learnings from the contribution of the FundaMentalSDG global initiative. *Global Mental Health (Cambridge, England)*, *3*(e26), e26.

Wang, T., Li, S., Jiang, G., Lin, C., Li, M., Ma, X., Zhan, W., Fang, J., Li, L., Li, C., & Tian, J. (2016). Regional homogeneity changes in patients with primary insomnia. *European Radiology*, *26*(5), 1292–1300.

Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale*. *110*. https://psycnet.apa.org/fulltext/1955-07334-000.pdf

Weigend, S., Holst, S. C., Treyer, V., O'Gorman Tuura, R. L., Meier, J., Ametamey, S. M., Buck, A., & Landolt, H.-P. (2019). Dynamic changes in cerebral and peripheral markers of glutamatergic signaling across the human sleep-wake cycle. *Sleep*, *42*(11), zsz161.

Wilcox, R. R. (2011). *Introduction to Robust Estimation and Hypothesis Testing*. Academic Press.

Wilcox, R. R., & Rousselet, G. A. (2023). An updated guide to robust statistical methods in neuroscience. *Current Protocols*, *3*(3), e719.

Wilkinson, I. D., Mitchel, N., Breivik, S., Greenwood, P., Griffiths, P. D., Winter, E. M., & Van Beek, E. J. R. (2006). Effects of creatine supplementation on cerebral white matter in

competitive sportsmen. *Clinical Journal of Sport Medicine: Official Journal of the Canadian Academy of Sport Medicine*, *16*(1), 63–67.

Williams, M., & Penman, D. (2011). *Mindfulness: A practical guide to finding peace in a frantic world*. Piatkus Books. https://books.google.com/books?hl=de&lr=&id=A6BcAgAAQBAJ&oi=fnd&pg=PT5&dq=mark+williams+mindfulness&ots=LsAUn35b1e&sig=e2O3LZhGCejsoI1dXiphWHyML6s

Wirz-Justice, A., Ajdacic, V., Rössler, W., Steinhausen, H.-C., & Angst, J. (2019). Prevalence of seasonal depression in a prospective cohort study. *European Archives of Psychiatry and Clinical Neuroscience*, *269*(7), 833–839.

Wirz-Justice, A., Graw, P., Röösli, H., Glauser, G., & Fleischhauer, J. (1999). An open trial of light therapy in hospitalised major depression. *Journal of Affective Disorders*, *52*(1-3), 291–292.

World Health Organization. (2017). *Depression and other common mental disorders: global health estimates*. Geneva: World Health Organization.

World Health Organization. (2022). *World mental health report: transforming mental health for all*. World Health Organization.

Youngstedt, S. D., & Kripke, D. F. (2007). Does bright light have an anxiolytic effect? - an open trial. *BMC Psychiatry*, *7*(1), 62.

Zaboski, B. A., 2nd, Kranzler, J. H., & Gage, N. A. (2018). Meta-analysis of the relationship between academic achievement and broad abilities of the Cattell-horn-Carroll theory. *Journal of School Psychology*, *71*, 42–56.

Zaccaro, A., Piarulli, A., Laurino, M., Garbella, E., Menicucci, D., Neri, B., & Gemignani, A. (2018). How Breath-Control Can Change Your Life: A Systematic Review on Psycho-Physiological Correlates of Slow Breathing. *Frontiers in Human Neuroscience*, *12*, 353.

Zeidan, F., Johnson, S. K., Diamond, B. J., David, Z., & Goolkasian, P. (2010). Mindfulness meditation improves cognition: evidence of brief mental training. *Consciousness and Cognition*, *19*(2), 597–605.

Zurrin, R., Wong, S. T. S., Roes, M. M., Percival, C. M., Chinchani, A., Arreaza, L., Kusi, M., Momeni, A., Rasheed, M., Mo, Z., Goghari, V. M., & Woodward, T. S. (2024). Functional brain networks involved in the Raven's standard progressive matrices task and their relation to theories of fluid intelligence. *Intelligence*, *103*(101807), 101807.

APPENDIX A: List of Figures and Tables

Figures

Tables

APPENDIX B: List of Abbreviations

5-HTTLPR	serotonin-transporter-linked promoter region
ACT	acceptance and commitment therapy
ATP	adenosine triphosphate
AGAT enzyme deficiency	arginine:glycine amidinotransferase deficiency
ANOVA	analysis of variance
AVLT, in German: VLMT	Auditory Verbal Learning Test
BDS	Backwards Digit Span
BF	Bayes factor
BROAD light therapy	BRight, whole-ROom, All-Day light therapy
(BVMT-R)	Brief-Visuospatial-Memory Test – Revised
CI	confidence interval
CBT	cognitive behavioural therapy
CONSORT	Consolidated Standards of Reporting Trials
COVID-19	coronavirus disease of 2019
Cr	creatine
dACC	dorsal anterior cingulate cortex
DLPFC	dorsolateral prefrontal cortex
DMPFC	dorsomedial prefrontal cortex
EMMs	estimated marginal means
GAMT enzyme deficiency	guanidinoacetate methyltransferase deficiency
HARKing	Hypothesising After Results are Known
K	Kelvin
LED	light-emitting diode
MWT-B	Mehrfachwahl-Wortschatz-Intelligenztest
NHST	Null Hypothesis Significance Testing

PCr	phosphocreatine
PFC	prefrontal cortex
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RAPM	Raven's Advanced Progressive Matrices
RCT	randomised controlled trial
RR	relative risk, also: odds ratio
SAD	seasonal affective disorder
SIGH-SAD-SR	Hamilton Depression Rating Scale-Seasonal Affective Disorders 29-items Version, self-report version
STAI	State-Trait Anxiety Inventory
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
SCN	suprachiasmatic nucleus
tCr	total creatine
TMT-A	Trail-Making-Test A
TMT-B	Trail-Making-Test B
VLPFC	left ventrolateral prefrontal cortex
YLD	year of healthy life lost to disability

APPENDIX C: Publication Study 1

Sandkühler, J. F., Brochhagen, S., Rohde, P., Muscheidt, R. C., Grömer, T. W., Müller, H., & Brauner, J. (2022). 100,000 lumens to treat seasonal affective disorder: A proof of concept RCT of Bright, whole-ROom, All-Day (BROAD) light therapy. *Depression and Anxiety, 39*(12), 760–769. https://doi.org/10.1002/da.23281

Received: 3 November 2021 | Revised: 27 April 2022 | Accepted: 30 June 2022

DOI: 10.1002/da.23281

RESEARCH ARTICLE

ANXIETY AND DEPRESSION ASSOCIATION OF AMERICA WILEY

100,000 lumens to treat seasonal affective disorder: A proof of concept RCT of Bright, whole-ROom, All-Day (BROAD) light therapy

Julia F. Sandkühler[1,2] | Sarah Brochhagen[2,3] | Paul Rohde[3,4] | Rosa C. Muscheidt[2,3] | Teja W. Grömer[3,5] | Helge Müller[2,3,4] | Jan M. Brauner[2,3,6]

[1]Department of Psychology, University of Bonn, Bonn, Germany

[2]Department of Health, Witten/Herdecke University, Witten, Germany

[3]Clinic for Psychiatry and Psychotherapy, Community Hospital Herdecke, Herdecke, Germany

[4]Department of Media and Communications, London School of Economics and Political Science, London, UK

[5]Department of Psychiatry and Psychotherapy, Friedrich-Alexander University of Erlangen-Nuremberg, Erlangen, Germany

[6]Department of Computer Science, University of Oxford, Oxford, UK

Correspondence

Julia F. Sandkühler, Abteilung Allgemeine Psychologie I, Institut für Psychologie, Rheinische Friedrich-Wilhelms-Universität Bonn, Kaiser-Karl-Ring 9, D-53111 Bonn, Germany.
Email: jf.sandkuehler@gmail.com

Funding information

Centre for Effective Altruism

Abstract

Background: Seasonal affective disorder (SAD) is common and debilitating. The standard of care includes light therapy provided by a light box; however, this treatment is restrictive and only moderately effective. Advances in LED technology enable lighting solutions that emit vastly more light than traditional light boxes. Here, we assess the feasibility of BROAD (Bright, whole-ROom, All-Day) light therapy and get a first estimate for its potential effectiveness.

Methods: Patients were randomly assigned to a treatment for 4 weeks; either a very brightly illuminated room in their home for at least 6 h per day (BROAD light therapy) or 30 min in front of a standard 10,000 lux SAD light box. Feasibility was assessed by monitoring recruitment, adherence, and side effects. SAD symptoms were measured at baseline and after 2 and 4 weeks, with the Hamilton Depression Rating Scale-Seasonal Affective Disorders 29-items, self-report version.

Results: All 62 patients who started treatment were available at 4-week follow-up and no significant adverse effects were reported. SAD symptoms of both groups improved similarly and considerably, in line with previous results. Exploratory analyses indicate that a higher illuminance (lux) is associated with a larger symptom improvement in the BROAD light therapy group.

Conclusions: BROAD light therapy is feasible and seems similarly effective as the standard of care while not confining the participants to 30 min in front of a light box. In follow-up trials, BROAD light therapy could be modified for increased illuminance, which would likely improve its effectiveness.

KEYWORDS

light room, randomized controlled trial, SAD, therapy lamp, winter depression, 10000 lux

1 | INTRODUCTION

Becoming depressed every winter is a common and debilitating experience. Standard light therapy is time-intensive for the patients and only moderately effective. Is the "dose of light" used in the standard therapy simply too low?

Seasonal affective disorder (SAD) is a seasonal pattern of recurrent major depressive episodes that most commonly occurs during autumn or winter with full remission during spring and summer (Rosenthal et al., 1984). SAD is thought to be caused primarily by a lack of daylight (Levitan, 2007); The prevalence of SAD based on diagnostic interviews is estimated to be, for example, 2.9% in Toronto, Canada (Levitt et al., 2000), 2.4% in North Wales (Michalak et al., 2001), 3.4% in Zurich, Switzerland (Wirz-Justice et al., 2019), and 0.4% in the contiguous United States (i.e., excluding among other places, Alaska; Blazer et al., 1998).

Light therapy is one of the most commonly recommended treatments for SAD (Anderson et al., 2008; Bauer et al., 2013; Gelenberg et al., 2010; Kurlansik & Ibay, 2012; Mårtensson et al., 2015; Praschak-Rieder & Willeit, 2003; Ravindran et al., 2009; Terman & Terman, 2005). For a short overview of the neurobiological effects of light therapy against SAD see Pjrek et al. (2020). Light therapy is usually provided by sitting in front of a light box for 30 min in the morning, restricting the number of activities during this time. This proof of concept trial introduces a new light therapy that enables patients to go about their daily activities during treatment.

Meta-analyses showed that the standard treatment regimen is only moderately effective (Golden et al., 2005; Mårtensson et al., 2015; Pjrek et al., 2020). Since SAD patients remit in summer, the reason why the current treatment is not fully effective might be that it differs too much from summer light. The new light therapy in this trial is more similar to summer than standard light therapy in several ways.

Currently, relevant guidelines and light therapy lamp manufacturers recommend light therapy that has the following characteristics (Kurlansik & Ibay, 2012; MayoClinic, 2017; National Health Service, 2018; National Institute of Mental Health, 20-MH-8138; Philips, 2018):

- Light emitted by a SAD therapy lamp, which usually provides 10,000 lux at 20 cm distance to the eyes and only covers a small area of the visual field.
- Exposure for 30 min per day, preferably in the morning.

Alternative recommendations are 5000 lux for 1 h and 2500 lux for 2 h.

In contrast, consider the characteristics of natural sunlight on a summer's day:

- *Homogeneous illumination*: The environment is much more evenly lit, bright light reaches the eye from all of the visual field. It might be beneficial to have the whole visual field illuminated and not just the small area covered by a light box.
- *Duration*: People are exposed to bright illuminance for many hours of the day (Li et al., 2010; Matour et al., 2017).
- *Illuminance*: Sunlight in summer has an illuminance of up to 25,000 lux in the shadow[1] (Li et al., 2010; Matour et al., 2017).

Assuming that patients keep the correct distance from the lamps, the standard recommended light therapy broadly meets the illuminance (lux) of natural lighting on a summer's day. However, it covers only a small part of the visual field, and the duration is short. We aimed to make light therapy more similar to summer lighting by using "more light"—bright light that covers the whole visual field for many hours. We term this experimental treatment BROAD (Bright, whole-ROom, All-Day) light therapy.

Indeed, there are online reports from patients suffering from SAD, who have experimented with such solutions and have reported that they found them much more effective than the standard SAD light box.[2] We interviewed several of these patients, and their experiences directly informed the experimental treatment we evaluated in the trial. The setup for BROAD light therapy, as evaluated in this trial, costs around 350 euros and consists of a socket cord with up to 70 very bright LED light bulbs, which participants installed in a room in their home where they spent at least 6 h per day (Figure 1).

BROAD light therapy is related to an existing treatment in the literature—light room therapy, that is, bright light that covers the whole visual field, but which is clinic-based and with shorter treatment durations. Light rooms have been studied for the treatment of SAD, with promising initial results. Six studies on light room therapy against SAD and six studies on light room therapy against other disorders found many positive psychological and neurological effects (Canazei et al., 2017; Kripke et al., 1983, 1992; Partonen et al., 1992, 1993; Rastad et al., 2008, 2011, 2017; Stain-Malmgren et al., 1998; Thalén et al., 1995; Van Someren et al., 1997; Wirz-Justice et al., 1999). These studies gave patients exposure for up to 3 h with illuminance ranging from 1100 to 4300 lux (Canazei et al., 2017; Kripke et al., 1983, 1992; Rastad et al., 2008, 2011, 2017; Stain-Malmgren et al., 1998; Thalén et al., 1995; Van Someren et al., 1997; Wirz-Justice et al., 1999). No study compared the efficacy of a light room against the more common SAD light box treatment. In all these previous studies, the patients had to travel to a central location, usually a clinic, to spend time in a light room specifically installed for this purpose. In contrast, we brightly illuminated a room in the patient's home for 6 or more hours per day while they went about their daily activities. This new treatment is cheaper, more convenient for the patients, and enables longer exposure times.

[1]Due to their position in the eye socket, the eye balls are usually in the shadow. On surfaces with direct sun exposure, much higher illuminances of more than 100,000 lux are reached during the summer (Li et al., 2010; Matour et al., 2017).

[2]Example from a SAD patient: https://meaningness.com/metablog/sad-light-led-lux. Example from a healthy person: https://www.benkuhn.net/lux/

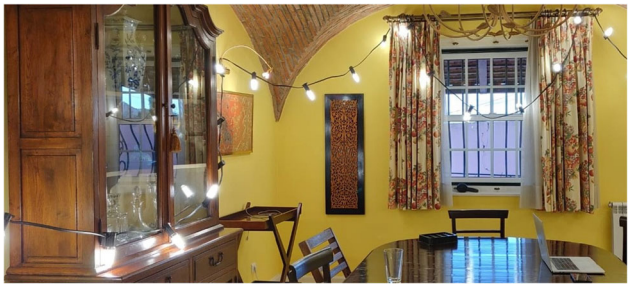


FIGURE 1 Example BROAD light therapy setup. Participants were provided with one or two such socket cords, a total of 40–70 bright LED light bulbs, and materials to attach them to their room. Only a fraction of the bulbs is shown in this picture. While it is hard to capture on camera, the BROAD setup results in very brightly illuminated rooms.

BROAD light therapy seems like a fairly natural extension of light room therapy, so one might ask why it has not been studied before. We think this might be due to past technical constraints. Producing bright light was much harder and more expensive in the past, which made the treatment tested in this trial prohibitively expensive. However, recent advances in LED technology enable affordable and powerful lighting solutions.

Developing less restrictive forms of light therapy is valuable. In addition, we hypothesize that BROAD light therapy can be made more effective at treating SAD than a SAD light box. Should our hypothesis be confirmed by our proof of concept trial or further follow-up trials, these findings would have the potential to revolutionize SAD therapy and reduce the symptom severity of millions of patients worldwide. The changes to clinical practice would be straightforward to implement: Physicians could simply recommend different lighting solutions to their patients. In this proof of concept trial, we aimed to explore the feasibility and acceptance of BROAD light therapy, get a first estimate for its potential effect to inform whether larger follow-up trials are warranted, and explore how our concrete implementation of BROAD light therapy can be improved.

2 | METHODS

This is a short summary of the study design and protocol. Further details on all sections can be found in the Supporting Information: Appendix.

2.1 | Study design

We did a randomized, partially blind (see section on blinding), parallel-group, proof of concept trial with a primary endpoint at 4 weeks from randomization. The trial evaluated a BROAD light therapy compared with an active control—the standard of care (a SAD light box). The trial design and participant flow are summarized in Figure 2. The trial was prospectively registered (drks.de identifier: DRKS00023075; https://osf.io/ndjm4/) and ethical approval was obtained from the ethics committee of Witten/Herdecke University (184/2020). Apart from telephone calls and the intervention itself, all aspects of the study, including screening, consent, diagnosis video calls, treatment allocation, and symptom assessment were conducted online. Participants received the lamps via mail and set them up at home.

2.2 | Participants

Participants were recruited through general practitioners, clinics, and social media between 29/09/2020 and 29/01/2021 and treated between 17/11/2020 and 17/03/2021. A screening questionnaire briefly assessed if the eligibility criteria were met. Participants who met these criteria received the Hamilton Depression Rating Scale-Seasonal Affective Disorders 29-items Version, self-report version (SIGH-SAD-SR) to assess their SAD symptom severity for screening and randomization purposes. If participants' symptoms were not yet severe enough, they were asked to answer the screening SIGH-SAD-SR again if they started to feel more depressed.

Participants were 18 years or older with a diagnosis of "major depressive disorder (recurrent episode) with a winter seasonal pattern," as assessed by a structured clinical interview based on the DSM-5 (American Psychiatric Association, 2013), which was conducted by psychologists in a video call. Patients were eligible to participate in the trial if they were at home for at least 6 h in morning and afternoon (before 19:00), on at least 5 days per week. Exclusion criteria included a history of manic episodes, light therapy in the previous 4 months, and recent changes in antidepressant medication (see Supporting Information: Appendix for full list).

At first, the screening information participants received did not adequately convey that they might have to put a long socket cord into their room, which led to some participants in the BROAD light therapy group withdrawing before starting the treatment. After we adjusted the recruitment and screening material, there were no more withdrawals.

2.3 | Interventions

2.3.1 | Active control group

The active control group received a commercially available SAD light box with a color temperature of 6500K and a color rendering index of 80 which provided 10,000 lux at 20 cm. With greater distance, the illuminance rapidly declined. The lamps were purchased from various manufacturers according to availability and were most commonly the Beurer TL 41 or equivalent lamps. Participants were asked to sit in front of this lamp in the morning at an eye distance of 20 cm while reading or having breakfast. They were asked not to look directly into it apart from occasional squints. The treatment duration was 30min at least 5 days a week for at least 4 weeks.

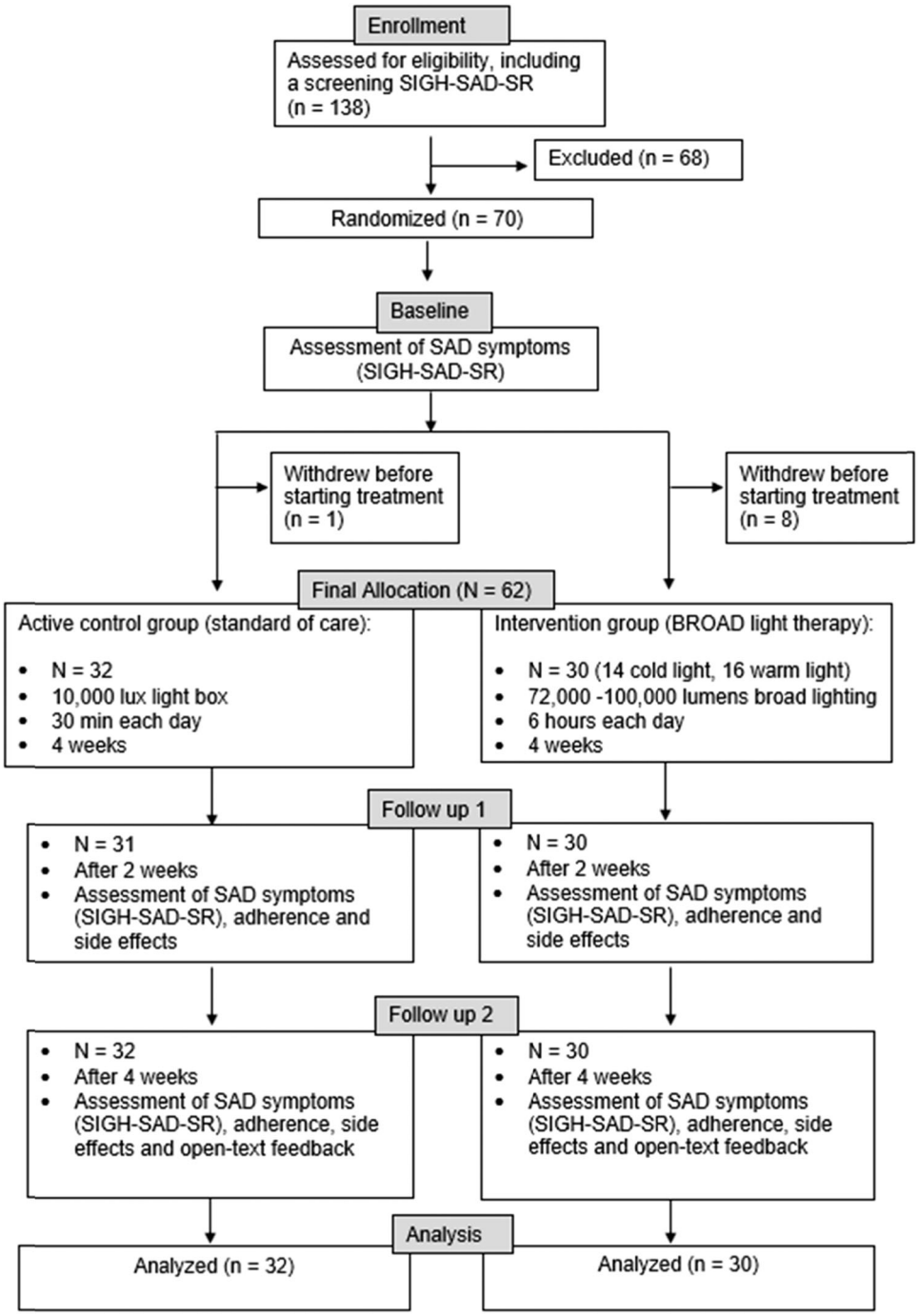


FIGURE 2 Participant flow through the study

2.3.2 | Intervention group

There were two BROAD light therapy conditions: a warm condition with a color temperature of 4000K and a cold condition with a color temperature of 6000K. Both had a color rendering index of 80–90. We included these two versions to investigate if there were large differences between them in how acceptable BROAD light therapy is to patients, as warm light is sometimes perceived as more comfortable.

The light in the BROAD light therapy was provided by one or two 10- to 20-m-long socket cords with 20–40 light bulbs each (see Figure 1). The light emitted in the cold condition was 72,000 lumens and 102,240 lumens in the warm condition, so the amount of blue light was the same in both. This is important because blue light is hypothesized to be the effective wavelength in SAD treatment (Strong et al., 2009). The mean illuminance at eye level when sitting was 1433 lux (minimum 550 lux, maximum 4061 lux) in the cold light condition and 1829 lux (minimum 500 lux, maximum 6800 lux) in the

warm light condition, as measured by the participants with a smartphone app. Despite the BROAD light therapy bulbs emitting vastly more light than the light box (72,000–102,000 vs. 850–1530 lumens), the average amount of light that reached the eye at any given point in time (illuminance) was lower for the BROAD light therapy than the light boxes (1433–1829 vs 10,000 lux). This was because the light bulbs in the BROAD light therapy setup were much further away from the eyes than the SAD light box (20 cm). The larger distance had the advantage that patients were able to go about their daily activities, enabling a longer treatment duration and more overall illuminance (lux) over time. The treatment duration was at least 6 h on at least 5 days a week for at least 4 weeks. Participants had the option to diffuse the light with paper lanterns but only very few participants chose to do this.

2.3.3 | Similarity of treatment groups

Apart from the inevitable differences in the nature of the treatments, the treatment groups were made as similar as possible, receiving the same questionnaires, checks on the setup, and reminders.

2.4 | Outcomes

To determine the feasibility of a larger study, we assessed recruitment, adherence, side effects, and participant feedback.

To get a first estimate of potential effectiveness, we had one preregistered primary outcome: Difference in SAD symptom severity between baseline and after 4 weeks of treatment. This outcome was recorded online using the Hamilton Depression Rating Scale-Seasonal Affective Disorders 29-items Version, self-report version (SIGH-SAD-SR) (Terman & White, 2008). Our three preregistered secondary outcomes were the difference in SAD symptom severity between baseline and after 2 weeks of treatment; and the fraction of patients in remission after 2 and 4 weeks.

2.5 | Randomization and blinding

Patients were randomly assigned (1:1:2) to receive either warm white (4000K) or cold white (6000K) BROAD light therapy or the active control procedure. Randomization was performed using a computer-generated list, in blocks of four and eight participants and in three screening symptom severity strata (SIGH-SAD score ≤ 19, 20–29, and ≥30) by the senior author, who did not have contact with or knowledge of participants. The senior author received an email with the participant code and symptom severity and emailed back the allocation.

Apart from occasional corrections to the setup of the lamps and answers to participants' questions, unblinded personnel did not interact with participants. In particular, the outcome surveys were online and included no interaction with the study personnel. Participants did not know which treatment was the active control group nor what all the treatment options were. Participants did not know that there were three groups, or what the difference in treatment between the groups was. However, the lamps provided in the different groups looked different.

2.6 | Statistical analysis

The feasibility outcomes were reported descriptively and narratively. In accordance with the CONSORT guideline on pilot and feasibility trials (Eldridge et al., 2016), for the clinical primary and secondary endpoints, we report only descriptive statistics: mean (standard deviation) for continuous outcomes and raw count (%) for categorical outcomes. In an exploratory fashion, we additionally report descriptives of the primary endpoint adjusted for confounders (for details see Supporting Information: Appendix). In addition, we performed exploratory Pearson correlations between illuminance in the BROAD light therapy condition and the primary endpoint and between measures of adherence and the primary endpoint. The significance level for these tests of correlation was $p < .05$.

We would have included participants who had dropped out after starting the treatment, but there were none.

3 | RESULTS

We analyzed all 62 patients who started the treatment. For participant characteristics, see Table 1.

3.1 | Feasibility

The recruitment of participants went better than according to plan. The process we had set up to run the study worked and we see no obstacle here for a larger study.

Participants adhered very well to their instructed treatment durations. Both groups underwent treatment for more than the minimum required 5 days a week on average (see Table 2). As instructed, participants in the intervention group underwent treatment for over 6 h per day on average, while active control participants underwent treatment for a little over half an hour per day on average (see Table 2). Participants in both groups started the treatment each day at the mean and median time of 9 a.m.

No patients discontinued the study due to an adverse event. Some participants (eight in the intervention group and four in the active control group) reported side effects of a nature consistent with other studies on light therapy (see Table 3).

Qualitative and free-text feedback for both the intervention group and the active control group was generally positive. On a scale from 1 = Completely disagree to 7 = Completely agree, participants in both groups on average somewhat agreed that the treatment had helped them with their symptoms (see Table 4).

TABLE 1 Baseline characteristics

	Intervention	Active control	Total
N	30	32	62
Age in years (M, SD)	38.8 (14.5)	38.5 (15.5)	38.6 (14.9)
Female (*N*)	19	22	41
Male (*N*)	11	10	21
Baseline symptom severity in SIGH-SAD points (M, SD)	26.3 (6.09)	27.9 (8.95)	27.1 (7.68)

Note: Data are given as mean (standard deviation) or as number of participants.

TABLE 2 Adherence

	BROAD light therapy	Light box therapy
Days treated per week	5.72 (0.63)	5.88 (0.79)
Hours treated on treatment days	6.75 (1.28)	0.6 (0.31)

Note: Data are given as mean (standard deviation).

TABLE 3 Side effects

	BROAD light therapy	Light box therapy
Any side effect	8 (27%)	4 (13%)
Eye ache	2 (7%)	2 (6%)
Headache	5 (17%)	1 (3%)
Agitation	2 (7%)	1 (3%)

Note: Data are given as *N* (%).

3.2 | Effectiveness of BROAD light therapy

This is a proof of concept trial and it is not powered to detect moderate effect size differences between the two treatments. In accordance with the CONSORT guideline on pilot and feasibility trials (Eldridge et al., 2016), we therefore report only descriptive statistics and no tests of significance of the differences.

Our primary endpoint, the improvement in symptoms after 4 weeks of treatment, was considerable and similar for all treatment groups (see Table 5 and Figure 3). The magnitude of SAD symptom improvement after light therapy is in line with previous results (Pjrek et al., 2020). The same was true for our secondary endpoints, symptom improvement after 2 weeks of treatment, and remission (defined as a SIGH-SAD score below 9) at 2 and 4 weeks (see Table 5). The improvement in symptoms was similar for the cold and the warm color temperature BROAD therapy groups (see Supporting Information: Appendix).

When adjusting for baseline symptom severity, timepoint of treatment and delay between baseline assessment and start of treatment, the improvement after 4 weeks was even more similar between treatments (see Table 5 and Supporting Information: Appendix). The SIGH-SAD consists of two subscales, measuring typical and atypical symptoms of major depression respectively. The improvements in both scales were very similar for both treatments.

3.3 | More light, better results

In the intervention group, the illuminance (lux) at eye level depended on the room sizes and how/where the light bulbs were attached. Accordingly, the illuminance at eye level varied greatly between participants in the BROAD light therapy group (between 500 and 6800 lux). The illuminance was lower than for the active control group under ideal conditions, because the BROAD light therapy bulbs were further away from the eyes (see Section 2.3). This allowed participants to go about their daily activities and thus enabled a longer treatment duration per day and more total lux. The higher the illuminance in the BROAD light therapy group, the more participants' symptoms improved after 4 weeks, $r(28) = .43$, $p = .032$ (see Table 6). As shown in Figure 4, every doubling of illuminance (lux) in the intervention group predicted an increase in symptom improvement by 4 SIGH-SAD points on average. To further display this trend, we can split the participants that received BROAD light therapy into two groups: Participants who received above median illuminance at eye levels (>1400 lux) experienced a mean improvement of 16.7 points on the SIGH-SAD score (SD = 6.5), which is more than in the active control group (which improved by $M = 13.24$, SD = 9.64), while the participants with below median illuminance improved by only 8.9 points on average (SD = 8.5). In addition, treatment hours per day correlated positively and bordering on significance with symptom improvement for the BROAD light therapy group. However, days treated per week did not (see Table 6).

4 | DISCUSSION

This is the first study on the feasibility of BROAD light therapy (Bright, whole-ROom, All-Day light therapy). This is also the direct first comparison of any light-room treatment (whether clinic-based or at home, for 2 h or all day) with the current standard of care, a SAD light box.

TABLE 4 Participant feedback on their treatment

	BROAD light therapy	Light box therapy
"I think the light therapy in this study helped me with my symptoms."	5.33 (1.45)	5.86 (0.99)
"I would recommend this treatment to a friend with similar symptoms to mine."	5.47 (1.46)	5.90 (1.23)

Note: Data are given as mean (standard deviation). Answer options ranged from 1 = Completely disagree to 7 = Completely agree.

TABLE 5 Secondary endpoints

	BROAD light therapy	Light box therapy
Symptom improvement after 4 weeks	11.58 (8.39), [8.45; 14.72]	13.24 (9.64), [9.76; 16.72]
Symptom improvement after 4 weeks, adjusted (arbitrary units)	12.56 (7.66), [9.70; 15.42]	12.92 (9.19), [9.61; 16.23]
Subscale A symptom improvement after 4 weeks	7.22 (5.32), [5.23; 9.20]	8.21 (6.55), [5.85; 10.57]
Subscale B symptom improvement after 4 weeks	4.37 (4.49), [2.69; 6.04]	5.03 (4.32), [3.47; 6.59]
Symptom improvement after 2 weeks	8.92 (8.75), [5.65; 12.18]	10.8 (8.16), [7.80; 13.79]
Remission after 2 weeks	3 (10%)	3 (10%)
Remission after 4 weeks	7 (22%)	6 (20%)

Note: Mean (standard deviation) [95% confidence intervals of the mean improvement] in SIGH-SAD points given for unadjusted symptom improvement and *N* (%) given for remission. The adjusted symptom improvement does not correspond to SIGH-SAD points, but only serves as a unitless measure of similarity.

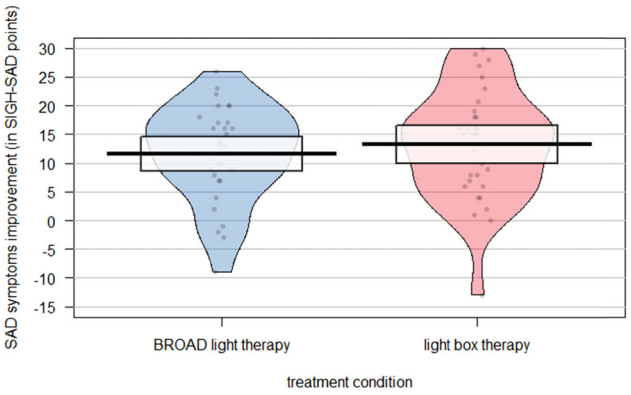


FIGURE 3 Improvement in SIGH-SAD score after 4 weeks of treatment. The plot shows the means, 95% confidence intervals of the means and distribution for both the light box group and the BROAD group.

TABLE 6 Correlations with symptom improvement 4 weeks after the start of treatment for the BROAD light therapy group

	Pearson's *r*	*p* Value
Log_2(lux) and symptom improvement	.43	.032*
Hours treated per day and symptom improvement	.34	.070
Days treated per week and symptom improvement	.05	.801

*Significant at $p < .05$.

In this proof of concept trial, treatments were roughly equally effective on average. The improvement in symptom severity (measured by the SIGH-SAD-SR score) in our study was similar to those of other light therapy studies (Pjrek et al., 2020). Could non-inferiority of BROAD light therapy treatment to standard of care be confirmed in a larger trial, this would be an exciting result for patients. To receive 10,000 lux, patients need to sit very close to the light box (distance from eyes to light box of 20 cm is common for commercial light boxes), meaning there is half an hour each morning in which they can do little else. BROAD light therapy puts no such constraint on patients' daily activities, possibly improving long-term adherence. To illustrate: In the course of the 4 weeks of therapy in this study, control participants spent 14 h in front of the light box. For the same reason, BROAD light therapy could be a good candidate for prevention, which many patients are likely to only do if it is very low effort. Additionally, we carefully instructed participants on how to use the light box to receive 10,000 lux—in non-study conditions, many participants will likely position themselves further away from the light box than 20 cm, thus receiving a lower illuminance. Again, there is no analogous concern for BROAD light therapy.

Furthermore, we have reason to believe that the BROAD light therapy treatment in this study can be modified to improve its effectiveness. In the BROAD light therapy group, the illuminance (lux) at eye level depended on the room sizes and how/where the lights were attached. Accordingly, the illuminance at eye level varied greatly between participants in this group (between 500 and 6800 lux). Exploratory analyses found an interesting significant correlation: The more light reached the participants' eyes, the more participants' symptoms improved after 4 weeks. Participants who received above median illuminance at eye levels experienced a larger improvement than in the active control group, while the participants with below median

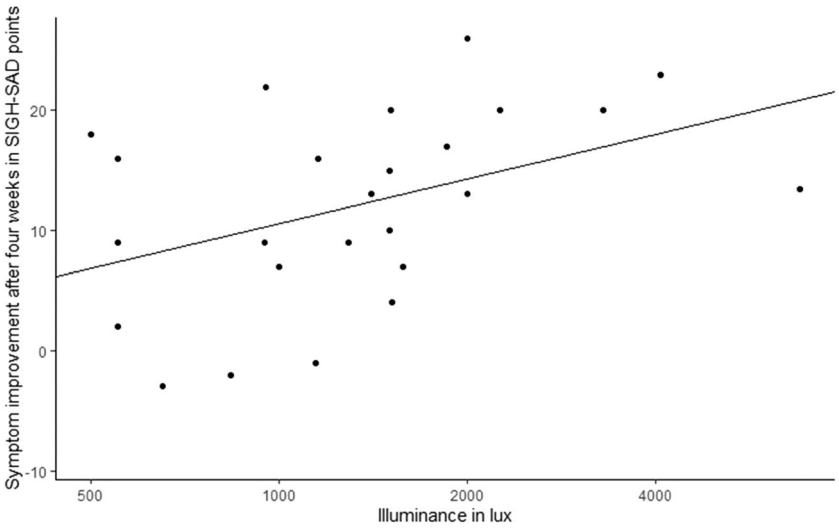


FIGURE 4 Correlation between symptom improvement after 4 weeks and the illuminance in the BROAD light therapy group in lux. Participants who received more lux had better outcomes. Scatter plot with trendline.

illuminance improved by less than the active control group. Additionally, undergoing BROAD light therapy for more hours per day was borderline significantly correlated with treatment success—however, here the causation may be the other way around, with participants who experienced the treatment to be more effective using it more regularly or for longer periods. These results offer some support for the idea behind our study, that "more light" leads to better results for patients.

4.1 | Limitations

As expected, the variance in our proof of concept trial was too large given our sample size to give much indication of whether one of the treatments performs better than the others. Additionally, illuminance levels in the BROAD light therapy group were likely too low for some participants, so this treatment was not as effective as it could be.

In addition, the light boxes were marketed on their packaging for the use against SAD, while the components of the BROAD light therapy were not, possibly leading to a higher placebo effect in the active control group than in the BROAD light therapy group. In a follow-up study, this limitation can be addressed by unifying the packaging that lamps arrive in.

Finally, we did not include a placebo or waitlist control group, so we cannot be certain that the improvement in our participants was due to the treatments. However, the effectiveness of the active control treatment has been well-studied and the size of the improvement in our study is consistent with previous findings.

4.2 | Conclusion

It is plausible that increasing the amount of light in light therapy makes it more effective against SAD. Our proof of concept trial did not find BROAD light therapy to be superior to light boxes, but it adds some first empirical evidence that BROAD light therapy is feasible and that the amount of light that reaches participants' eyes (illuminance/lux) predicts how effective it is at treating SAD. There are several straightforward ways to further increase the brightness of the treatment studied in this paper, potentially creating a treatment that is superior to the current standard of care.

ACKNOWLEDGMENTS

The authors are grateful to Eliezer Yudkowsky, as the authors took the inspiration for this trial from his book "Inadequate Equilibria," and the author helped with raising funding and contributed ideas to the initial study design. The authors thank Johann-Lukas Voigt and Simon Grosse-Holz for their help with calculating the fraction of blue light in LEDs of various color temperatures. The authors are grateful to Prof. Ulrich Ettinger, Prof. Markus Canazei, Prof. Daniel Kripke, and to the reviewer of our paper for their helpful feedback. Funding was provided by the non-profit organization Centre for Effective Altruism, 2443 Fillmore St., #380-16662, San Francisco, CA 94115, and by one private donor. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data of this study are openly available at the Open Science Framework, https://osf.io/ndjm4/.

ORCID

Julia F. Sandkühler https://orcid.org/0000-0002-5585-9539
Paul Rohde https://orcid.org/0000-0002-9156-7998

REFERENCES

Anderson, I. M., Ferrier, I. N., Baldwin, R. C., Cowen, P. J., Howard, L., Lewis, G., Matthews, K., McAllister-Williams, R. H., Peveler, R. C., Scott, J., & Tylee, A. (2008). Evidence-based guidelines for treating depressive disorders with antidepressants: A revision of the 2000 British Association for Psychopharmacology guidelines. *Journal of Psychopharmacology*, *22*(4), 343–396. https://doi.org/10.1177/0269881107088441

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (Vol. 5).

Bauer, M., Pfennig, A., Severus, E., Whybrow, P. C., Angst, J., Moeller, H. -J., & World Federation of Societies of Biological Psychiatry. Task Force on Unipolar Depressive Disorders. (2013). World Federation of Societies of Biological Psychiatry (WFSBP) guidelines for biological treatment of unipolar depressive disorders, part 1: Update 2013 on the acute and continuation treatment of unipolar depressive disorders. *The World Journal of Biological Psychiatry*, *14*(5), 334–385. https://www.tandfonline.com/doi/abs/10.3109/15622975.2013.804195

Blazer, D. G., Kessler, R. C., & Swartz, M. S. (1998). Epidemiology of recurrent major and minor depression with a seasonal pattern. The National Comorbidity Survey. *The British Journal of Psychiatry: The Journal of Mental Science*, *172*, 164–167. https://doi.org/10.1192/bjp.172.2.164

Canazei, M., Pohl, W., Bauernhofer, K., Papousek, I., Lackner, H. K., Bliem, H. R., Marksteiner, J., & Weiss, E. M. (2017). Psychophysiological effects of a single, short, and moderately bright room light exposure on mildly depressed geriatric inpatients: A pilot study. *Gerontology*, *63*(4), 308–317. https://doi.org/10.1159/000455231

Eldridge, S. M., Chan, C. L., Campbell, M. J., Bond, C. M., Hopewell, S., Thabane, L., Lancaster, G. A., & PAFS Consensus Group. (2016). CONSORT 2010 statement: Extension to randomised pilot and feasibility trials. *Pilot and Feasibility Studies*, *2*, 64. https://doi.org/10.1186/s40814-016-0105-8

Gelenberg, A. J., Freeman, M. P., Markowitz, J. C., Rosenbaum, J. F., Thase, M. E., Trivedi, M. H., & Van Rhoads, R. S. (2010). Practice guideline for the treatment of patients with major depressive disorder (3rd). *American Journal of Psychiatry*, *157*(4), 1–45.

Golden, R. N., Gaynes, B. N., Ekstrom, R. D., Hamer, R. M., Jacobsen, F. M., Suppes, T., Wisner, K. L., & Nemeroff, C. B. (2005). The efficacy of light therapy in the treatment of mood disorders: A review and meta-analysis of the evidence. *The American Journal of Psychiatry*, *162*(4), 656–662. https://doi.org/10.1176/appi.ajp.162.4.656

Kripke, D. F., Risch, S. C., & Janowsky, D. (1983). Bright white light alleviates depression. *Psychiatry Research*, *10*(2), 105–112. https://doi.org/10.1016/0165-1781(83)90109-9

Kripke, D. F., Mullaney, D. J., Klauber, M. R., Risch, S. C., & Gillin, J. C. (1992). Controlled trial of bright light for nonseasonal major depressive disorders. *Biological Psychiatry*, *31*(2), 119–134. https://doi.org/10.1016/0006-3223(92)90199-a

Kurlansik, S. L., & Ibay, A. D. (2012). Seasonal affective disorder. *American Family Physician*, *86*(11), 1037–1041. https://www.ncbi.nlm.nih.gov/pubmed/23198671

Levitan, R. D. (2007). The chronobiology and neurobiology of winter seasonal affective disorder. *Dialogues in Clinical Neuroscience*, *9*(3), 315–324. https://www.ncbi.nlm.nih.gov/pubmed/17969868

Levitt, A. J., Boyle, M. H., Joffe, R. T., & Baumal, Z. (2000). Estimated prevalence of the seasonal subtype of major depression in a Canadian community sample. *Canadian Journal of Psychiatry. Revue Canadienne de Psychiatrie*, *45*(7), 650–654. https://doi.org/10.1177/070674370004500708

Li, D. H. W., Cheung, G. H. W., Cheung, K. L., & Lam, T. N. T. (2010). Determination of vertical daylight illuminance under non-overcast sky conditions. *Building and Environment*, *45*(2), 498–508. https://doi.org/10.1016/j.buildenv.2009.07.008

Mårtensson, B., Pettersson, A., Berglund, L., & Ekselius, L. (2015). Bright white light therapy in depression: A critical review of the evidence. *Journal of Affective Disorders*, *182*, 1–7. https://doi.org/10.1016/j.jad.2015.04.013

Matour, S., Mahdavinejad, M., & Fayaz, R. (2017). Horizontal and vertical illuminance correlations in a case study for shaded south facing surfaces. *International Journal of Energy and Environmental Engineering*, *11*(12), 1046–1051. https://waset.org/Publications/horizontal-and-vertical-illuminance-correlations-in-a-case-study-for-shaded-south-facing-surfaces/10008302

MayoClinic. (2017, February 8). *Light therapy*. https://www.mayoclinic.org/tests-procedures/light-therapy/about/pac-20384604

Michalak, E. E., Wilkinson, C., Dowrick, C., & Wilkinson, G. (2001). Seasonal affective disorder: Prevalence, detection and current treatment in North Wales. *The British Journal of Psychiatry: The Journal of Mental Science*, *179*, 31–34. https://doi.org/10.1192/bjp.179.1.31

National Health Service. (2018, July 30). *Seasonal affective disorder (SAD) – Treatment*. Nhs.uk. https://www.nhs.uk/conditions/seasonal-affective-disorder-sad/treatment/

National Institute of Mental Health (20-MH-8138). *Seasonal affective disorder*. U.S. Department of Health and Human Services, National Institutes of Health, National Institute of Mental Health. https://www.nimh.nih.gov/health/topics/seasonal-affective-disorder/index.shtml

Partonen, T., Appelberg, B., & Partinen, M. (1993). Effects of light treatment on sleep structure in seasonal affective disorder. *European Archives of Psychiatry and Clinical Neuroscience*, *242*(5), 310–313. https://doi.org/10.1007/BF02190391

Partonen, T., Appelberg, B., Kajaste, S., Partinen, M., Härmä, M., & Laitinen, J. (1992). Effects of light treatment on circadian rhythmicity in seasonal affective disorder. *European Psychiatry: The Journal of the Association of European Psychiatrists*, *7*(3), 141–142. https://doi.org/10.1017/S0924933800003084

Philips. (2018, June 30). *Philips EnergyUp*. https://www.philips.de/c-p/HF3419_02/energyup-energylight

Pjrek, E., Friedrich, M. -E., Cambioli, L., Dold, M., Jäger, F., Komorowski, A., Lanzenberger, R., Kasper, S., & Winkler, D. (2020). The efficacy of light therapy in the treatment of seasonal affective disorder: A meta-analysis of randomized controlled trials. *Psychotherapy and Psychosomatics*, *89*(1), 17–24. https://doi.org/10.1159/000502891

Praschak-Rieder, N., & Willeit, M. (2003). Treatment of seasonal affective disorders. *Dialogues in Clinical Neuroscience*, *5*(4), 389–398. https://www.ncbi.nlm.nih.gov/pubmed/22033639

Rastad, C., Ulfberg, J., & Lindberg, P. (2008). Light room therapy effective in mild forms of seasonal affective disorder—A randomised controlled study. *Journal of Affective Disorders*, *108*(3), 291–296. https://doi.org/10.1016/j.jad.2007.10.009

Rastad, C., Ulfberg, J., & Lindberg, P. (2011). Improvement in fatigue, sleepiness, and health-related quality of life with bright light treatment in persons with seasonal affective disorder and subsyndromal SAD. *Depression Research and Treatment*, *2011*, 543906. https://doi.org/10.1155/2011/543906

Rastad, C., Wetterberg, L., & Martin, C. (2017). Patients' experience of winter depression and light room treatment. *Psychiatry Journal*, *2017*, 6867957. https://doi.org/10.1155/2017/6867957

Ravindran, A. V., Lam, R. W., Filteau, M. J., Lespérance, F., Kennedy, S. H., Parikh, S. V., & Patten, S. B. (2009). Canadian Network for Mood and Anxiety Treatments (CANMAT) clinical guidelines for the management of major depressive disorder in adults.: V. Complementary and alternative medicine treatments. *Journal of Affective Disorders*, *117*, S54–S64. https://doi.org/10.1016/j.jad.2009.06.040

Rosenthal, N. E., Sack, D. A., Gillin, J. C., Lewy, A. J., Goodwin, F. K., Davenport, Y., Mueller, P. S., Newsome, D. A., & Wehr, T. A. (1984). Seasonal affective disorder. A description of the syndrome and preliminary findings with light therapy. *Archives of General Psychiatry*, *41*(1), 72–80. https://doi.org/10.1001/archpsyc.1984.01790120076010

Stain-Malmgren, R., Kjellman, B. F., & Aberg-Wistedt, A. (1998). Platelet serotonergic functions and light therapy in seasonal affective disorder. *Psychiatry Research*, *78*(3), 163–172. https://doi.org/10.1016/s0165-1781(98)00017-1

Strong, R. E., Marchant, B. K., Reimherr, F. W., Williams, E., Soni, P., & Mestas, R. (2009). Narrow-band blue-light treatment of seasonal

affective disorder in adults and the influence of additional nonseasonal symptoms. *Depression and Anxiety*, *26*(3), 273–278. https://doi.org/10.1002/da.20538

Terman, M., & Terman, J. S. (2005). Light therapy for seasonal and nonseasonal depression: Efficacy, protocol, safety, and side effects. *CNS Spectrums*, *10*(8), 647–663. https://www.ncbi.nlm.nih.gov/pubmed/16041296

Terman, W., & White, G. (2008). *Selbstbewertungs-Fragebogen (SIGH-SAD-SA 2008)* (A. B. A. V. Wirz, Trans.). Center for Environmental Therapeutics. https://cet.org/wp-content/uploads/2018/01/SIGH-SAD-SA-DE.pdf

Thalén, B. E., Kjellman, B. F., Mørkrid, L., & Wetterberg, L. (1995). Melatonin in light treatment of patients with seasonal and nonseasonal depression. *Acta Psychiatrica Scandinavica*, *92*(4), 274–284. https://doi.org/10.1111/j.1600-0447.1995.tb09583.x

Van Someren, E. J., Kessler, A., Mirmiran, M., & Swaab, D. F. (1997). Indirect bright light improves circadian rest-activity rhythm disturbances in demented patients. *Biological Psychiatry*, *41*(9), 955–963. https://doi.org/10.1016/S0006-3223(97)89928-3

Wirz-Justice, A., Graw, P., Röösli, H., Glauser, G., & Fleischhauer, J. (1999). An open trial of light therapy in hospitalised major depression. *Journal of Affective Disorders*, *52*(1–3), 291–292. https://doi.org/10.1016/s0165-0327(98)00128-1

Wirz-Justice, A., Ajdacic, V., Rössler, W., Steinhausen, H. -C., & Angst, J. (2019). Prevalence of seasonal depression in a prospective cohort study. *European Archives of Psychiatry and Clinical Neuroscience*, *269*(7), 833–839. https://doi.org/10.1007/s00406-018-0921-3

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Sandkühler, J. F., Brochhagen, S., Rohde, P., Muscheidt, R. C., Grömer, T. W., Müller, H., & Brauner, J. M. (2022). 100,000 lumens to treat seasonal affective disorder: A proof of concept RCT of Bright, whole-ROom, All-Day (BROAD) light therapy. *Depression and Anxiety*, 39, 760–769. https://doi.org/10.1002/da.23281

APPENDIX D: Publication Study 2

Sandkühler, J. F., Kersting, X., Faust, A., Königs, E. K., Altman, G., Ettinger, U., Lux, S., Philipsen, A., Müller, H., & Brauner, J. (2023). The effect of creatine supplementation on cognitive performance—a randomised controlled study. *BMC Medicine*. https://doi.org/10.1186/s12916-023-03146-5

BMC Medicine

RESEARCH ARTICLE

Open Access

The effects of creatine supplementation on cognitive performance—a randomised controlled study

Julia Fabienne Sandkühler[1,2*], Xenia Kersting[2,3], Annika Faust[2], Eva Kathrin Königs[2], George Altman[4], Ulrich Ettinger[1], Silke Lux[2], Alexandra Philipsen[2], Helge Müller[2,5] and Jan Brauner[5,6,7]

Abstract

Background Creatine is an organic compound that facilitates the recycling of energy-providing adenosine triphosphate (ATP) in muscle and brain tissue. It is a safe, well-studied supplement for strength training. Previous studies have shown that supplementation increases brain creatine levels, which might increase cognitive performance. The results of studies that have tested cognitive performance differ greatly, possibly due to different populations, supplementation regimens, and cognitive tasks. This is the largest study on the effect of creatine supplementation on cognitive performance to date.

Methods Our trial was preregistered, cross-over, double-blind, placebo-controlled, and randomised, with daily supplementation of 5 g for 6 weeks each. We tested participants on Raven's Advanced Progressive Matrices (RAPM) and on the Backward Digit Span (BDS). In addition, we included eight exploratory cognitive tests. About half of our 123 participants were vegetarians and half were omnivores.

Results Bayesian evidence supported a small beneficial effect of creatine. The creatine effect bordered significance for BDS ($p = 0.064$, $\eta^2_p = 0.029$) but not RAPM ($p = 0.327$, $\eta^2_p = 0.008$). There was no indication that creatine improved the performance of our exploratory cognitive tasks. Side effects were reported significantly more often for creatine than for placebo supplementation ($p = 0.002$, RR = 4.25). Vegetarians did not benefit more from creatine than omnivores.

Conclusions Our study, in combination with the literature, implies that creatine might have a small beneficial effect. Larger studies are needed to confirm or rule out this effect. Given the safety and broad availability of creatine, this is well worth investigating; a small effect could have large benefits when scaled over time and over many people.

Trial registration The trial was prospectively registered (drks.de identifier: DRKS00017250, https://osf.io/xpwkc/).

Keywords Creatine, Cognition, Intelligence, Cognitive performance, Raven's Advanced Progressive Matrices, Backward Digit Span, Working memory, Deductive reasoning, Randomised controlled trial, RCT

*Correspondence:
Julia Fabienne Sandkühler
jf.sandkuehler@gmail.com
Full list of author information is available at the end of the article

Background

Given the important role cognition plays in daily life, enhancing cognition safely and cheaply is highly desirable. Creatine is safe, well-tolerated, and cheap [1]. Strength athletes have benefited from creatine supplementation for over 30 years [2, 3]. Slight weight gain due to water retention is the only consistently reported side effect [1, 4–6].

While the safety and athletic benefits of creatine are well established, its potential cognitive benefits are still unclear. A systematic review tentatively suggests that creatine supplementation may improve "short-term memory"/working memory and "intelligence/reasoning" in healthy individuals [7]. The few studies that have tested this have had heterogeneous results, but they have also used very different populations (such as vegetarians, omnivores, varying age groups), supplementation doses and durations, and cognitive tasks (including different kinds of memory, reaction time, reasoning, inhibitory control, attention, and task switching). The study with the largest effect, Rae et al. [8], tested the effect of creatine supplementation in 45 young vegetarian adults on working memory and abstract reasoning using the Backwards Digit Span (BDS) and Raven's Advanced Progressive Matrices (RAPM), respectively. Their study was placebo-controlled, randomised, and double-blind. Rae et al. [8] found creatine supplementation had a large and highly significant (i.e. $p < 0.001$) positive effect on both tasks. We deemed this study particularly worth replicating.

Supplementing creatine may benefit cognition as muscle and brain cells use creatine to access more energy when demand is high. They store creatine as phosphocreatine, which acts to regenerate the energy-providing adenosine triphosphate (ATP) [9, 10]. The energy demand of neurons can increase rapidly; maintaining ATP concentration despite increased demand may explain the potential effect of creatine intake on cognition [11]. The crucial role of creatine in brain metabolism is supported by evidence from Cerebral Creatine Deficiency Syndromes. Conditions causing brain creatine deficiency result in profound intellectual disability which can be reversed by creatine supplementation [12].

Dietary creatine is primarily contained in meat, fish, and a small amount in some dairy products [13, 14]. However, typical supplementation doses of creatine (5 g per day) are equivalent to more than 1 kg of meat consumption per day [13], which is substantially higher than the combined dietary intake and synthesis in most people [13]. So, one might expect creatine supplementation to make a difference despite creatine being produced by the body and being present in common foods.

Creatine intake increases the level of creatine in the blood serum [15, 16]. Crucially, Dechent et al. [17] found brain creatine increased by 8.7% following a 20 g/day 4-week supplementation regime; two further studies have confirmed varying supplementation regimes can increase brain creatine [18, 19] (however, see [20–22]).

It is unclear if creatine supplementation has similar effects on omnivores and vegetarians. Rae et al. [8] only included vegetarians. Another study comparing memory improvement under creatine supplementation in omnivores and vegetarians found that creatine supplementation benefited memory only for vegetarians but not omnivores [23]. Vegetarians have been found to have lower serum and muscle creatine concentration, but comparable total brain creatine to omnivores [22, 24, 25]. In this study, we included both omnivores and vegetarians to allow comparison. We hypothesised that creatine supplementation would improve working memory and reasoning ability in vegetarians. We also hypothesised that the improvement would be greater in vegetarians than in omnivores.

To test these hypotheses, we approximately replicated the study design and treatment (5 g per day of creatine for 6 weeks) used by Rae et al. [8]. We included the same primary outcome measures, the Backwards Digit Span and 10-min standardised subtests of Raven's Advanced Progressive Matrices. In addition, to investigate a broader range of cognitive functions, we included exploratory tests on attention, verbal fluency, task switching, and memory.

Methods

Trial design

We conducted a randomised, placebo-controlled, double-blind, cross-over study. The primary endpoints are the scores in the cognitive tasks after 6 weeks of each supplementation. Six weeks is the duration used by Rae et al. [8]. We learned from private correspondence with Turner et al. [26] that they chose a liberal 5-week washout period for their brain creatine study based on the muscle creatine literature, which suggests 5 weeks is sufficient [27–29]. Given the way creatine is transported, stored, and excreted, they expected the same washout period to be sufficient for the brain. They confirmed that in their study, brain creatine levels were back to normal after the washout. This is evidence that 5 weeks is an upper bound for how long brain creatine takes to wash out. We also see no reason why brain creatine would take longer to wash out than muscle creatine. Note that Turner et al. [26] used a shorter (7 days) but higher-dosed (20 g) supplementation regimen compared to the present study. However, there is evidence that the supplementation regimens are equivalent in terms of creatine saturation (e.g. [27]). Unlike Rae et al. [8], we did not have an extra washout period nor second baseline testing after the first

supplementation. Instead, we relied on the 6 weeks of placebo supplementation for washing out the creatine. The trial evaluated cognitive performance after creatine compared to placebo. The trial design and participant flow are summarised in Fig. 1. We follow the CONSORT reporting guidelines [30].

Participants

Participants were 18 years or older (see Appendix for full list of inclusion criteria). Half of them reported being on a vegetarian diet and half of them on an omnivore diet. A screening questionnaire assessed if the eligibility criteria were met. Participants who met these criteria went through the baseline assessment and were given their first supplement to take home (or for participants tested online, received the two supplements via the mail). Cognitive assessments of participants took place in the clinic laboratory. Due to the contact restrictions due to the COVID-19 pandemic, after 04/2020, participants were tested online via video call instead.

Interventions and similarity of treatment groups

Participants took the supplements daily for 6 weeks, including the day of the testing. The creatine supplement consisted of creatine monohydrate powder "CreaPure PG" produced by the company Alzchem (Trostberg, Germany). The placebo supplement consisted of maltodextrin powder "Maltodextrin 6" produced by the company Nutricia (Frankfurt am Main, Germany).

The cans looked exactly the same except for clear markings of which one was the first and which one the second supplement. The two powders looked exactly the same and were flavourless. The solubility was somewhat different: While the placebo powder was completely soluble in water and did not settle, the creatine powder slowly settled. We were initially not aware of this difference in solubility. After we noticed it (after the first 40 participants), we asked participants to stir the powder into yoghurt or food with a similar consistency, as we had found no perceptible difference then. To check to what extent blinding was achieved, directly after the last testing, participants were asked to guess what their first supplement had been.

Outcomes

We had two primary outcomes:

- A standardised 10-min subtest of Raven Advanced Progressive Matrices (RAPM) [8]
- The Wechsler auditory Backward Digit Span (BDS) [31]

RAPM is a test of abstract reasoning. Each item in the test consists of a 3×3 matrix with pictures of geometric forms. One of the pictures is missing and the task consists of choosing the right picture to fill this gap out of eight alternatives. The full RAPM consists of 80 items and has a time limit of 40 min. We used the same standardised 10-min subtests of the RAPM as Rae et al. [8], consisting of 20 items each. The subtests are constructed to have equal levels of difficulty based on the published normative performance data and Rae et al. [8] additionally verified this in an independent sample ($N=20$). The RAPM score consists of the sum of correct responses.

The Backward Digit Span is a test of working memory. The tester reads increasingly longer series of digits to the participant whose task it is to remember and repeat them in reverse order. The task starts with two digits. Each length has two series of digits. The test ends after wrong answers to two series of the same length. The BDS score consists of the sum of correct responses.

We had eight further exploratory outcomes:

- The D2 Test of Attention [32], a test of sustained attention
- The Trail-Making-Test A (TMT-A), a test of visual attention [33]
- The Trail-Making-Test B (TMT-B), a test of task switching [33]
- The Block-Tapping-Test, a test of visuospatial working memory [34]
- The Auditory Verbal Learning Test (AVLT, in German: VLMT), a word-learning test including immediate recall, delayed recall, and recognition [35]
- The Brief-Visuospatial-Memory Test—Revised (BVMT-R), a test of visuospatial memory [36]
- The Stroop test (in German: Farb-Wort-Interferenz Test, a test of inhibitory control [37]
- Regensburger Wortflüssigkeitstest, a test of verbal fluency [38]

Participants reported side effects experienced during the supplementation period in a free text form on the day of testing. How side effects would be grouped for the report was determined after evaluating all entries. At baseline testing, participants performed a test of crystallised intelligence called "Mehrfachwahl-Wortschatztest (MWT-B)" [39]. In this test, participants had to identify real German words among made-up words.

Sample size

The sample size of 123 was powered (with power = 0.8, alpha = 0.05, calculated with GPower [40]) to detect the effects of Cohen's $d=0.45$. The sample size (preregistered as 120) was chosen based on a conservative estimate (see Appendix) of the effect size in Rae et al. [8] ($d=1$ for both

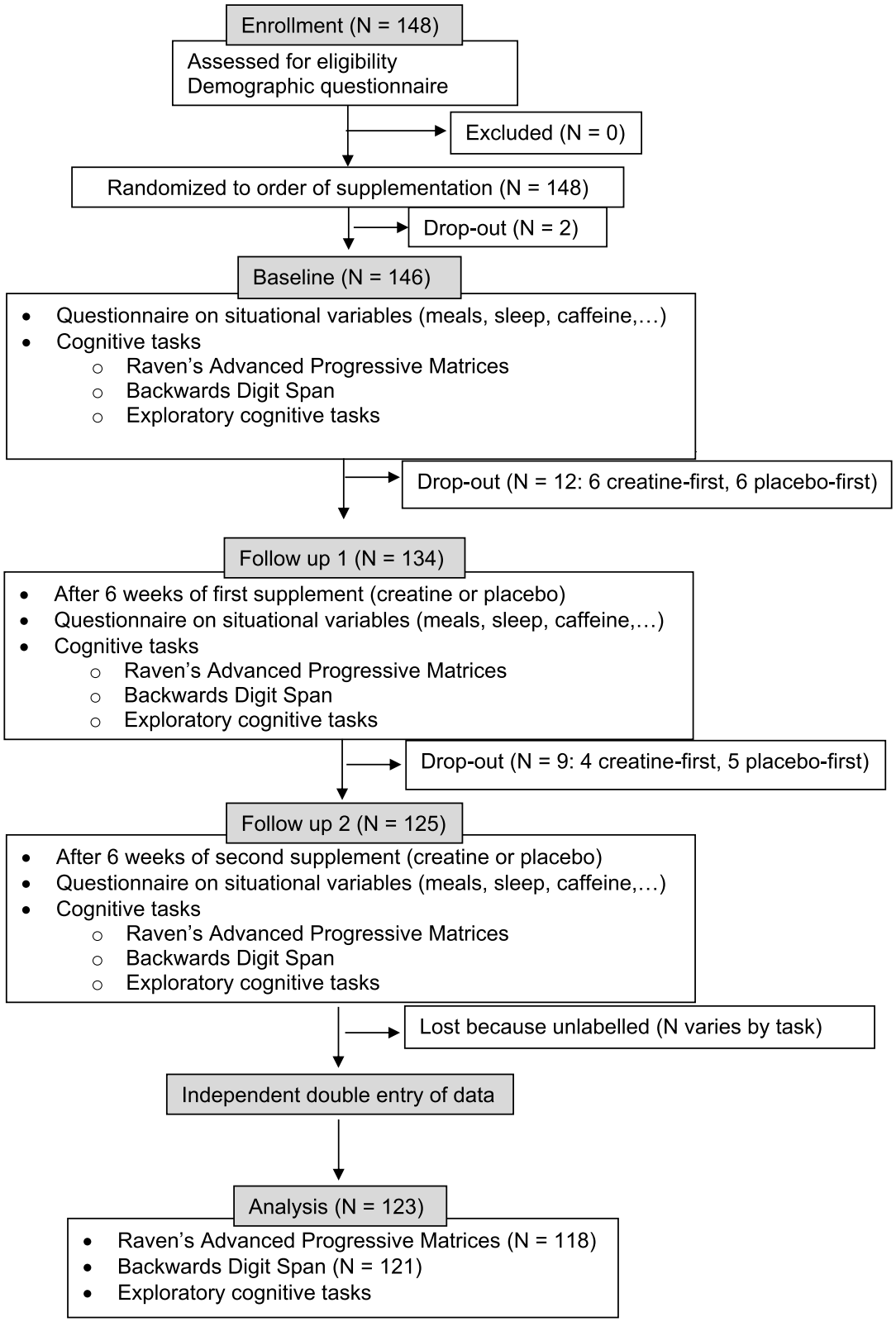


Fig. 1 Participant flow through the study

RAPM and BDS) with a substantial buffer to account for smaller effects.

Block-tapping was originally performed with physical blocks and later on the website Psytoolkit [41, 42] as part of remote testing during the COVID-19 pandemic. Because the remote version was not immediately available, the participant number is lower for this task.

Randomisation and blinding

The order of the two supplements was randomised with Excel by the pharmacy of the University Hospital Heidelberg. They labelled each of the cans of supplements with the participant code and "A" or "B", corresponding to the first and second supplement. The staff members who tested participants also provided the participants with the supplement cans. Allocation concealment was performed using sequentially numbered, opaque sealed envelopes (SNOSE). Participants and all staff who interacted with them were kept blinded to the allocation (also see intervention section).

Statistical methods

For each cognitive test, we conducted a mixed ANOVA with test score after supplementation as the dependent variable, supplement (creatine vs placebo) as the within-subjects factor and supplement order (creatine-first vs placebo-first) as the between-subjects factor[1]. We did not remove outliers in our main analysis, but conducted robustness checks which included trimming and winsorising. We applied the Greenhouse–Geisser correction to all our analyses but the correction did not change any value. We did not exclude participants based on adherence (intention-to-treat analysis). We analysed participants' actual and not assigned order of supplements. Actual and assigned order differed for only one participant.

We conducted a frequentist analysis, because it is widely understood and it offers the insight of how likely the data is under the null hypothesis. However, we also wanted to know whether the data was more likely under the null or under the alternative hypotheses, i.e. in which direction to update our credence, and to what extent. We conducted a Bayesian analysis, because it answers this question. We were interested in comparing the null hypothesis to different alternative hypotheses—postulating a large effect like in the study we replicated [8] and postulating effects of smaller, in our view more realistic, sizes. A significant result does not imply Bayesian evidence in favour of the alternative hypothesis, nor does a nonsignificant result imply Bayesian evidence against the alternative hypothesis [43]. If the alternative hypothesis postulates a relatively large effect, the data may be more likely under the null than the alternative hypothesis, despite a p-value below 0.05. If the alternative hypothesis postulates a relatively small effect, the data may be more likely under the alternative than the null hypothesis, despite a p-value above 0.05.

Confirmatory analyses

As preregistered, our two confirmatory cognitive tasks are the Backward Digit Span and Raven's Advanced Progressive Matrices. All other cognitive tasks are analysed in an exploratory fashion. There is one deviation from our preregistered analyses. We had preregistered t-tests, but this was a mistake in the preregistration. The t-test is not appropriate here because imbalances in the supplement order group sizes would bias the results. Instead, we conducted mixed ANOVAs with supplement (creatine vs placebo) as the within-subjects variable, supplement order (creatine-first vs placebo-first) and diet (vegetarian vs omnivore) as the between-subjects variables, and test score after supplementation as the dependent variable.

Robustness checks

We checked the robustness of our normality-assuming ANOVAs by performing: an ANOVA on 20%-trimmed data, an ANOVA on 5%- and 20%-winsorised data, and a robust ANOVA which uses trimming and bootstrapping (performed with the sppb functions in the WRS2 R package). The latter ANOVA provides the most robust estimate of these methods [44, 45].

Bayes factors

For the calculation of the Bayes factors, we used the estimated marginal means (EMMs) of the creatine and placebo score. The EMMs are the means weighed for the order groups (creatine-first and placebo-first), so that imbalances in the sizes of the order groups do not affect the means. So, we only had two groups for the Bayes factor calculation (creatine and placebo), simplifying the analysis. The mean difference and standard error of the mean difference were used to describe the data. Using the Bayesplay package [46], we calculated the Bayes factors in several different ways. Approach 1 used point models for the null hypothesis and the alternative hypotheses. Approach 2 compared a point null model against half normal distributions centred on zero and with the standard deviation set to half the maximum expected effect size. For the reasons behind this, see the Appendix.

[1] We used the score after supplementation and not change from baseline, because subtracting the same baseline from both after-supplement scores in a crossover study would cancel out, which would give the same result but complicate the analysis unnecessarily. The baseline was used for describing participants' baseline characteristics and for exploratory analyses.

Table 1 Participant baseline characteristics

N	Total 125	Creatine-first 63	Placebo-first 62
Age in years (M, SD)	30.6 (10.1)	31.5 (10.4)	29.8 (9.7)
Sex (% female)	57%	54%	60%
Weight in kg (M, SD)	70.3 (13.7)	71.8 (15.5)	68.8 (11.4)
MWT-B (M, SD)	26.31 (4.35)	26.32 (3.99)	26.31 (4.72)

Data is given as mean (standard deviation) or as percentage. The MWT-B (Mehrfach-Wahl-Wortschatztest) is a test of crystallised intelligence [39]

Table 2 Adherence and negative side effects

	Creatine	Placebo
Days supplemented per week (M, SD)	6.89 (0.26)	6.87 (0.26)
Any side effects	17%	4%
Of these		
Digestion problems	6%	2%
Weight gain	3%	0%
Other		
- Tiredness	1x	1x
- Thirst	1x	1x
- Weight loss	1x	0x
- Nightmares	1x	0x
- Cramps	1x	0x
- Thoughts racing	1x	0x
- Problems concentrating	1x	0x
- Nervousness	1x	0x

Exploratory analyses

In addition to the confirmatory analyses of BDS and RAPM, we analysed the other cognitive tasks in the same way in an exploratory fashion.

We also looked in an exploratory fashion at the first supplementation and the second supplementation separately and at participants with a low and high baseline performance separately (see Appendix).

Results

Participant flow

See participant flow in Fig. 1.

Drop-outs were due to supplements failing to arrive ($N=2$), dealing with stressful personal events ($N=2$), no reason given ($N=2$), and no time ($N=18$).

Recruitment

Participants were recruited through flyers and social media between 05/2019 and 05/2022 and tested between 05/2019 and 08/2022.

Baseline data

We analysed all available participant data apart from two minor exceptions (see Appendix). Participants were included irrespective of their adherence. The median number of days per week with meat consumption for omnivore participants was 3.5 (mean = 3.7, SD = 2.0). For further participant characteristics, see Table 1.

Blinding, adherence, and side effects

After their final testing session, the last 73 participants were asked to guess the order of their supplements (as the idea did not occur to us before). Forty-three (59%) guessed correctly and 30 (41%) guessed incorrectly. A binomial test reveals that the probability of 43 or more correct guesses out of 73 by pure chance is $p=0.080$. However, most participants who guessed correctly reported being very unsure about their guess. We recorded the reasons for the guesses of the last of the 33 participants. Of those participants who had a reason for their guess, solubility was the most common, followed by negative side effects and positive side effects. All three reasons seemed to improve guess accuracy (see Appendix).

A z-score test for two population proportions revealed that the proportion of participants reporting any negative side effect was significantly higher for the creatine than the placebo condition, $p=0.002$, RR = 4.25 (Table 2). In addition, although we did not assess this systematically, some participants reported positive side effects such as improvements in strength (several participants) and mood (one participant). No patients discontinued the study due to an adverse event.

Adherence (self-reported) was high (Table 2). All but one participant took the supplements in the order assigned to them. This participant was analysed with their actual, not their assigned, supplement order.

Interaction with diet

There was no significant interaction between diet and supplement nor between diet, supplement, and supplement order for neither BDS ($p=0.808$ and $p=0.559$) nor RAPM ($p=0.392$ and $p=0.606$), nor was the interaction in the predicted direction (we had hypothesised that vegetarian participants would benefit more from creatine than omnivore participants). This was also true when using the robust ANOVA based on bootstrapping. Bayes factors favoured the null hypothesis. To be precise: They indicated strong support in favour of the null hypothesis over the effect size in Benton and Donohoe [23] ($d=0.36$) and weak to strong support in favour of the null hypothesis over smaller effect sizes (see Appendix). There was no indication for an effect of diet in the exploratory cognitive tasks either. For more details on the analysis of diet, see the Appendix.

Table 3 Results of confirmatory analysis

Task	*N*			Supplement effect 3-way ANOVA, inkl. diet			Supplement effect 2-way ANOVA			Crea. score	Pl. score	Crea.-Pl. scores M(SE) [95% CI]
	Total	Crea. first	Pl. first	*F* (df)	*p*	η^2_P	*F* (df)	*p*	η^2_P			
BDS	121	61	60	3.41 (1, 117)	**0.067**	**0.028**	3.49 (1, 119)	**0.064**	**0.029**	8.85 (0.28)	8.44 (0.25)	0.41 (0.22) [−0.24; 0.84]
RAPM	118	60	58	1.02 (1, 114)	**0.315**	**0.009**	0.97 (1, 116)	**0.327**	**0.008**	12.39 (0.28)	12.16 (0.28)	0.23(0.23) [−0.24; 0.70]

Mixed 3-way ANOVA with supplement (creatine vs placebo) as the within-subjects variable, supplement order (creatine-first vs placebo-first) and diet (vegetarian vs omnivore) as the between-subjects variable and test score after supplementation as the dependent variable. Mixed 2-way ANOVA without diet. The test score is given as estimated marginal mean (standard error). *P*-values are two-tailed. The two cognitive tasks are the Backward Digit Span and Raven's Advanced Progressive Matrices

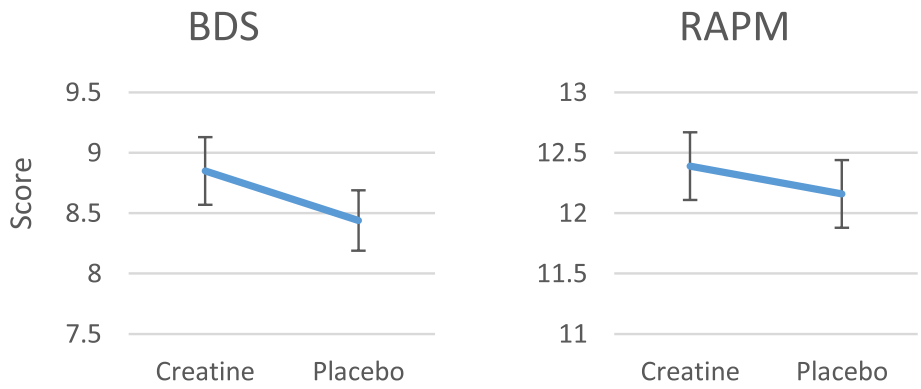


Fig. 2 **a** Estimated marginal means for the Backward Digit Span (BDS) score. **b** Estimated marginal means for Raven's Advanced Progressive Matrices (RAPM) score. Error bars represent standard errors

Confirmatory analysis

There was a significant interaction between supplement and supplement order for both BDS and RAPM. This seems to reflect a learning effect (see Appendix). The effect of most interest, the main effect of the supplement, was in the expected direction but not significant. However, it bordered on significance for BDS ($p=0.067$, $\eta^2_P=0.028$). This means that 2.8% of the variance in BDS scores that was not already explained by other variables was explained by the supplement. For RAPM, it was 0.9%. The supplement effect was virtually the same whether diet was included as a variable or not (Table 3). Thus, we simplified additional analyses (estimated marginal means, Bayes factors, and robustness checks) by dropping diet as a variable for these analyses.

In terms of raw scores, the effect size for BDS was 0.41 additional correct items, i.e. a 0.2-digit longer digit span, because there were always two-digit spans of the same length. For RAPM, the effect was 0.23 more matrices solved (Fig. 2). If these were IQ tests, this increase in raw scores would mean 2.5 IQ points for BDS (using the standard deviations of a normative study [47] or our own baseline gives the same result, see Appendix). For RAPM, the improvement would be 1 IQ point (using the standard deviation of our own baseline, see Appendix). Cohen's *d* based on the estimated marginal means of the creatine and placebo scores was 0.09 for RAPM and 0.17 for BDS.

Bayes factors

To facilitate the interpretation of the results of the confirmatory analysis, we provide Bayes factors. A Bayes factor (BF_{10}) indicates how likely a null hypothesis is compared to an alternative hypothesis given the data. A BF_{10} between 1/3 and 3 indicates low sensitivity of the data (i.e. not enough data to be certain), with weak evidence in favour of the null hypothesis if BF_{10} is below 1 and weak evidence in favour of the alternative hypothesis if it is above 1. A BF_{10} above 3 (below 1/3) is considered moderate and above 10 (below 1/10) strong evidence [48].

We compare several alternative hypotheses postulating small beneficial effects of creatine to the null hypothesis. For RAPM, the data was very insensitive, very weakly favouring the alternative hypotheses. For BDS, the data was more sensitive, providing weak to moderate support in favour of the alternative hypotheses. Two different approaches to calculating these Bayes factors were used (see statistical analysis) and the results were similar (Table 4).

There was strong evidence in favour of the null hypothesis compared to the alternative hypothesis

Table 4 Results of Bayesian analysis

Task	Approach 1: point models				Approach 2: half normal		
	Small effects			Rae-sized	Small effects		Max. = 2 × Rae-size
	0.1	0.2	0.4	2.5	max. 0.4	max. 1	max. 5
BDS	2.1	3.6	5.7	< 2e − 7	2.9	3.3	1.0
RAPM	1.4	1.6	1.3	< 2e − 7	1.4	1	0.3

Bayes factors (BF_{10}) comparing a range of alternative hypotheses to the null hypothesis. The effect size is given as the raw score difference. Approach 1 compared a point null model to point alternative models with a range of small effect sizes (0.1–0.4, i.e. $d = 0.04$–0.17) as well as an equivalent of Rae et al.'s effect size (2.5, i.e. $d = 1$, see calculation in Appendix). Approach 2 compared a point null model against half-normal distributions centred on zero and with the SD set to half the maximum expected effect size

Table 5 Results of robustness checks

Task	Better score	Max. skew	*p* (Supplement)				
			Normal	20% trim	5% winsorisation	20% winsorisation	Bootstrap and 20% trim
RAPM	creatine	−059	0.327	0.412	0.361	0.672	0.354
BDS	creatine	1.14	0.064	0.17	0.009	0.05	0.37

Creatine effect *p*-values (two-tailed) for different ANOVAs. The given trim and winsorisation percentages are applied to each side. Better score based on estimated marginal means. "Max. skew" gives the highest skewness statistic in any combination of conditions (supplement and supplement order)

postulating the effect size found by Rae et al. [8]. The data was insensitive (BDS) or weakly favoured the null hypothesis (RAPM) when compared to the half-normal model based on Rae et al. [8]. The half-normal model based on Rae et al. [8] does not assume their effect size is the true effect size in the population. Instead, the model assumes their effect size is a moderate overestimation of the true effect size. The model uses their effect size as a reference point to assign probabilities to effect sizes. It assigns most of the probability weight to effect sizes that are smaller than this effect size, and some probability to effect sizes up to twice that effect size. This is a common alternative model when replicating studies. However, we did not use it as our only model, because we were also interested in assessing the likelihood of smaller effect sizes and of the possibility that the effect size in Rae et al. [8] was the true population effect size.

The results were similar whether using normal or Cauchy distributions. For more details on this and the aforementioned calculations, see the Appendix.

In summary, this study provides weak to moderate evidence for a small cognitive benefit of creatine and strong evidence against the effect size by Rae et al. [8] being representative.

Robustness checks

We checked the robustness of our confirmatory analysis (the normal ANOVA) by performing an ANOVA on 20%-trimmed data, an ANOVA on 5%- and 20%-winsorised data, and an ANOVA which uses bootstrapping and 20% trimming.

For RAPM, all of these methods gave overall similar results to that of the normal ANOVA (Table 5).

For BDS, whose skewness statistic was slightly further from 0 than that of RAPM, these methods gave results that differ from each other and from the normal ANOVA to a relevant extent (Table 5). Most notably, the *p*-value for the supplement effect was 0.009 for the 5% winsorisation and 0.370 for the bootstrap ANOVA. This seems to suggest that in the normal ANOVA, the most extreme values made the effect of creatine appear smaller by inflating the variance, while relying on possibly unjustified assumptions of normality made the effect of creatine appear larger.

Thus, the result for RAPM was robust and for BDS much less so.

Exploratory cognitive tasks

There was no indication that creatine improved the performance of our exploratory cognitive tasks. The distribution of *p*-values was what one would expect if there was no effect. For the exploratory cognitive tasks, Table 6 only includes the *p*-values of the supplement effect. For the full results, including the interaction effect (reflecting a learning effect) and the order of supplement effect, see the Appendix.

Table 6 Results for exploratory cognitive tasks

Task	*N*	Better score	*p* (Supplement)				
			Normal	20% trim	5% winsorisation	20% winsorisation	Bootstrap and 20% trim
Block-tapping forward	71	creatine	0.779	0.564	0.865	0.678	0.826
Block-tapping backward	70	placebo	0.83	0.87	0.482	0.482	0.59
BVMT-R	119	creatine	0.543	0.809	0.746	0.112	0.67
D2 test	104	placebo	0.394	0.382	0.446	0.291	0.62
Forward digit span	117	placebo	0.714	0.795	0.838	0.721	0.52
Stroop—colors	118	placebo	0.813	0.184	0.432	0.054	0.568
Stroop—colorletters	119	placebo	0.626	0.877	0.861	0.547	0.856
TMT A	123	placebo	0.129	0.04	0.068	0.021	0.224
TMT B	122	creatine	0.855	0.622	0.745	0.567	0.56
VLMT immediate recall	119	creatine	0.87	0.744	0.996	0.883	0.996
VLMT recall after interference	119	creatine	0.694	0.622	0.854	0.323	0.806
VLMT delayed recall	118	placebo	0.339	0.346	0.462	0.133	0.54
VLMT recognition	117	creatine	0.722	0.327	0.398	0.635	0.348
Word fluency	122	creatine	0.227	0.631	0.272	0.119	0.692

Creatine effect *p*-values (two-tailed) for different ANOVAs. The given trim and winsorisation percentages are applied to each side. Higher score based on estimated marginal means

Discussion

Summary of results

This is the largest study on the cognitive effects of creatine to date. As part of our study, we aimed to replicate Rae et al. [8], who found a large positive effect of creatine on the abstract reasoning task Raven's Advanced Progressive Matrices (RAPM) and on the working memory task Backward Digit Span (BDS) in healthy young adult vegetarians.

We found Bayesian evidence for a small beneficial effect of creatine on cognition for both tasks. Cohen's *d* based on the estimated marginal means of the creatine and placebo scores was 0.09 for RAPM and 0.17 for BDS. If these were IQ tests, the increase in raw scores would mean 1 and 2.5 IQ points. The preregistered frequentist analysis of RAPM and BDS found no significant effect at $p < 0.05$ (two-tailed), although the effect bordered significance for BDS. There was no influence of diet (vegetarian vs omnivore), age, or sex on this effect. There was no indication that there was a creatine effect in several other cognitive tasks that we studied in an exploratory analysis.

Effect of diet

Dietary creatine is primarily contained in meat, fish, and a small amount in some dairy products [13, 14]. So, creatine is almost non-existent in a vegetarian diet. This might lead one to expect vegetarians to benefit more from creatine supplementation than omnivores do. However, even in an omnivore diet, creatine intake through the diet is low compared to common doses for creatine supplementation. Therefore, it is possible that the low dose in the diet does not affect cognition, while the higher dose of supplementation does. As opposed to muscles, which always receive creatine from other parts of the body, the brain synthesises its own creatine and for that reason might be more resistant to exogenous creatine [22, 49]. This might mean that higher doses of creatine are necessary to increase brain creatine levels [22, 50, 51]. In our study, half of the participants were vegetarians and half of them were omnivores. We found no indication that our vegetarian participants benefited more from creatine than our omnivore participants (in fact, the creatine effect was smaller in vegetarians than omnivores to a non-statistically significant extent). This is in line with Solis et al. [22, 24] who did not find a difference in brain creatine content between omnivores and vegetarians. Our Bayesian analysis of their data provides moderate support for the lack of a difference (see Appendix). In contrast, Benton and Donohoe [23] found that creatine supplementation benefited memory in vegetarians more than in omnivores, with no difference in baseline performance. However, given the high number of cognitive tasks in that study, the chance of a false positive was high, so we regard their finding as exploratory. Apart from the present study and Benton and Donohoe [23], we are not aware of any other RCT comparing the effect of creatine supplementation on cognition between vegetarians and omnivores.

Rae et al. [8] is the only RCT on the effect of creatine supplementation on cognition with only vegetarian participants (healthy, under normal conditions) and they found a large creatine effect. Under the same conditions, a number of other studies with omnivore participants [52–55] or unspecified and presumably omnivore participants [56–58] also found a large creatine effect, while other studies with omnivore participants [21, 59, 60] or unspecified and presumably omnivore participants [61] failed to find a creatine effect[2,3]. The omnivore findings are mixed while the only study on vegetarians [8] is positive, but as it is only one study, it is not a strong indication.

Observational data on the role of dietary creatine in cognition is conflicting. Ostojic et al. [62] (1340 elderly adults in the USA) found a significant positive correlation between performance in the Digit Symbol Substitution Test and amount of meat consumed. By contrast, two studies by Giem et al. [63] in Seventh Day Adventist adults in the USA (272 participants matched for age, sex, and zip code and 2984 unmatched participants) both found a trend towards delayed onset of dementia in vegetarians compared to omnivores. Many confounders are possible in these observational studies, such as health influencing both diet and cognitive performance, differences in lifestyle, and components in meat other than creatine. None of the studies assessed creatine supplementation.

Overall, observational evidence is mixed. Evidence from RCTs and brain creatine studies does not support the idea that dietary creatine affects brain health.

BDS and other short-term memory tasks

While the creatine effect for BDS in our study bordered significance, it was smaller ($d = 0.17$) compared to the two other studies in the literature that have tested the effect of creatine supplementation on BDS in young, healthy adults (Rae et al. [8]; Hammett et al. [56])[4].

One reason for this could be that there might have been more noise in our study, maybe due to the COVID pandemic starting in the midst of the study. This reason might apply to all of our tests. However, the standard deviation for BDS in this study was almost exactly the same as in the study we aimed to replicate[5] [8] (2.42 vs 2.38) and our sample size was much larger, so noise does not explain the difference in results. Our supplementation protocol was different from that of Hammett et al. [56], which could lead to different results. However, the other study [8] used the same supplementation protocol as ours (5 g/day for 6 weeks), so this reason seems unlikely. Compared to Rae et al. [8], one might at first glance think that their effect was larger because all of their participants were vegetarians whereas only half of our participants were vegetarians. However, the number of vegetarian participants in our study was still higher than theirs and the creatine effect was not larger in this subgroup (in fact, it was smaller to a non-statistically significant extent). There are reasons to think that creatine supplementation becomes more beneficial with age (more on this below). However, the age of our participants was very similar (mean = 31, sd = 10, median = 28) but slightly higher than those of Rae et al. [8] (median = 25.5) and Hammett et al. [56] (median = 26), so age does not seem to explain the difference in results. Overall, we did not find a convincing reason for why our BDS result is different from the two other studies with young, healthy adults. The difference is not explained by diet, age, or noise.

Three further studies tested the effect of creatine supplementation on BDS in a different population than healthy young adults. In two of these studies, participants were healthy and elderly [55, 61] and in one study sleep-deprived [64]. In these three studies, there was no creatine effect. This is surprising, seeing as theory and evidence generally suggest that sleep-deprived and older participants have a larger cognitive benefit from creatine supplementation (more on this below). One of the two studies on healthy elderly participants, Alves et al. [61], also tested the effect of creatine on strength and found no effect, which goes against a large body of literature. The effect of creatine supplementation on strength is well-studied and generally large. Given that this effect was not present in Alves et al. [61] makes it less surprising that they did not find an effect for cognition. The results of the two studies by McMorris et al. are more difficult to make sense of. The creatine effects for BDS in these studies were smaller than in the present study, while the effects in studies with the same kind of population as ours were

[2] Samadi et al. [59] mistakenly concluded they had found a creatine effect because their creatine group improved significantly while their placebo group did not. However, the creatine effect is correctly measured by the difference in change between the two groups (i.e. the time × supplement interaction) and this was not significant ($p = 0.84$), nor does it likely give Bayesian evidence in favour of a creatine effect, although we did not check this. Note also that in addition to creatine and placebo, both groups received beta-alanine.

[3] One of the studies with omnivores that did not find an effect was on children [48]. Another one also found no effect of creatine on strength, in contrast to a large body of literature [50].

[4] Rae et al. [8] was a crossover study like ours. As opposed to us, they included two baselines. To better compare our results to them, we did not just look at their reported effect but also calculated their effect size using just the after-supplement scores like we did in this study (see Appendix). Their effect was still much larger than ours.

[5] Using the doubled effect size for Rae et al. [8], because they did not present the same sequence length twice by default like we did.

larger than in the present study. Ignoring the differences in populations for a moment, it might simply be that the real effect of creatine on BDS performance lies in the middle like our study.

In a review by Avgerinos et al. [7] on the effects of creatine supplementation on cognition, studies testing BDS were grouped together with two studies with other short-term memory tasks: Benton and Donohoe [23] and Rawson et al. [60]. The short-term memory task in Benton and Donohoe [23] consisted of the recall of 30 words. There was a creatine effect only for their vegetarian participants for this task. However, as mentioned in the discussion section on diet, this study included a large number of tasks, so we consider this finding exploratory. The short-term memory task in Rawson et al. [60] was the Sternberg task, where participants had to memorise 6 letters displayed for 20 s and subsequently decide if "probe" letters were included or not in those 6 letters. Rawson et al. [60] found no effect of creatine, which Avgerinos et al. [7] attributed to the participants' young age (M = 20, SD = 2.2). While this is plausible in theory and there is evidence supporting older people benefiting more from creatine, it is unclear to what extent it explains the differences in findings in this case, as two studies with participants who were only slightly older [8, 56] found large creatine effects for short-term memory while the only two studies with elderly participants found no effect.

Five of our exploratory tasks, the forward digit span, forwards and backwards block-tapping task (spatial), the BVMT-R, and the immediate recall part of the VLMT also tested short-term memory and there was no indication of an effect for these tasks. A review by Dolan et al. [50] reports that a creatine effect is more likely for more cognitively demanding tasks. In line with this, we found some indication of a creatine effect for the backward digit span (BDS) but not for the less demanding forwards digit span and block-tapping tasks. The VLMT and BVMT-R may also be less cognitively demanding than the BDS, but this comparison is less obvious to make. Going against this idea, McMorris et al. [55] found a creatine effect for short-term memory tasks that we deem less demanding but not for BDS.

Overall, the effect size for BDS in the present study lies in between that of other studies, some of which found very large effects and some of which found no effect. The reason why we found an effect for BDS, which is a short-term memory task, but not for our exploratory short-term memory tasks, might be that BDS is probably more difficult and thus requires a higher ATP turnover, which is where creatine helps. Short-term memory is critical for language comprehension, learning, planning, reasoning, and general fluid intelligence [65], so even a small improvement in the most difficult tasks might be very valuable.

RAPM and other abstract reasoning tasks

In our study, the creatine effect for Raven's Advanced Progressive Matrices (RAPM) was much smaller ($d = 0.09$) compared to two of the three other studies in the literature that have tested the effect of creatine supplementation on RAPM in healthy adults [8, 54, 56]. It is unclear how our effect for RAPM compares to that of the third study [56] because they did not report this effect size or any other information that could be used to calculate it. However, as their creatine effect for RAPM was in the predicted direction but not significant and presumably also not bordering significance, the effect in that study was likely not very different from that of the present study. So, our results are likely in line with those of Hammett et al. [56]. Hammett et al. [56] used a shorter version of the RAPM (5 min) compared to Rae et al. [8] (10 min). It seems plausible that this might have made the task less robust and thereby reduced the effect.

Rae et al. [8] have already been discussed above for BDS—the same points apply here. In short: No reason is evident why the effect would be smaller in the present study. Noise in RAPM was lower in the present study than in Rae et al. [8], so this does not explain the result. Again, age does not explain the difference in results. In line with the idea that cognition in young people benefits less from creatine, a large study with children (aged 10–12) did not find a significant creatine effect on RAPM [21]. However, the age of our participants (mean = 31, sd = 10, median = 28) and Hammett et al. [56] (median = 26) was slightly to moderately higher than in the other two adult studies (Rae et al.: median = 25.5; Ling et al.: mean = 21).

In a review by Avgerinos et al. [7], studies testing RAPM were grouped together with a study by Rawson et al. [60] using a logical reasoning task and a mathematical processing task, which did not find a creatine effect. Avgerinos et al. [7] explained this difference in findings with the participants' age. The participants in Rawson et al. were younger than those of the other studies in the review [8, 54]. However, compared to Ling et al., the age difference is so small (mean age of 20 (SD = 2.2) vs 21 (SD = 1.4) years) that we do not think it explains the difference in the results of the two studies. A reason we find more compelling is the difference between the tasks. The mathematical processing task requires participants to decide if a three-step equation involving addition/subtraction gives a result lesser or greater than five. The logical reasoning task consists of a series of symbols, for example #@, followed by two statements about their

order, for example # before @ and @ after #. The task is to determine if both or only one of the statements is consistent with the series. Both tasks seem to differ from RAPM in being easier to solve, less varied and complex, requiring less creativity and providing no progression in difficulty. Alternatively, the negative finding by Rawson et al. could also be explained by the creatine effect in the reasoning domain being smaller than suggested by the findings of Rae et al. and Ling et al. The present study and Hammett et al. [56] are in line with this explanation.

In sum, two studies found much larger creatine effects for RAPM than that in the present study, while a third study did not find a significant effect. Possibly the real effect lies in the middle like our study. RAPM, while not a pure measure of g [66], substantially correlates with general intelligence [66] and predicts academic achievement [67]. Even small improvements would be very valuable.

Exploratory cognitive tasks other than short-term memory

We did not find a creatine effect for our exploratory tasks. Our negative finding for the verbal fluency task is in line with the only other study using the same task [23]. Our negative finding for the long-term memory part of the VLMT is in line with two studies assessing long-term memory [21, 61] and in contrast with one study which found a creatine effect for this domain [55]. The study which found an effect on long-term memory used an idiosyncratic task in which participants had to remember a combination of occupations and photographs of faces. It might be that this task has characteristics that make it more susceptible to benefit from creatine—e.g. perhaps it is more difficult. Our negative finding for the trail-making task, which tests task switching, is in line with the same two studies [21, 61]. Our negative finding for the Stroop task, a test of inhibition, is again in line with the same two studies [21, 61] and in contrast with one study which found a creatine effect for this task [57]. In contrast to Van Cutsem et al. [57], Alves et al. [61] found no creatine effect on strength and the participants in Merege-Filho et al. [21] were children (aged 10–12). These differences might explain why Van Cutsem et al. [57] found an effect while the other two studies did not find an effect of creatine for any of their tasks. In addition, Van Cutsem et al. [57] modified the Stroop task in various ways that increased its difficulty. More difficult tasks have been hypothesised to benefit more from creatine supplementation [50], because they require more energy, i.e. a higher ATP turnover, which is benefited by creatine. This modification might be why Van Cutsem et al. [57] found a creatine effect for this task while the present study did not.

In sum, for the exploratory tasks, overall the evidence does not support a creatine effect. However, as the evidence for the Stroop task shows, this might be only when the tasks are made too easy for participants, so that creatine has no chance to help.

Effect of age

It has been claimed that creatine supplementation is more likely to benefit older adults more than younger ones [7, 68, 69]. So, one reason why creatine did not impact most of the cognitive tasks in this study might be that most of our participants were relatively young (mean = 31, sd = 10, median = 28).

One theory behind an effect of age is that brain creatine levels might decrease with age. There is evidence that this happens with muscle creatine levels [70–73] (although see [74, 75]), although it is unclear if this is an effect of ageing itself or a result of other reasons such as dietary choices or reduced physical activity [22]. Similarly, brain creatine might be affected by ageing directly or mediated by reduced brain activity. However, Solis et al. [22] found no difference between the brain creatine levels (nor muscle creatine) of young and elderly adults, which speaks against this theory.

Experimental evidence suggests an effect of age. A meta-analysis by Prokopidis et al. [68, 69] on RCTs with healthy participants on an omnivore diet found that creatine supplementation significantly improved memory in older adults (aged 66–76 years) but not younger adults. However, there are reasons for caution. The meta-analysis included only two studies ($N = 25$ and $N = 32$) with elderly participants [55, 61], only one of which found a creatine effect. For reasons unknown to us, the meta-analysis did not include Hammett et al. [56], who found a large creatine effect on memory in young adults and might have changed the conclusion. In addition, as described above, a number of other studies with participants younger than ours found large creatine effects. The observational evidence reported in the discussion of diet [62] is consistent with an effect of age. However, as they did not include young adults, we cannot know if they would not show the same correlation.

Effect of gender

In their review, Smith-Ryan et al. [76] theorise that women might benefit more from creatine supplementation than men. The limited evidence from the present study and previous studies does not support this idea.

There have only been three RCTs on the effect of creatine on cognition in healthy women and three in healthy men. One of the studies with women found a creatine effect and two did not find an effect (one of these with only elderly participants). Two of the studies with men (one of these with sleep-deprived participants) found a creatine effect and one did not find an effect. Apart from

the present study, which found no effect of sex, studies who included both men and women did not report the effect of sex. In their meta-analysis on the effect of creatine on memory performance, Prokopidis et al. [68, 69] found no effect of sex.

Limitations

There are a number of limitations to this study. Despite the large sample size compared to other studies, a larger sample size would be needed to be powered for effects that are smaller but still relevant. Some of the data (4%) could not be analysed because it was not labelled with the participant and timepoint. The COVID-19 pandemic started in the middle of the study, might have added noise to the data, and meant that we had to switch from in-person cognitive testing to testing via video call. However, we do not see this potential source of noise reflected in the standard deviations compared to pre-pandemic studies. We assessed baseline days per week of meat consumption (median = 3.5, mean = 3.7, SD = 2.0), but not grammes per day of consumption. However, creatine intake through meat is usually substantially lower than the supplemented dose [13]. Adherence was self-reported and not checked with blood samples. A major limitation is that brain creatine levels were not assessed. Another limitation is that the proportion of participants who correctly guessed their supplement order (59%) bordered on significance ($p = 0.08$). However, most participants who guessed correctly reported being very unsure about their guess. The largest contributing factor to correct guesses was likely the difference in the solubility between the powders, followed by negative and positive side effects. We attempted to counteract differences in solubility by recommending participants to stir the supplements into yoghurt. Mixing creatine in yoghurt rather than water might have negatively affected creatine absorption, because the lower water content and usually cold temperature of yoghurt would decrease creatine solubility [77]. However, yoghurt is still relatively liquid and has a high water content of about 70–90% [78]. Yoghurt provides carbohydrates (4–18 g/100 g) [78], which has been found to aid creatine retention [27, 79]. In addition, yoghurt has a lower pH value than water (4 vs. 6.5–9-5) [80, 81], a factor likely aiding absorption [82]. Overall, our educated guess is that creatine absorption when mixed in yoghurt is similarly effective as in cold water but worse than in warm water. For future studies, we recommend cellulose as the placebo and a mixture of cellulose and creatine as the treatment, as these two look extremely similar when dissolved in water. The alternative solution with capsules would require participants to consume many capsules per day. This would likely reduce adherence and massively increase costs. Unfortunately, it is difficult to achieve perfect blinding when side effects occur with higher frequency in the creatine condition. The side effects of creatine are well-known and not dangerous [1, 4–6].

Conclusions

Supplementing creatine is safe, easy, and very cheap. The real effect of creatine on cognition is likely smaller than that reported in Rae et al. [8]. However, even small improvements in cognition may be relevant, especially if accumulated over many people and over time. The results of this study do not allow any strong conclusions, but it would be worthwhile to test for a small effect of creatine in strategically designed, larger studies.

Abbreviations

95% CI	95% Confidence interval
ANOVA	Analysis of variance
ATP	Adenosine triphosphate
AVLT/German: VLMT	Auditory Verbal Learning Test/German: Verbaler Lern- und Merkfähigkeitstest
BDS	Backwards Digit Span
BF_{10}	Bayes factor that indicates how likely a null hypothesis is compared to an alternative hypothesis given the data
BVMT-R	Brief Visuospatial Memory Test-Revised
Crea.	Creatine
df	Degrees of freedom
EMM	Estimated marginal mean
F	*F*-value
IQ	Intelligence quotient
M	Mean
MWT-B	Mehrfach-Wahl-Wortschatztest, a test of crystallised intelligence
p	*p*-value
Pl.	Placebo
RAPM	Raven's Advanced Progressive Matrices
RCT	Randomised controlled trial
RR	Relative risk
SD	Standard deviation
SE	Standard error
SNOSE	Sequentially numbered, opaque sealed envelopes
TMT-A	Trail-Making-Test, version A
TMT-B	Trail-Making-Test, version B
η^2_P	Partial eta squared

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12916-023-03146-5.

Additional file 1.

Acknowledgements

We thank the doctors of the University Clinic Bonn who collected blood samples. We thank Prof. Dr. Rand R. Wilcox, Dr. Lincoln Colling, Dr. Christian Stark, Jan Speller, Maximilian Meier, and David Reinstein for their feedback on statistical questions. We thank Thomas Szpejewski and Tom Lieberum for their help with verifying data quality. We thank all data entry helpers.

Authors' contributions

JFS and JB conceived of the study. JFS, XK, SL, AP, HM, and JB designed the study. AP and HM managed staff members as the director and vice director of

the hospital. JFS and JB wrote the grant proposal. XK wrote the proposal for the ethics committee. JFS preregistered the study. XK prepared the logistics of the study and supervised EKK. JFS, AF, and EKK tested participants. JFS managed data entry and analysed the data. JFS wrote and revised the manuscript with help from XK, GA, SL, UE, and JB. UE contributed to the supervision of JFS. JB was the main supervisor and senior author of the study. All authors read and approved the final manuscript.

Authors' Twitter handles

Twitter handles: @FabienneSand (Julia Fabienne Sandkühler)
Twitter handles: @JanMBrauner (Jan Brauner)

Funding

Open Access funding enabled and organized by Projekt DEAL. Funding was provided by the non-profit organisation Effective Ventures Foundation, 2443 Fillmore St., #380–16662, San Francisco, CA 94115, the non-profit organisation Effective Altruism Foundation, Efringerstrasse 25, CH-4057 Basel, Switzerland, and private donor Paul Christiano. The trial funders had no role in the design of the study; the collection, analysis, or interpretation of data; the writing of the report; or the decision to submit the article for publication.

Availability of data and materials

The appendix, protocol, data, code, and output of this study are openly available at the Open Science Framework, https://osf.io/xpwkc/.

Declarations

Ethics approval and consent to participate

Ethical approval was obtained from the ethics committee of the University of Bonn (060/19). Participants gave informed consent to participate before being enrolled.

Consent for publication

This manuscript contains only anonymised data. Participants gave consent for the publication of this data.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Department of Psychology, University of Bonn, Kaiser-Karl-Ring 9, 53111 Bonn, Germany. [2]Department of Psychiatry and Psychotherapy, University Hospital Bonn, Bonn, Germany. [3]Department of Psychiatry and Psychotherapy, University Medical Center, Mainz, Germany. [4]Manchester University NHS Foundation Trust, Manchester, UK. [5]Department of Health, Witten/Herdecke University, Witten, Germany. [6]Department of Computer Science, University of Oxford, Oxford, UK. [7]Future of Humanity Institute, University of Oxford, Oxford, UK.

Received: 14 April 2023 Accepted: 31 October 2023
Published online: 15 November 2023

References

1. Kreider RB, Kalman DS, Antonio J, Ziegenfuss TN, Wildman R, Collins R, et al. International society of sports nutrition position stand: safety and efficacy of creatine supplementation in exercise, sport, and medicine. J Int Soc Sports Nutr. 2017;13(14):18.
2. Butts J, Jacobs B, Silvis M. Creatine use in sports. Sports Health. 2018;10(1):31–4.
3. Branch JD. Effect of creatine supplementation on body composition and performance: a meta-analysis. Int J Sport Nutr Exerc Metab. 2003;13(2):198–226.
4. Kutz MR, Gunter MJ. Creatine monohydrate supplementation on body weight and percent body fat. J Strength Cond Res. 2003;17(4):817–21.
5. Bender A, Samtleben W, Elstner M, Klopstock T. Long-term creatine supplementation is safe in aged patients with Parkinson disease. Nutr Res. 2008;28(3):172–8.
6. de Silva AD, Pertille A, Barbosa CG, de Oliveira Silva JA, de Jesus DV, Ribeiro AG, et al. Effects of creatine supplementation on renal function: a systematic review and meta-analysis. J Ren Nutr. 2019;29(6):480–9.
7. Avgerinos KI, Spyrou N, Bougioukas KI, Kapogiannis D. Effects of creatine supplementation on cognitive function of healthy individuals: a systematic review of randomized controlled trials. Exp Gerontol. 2018;15(108):166–73.
8. Rae C, Digney AL, McEwan SR, Bates TC. Oral creatine monohydrate supplementation improves brain performance: a double-blind, placebo-controlled, cross-over trial. Proc Biol Sci. 2003;270(1529):2147–50.
9. Persky AM, Brazeau GA. Clinical pharmacology of the dietary supplement creatine monohydrate. Pharmacol Rev. 2001;53(2):161–76.
10. Lowe MTJ, Kim EH, Faull RLM, Christie DL, Waldvogel HJ. Dissociated expression of mitochondrial and cytosolic creatine kinases in the human brain: a new perspective on the role of creatine in brain energy metabolism. J Cereb Blood Flow Metab. 2013;33(8):1295–306.
11. Ainsley Dean PJ, Arikan G, Opitz B, Sterr A. Potential for use of creatine supplementation following mild traumatic brain injury. Concussion. 2017;2(2):CNC34.
12. Clark JF, Cecil KM. Diagnostic methods and recommendations for the cerebral creatine deficiency syndromes. Pediatr Res. 2015;77(3):398–405.
13. Brosnan ME, Brosnan JT. The role of dietary creatine. Amino Acids. 2016;48(8):1785–91.
14. Balestrino M, Adriano E. Beyond sports: efficacy and safety of creatine supplementation in pathological or paraphysiological conditions of brain and muscle. Med Res Rev. 2019;39(6):2427–59.
15. Harris RC, Söderlund K, Hultman E. Elevation of creatine in resting and exercised muscle of normal subjects by creatine supplementation. Clin Sci. 1992;83(3):367–74.
16. Schedel JM, Tanaka H, Kiyonaga A, Shindo M, Schutz Y. Acute creatine ingestion in human: consequences on serum creatine and creatinine concentrations. Life Sci. 1999;65(23):2463–70.
17. Dechent P, Pouwels PJ, Wilken B, Hanefeld F, Frahm J. Increase of total creatine in human brain after oral supplementation of creatine-monohydrate. Am J Physiol. 1999;277(3):R698-704.
18. Turner CE, Russell BR, Gant N. Comparative quantification of dietary supplemented neural creatine concentrations with (1)H-MRS peak fitting and basis spectrum methods. Magn Reson Imaging. 2015;33(9):1163–7.
19. Lyoo IK, Kong SW, Sung SM, Hirashima F, Parow A, Hennen J, et al. Multinuclear magnetic resonance spectroscopy of high-energy phosphate metabolites in human brain following oral supplementation of creatine-monohydrate. Psychiatry Res. 2003;123(2):87–100.
20. Wilkinson ID, Mitchel N, Breivik S, Greenwood P, Griffiths PD, Winter EM, et al. Effects of creatine supplementation on cerebral white matter in competitive sportsmen. Clin J Sport Med. 2006;16(1):63–7.
21. Merege-Filho CAA, Otaduy MCG, de Sá-Pinto AL, de Oliveira MO, de Souza GL, Hayashi APT, et al. Does brain creatine content rely on exogenous creatine in healthy youth? A proof-of-principle study. Appl Physiol Nutr Metab. 2017;42(2):128–34.
22. Solis MY, Artioli GG, Otaduy MCG, da Leite CC, Arruda W, Veiga RR, et al. Effect of age, diet, and tissue type on PCr response to creatine supplementation. J Appl Physiol. 2017;123(2):407–14.
23. Benton D, Donohoe R. The influence of creatine supplementation on the cognitive functioning of vegetarians and omnivores. Br J Nutr. 2011;105(7):1100–5.
24. Solis MY, de Salles PV, Artioli GG, Roschel H, Otaduy MC, Gualano B. Brain creatine depletion in vegetarians? A cross-sectional 1H-magnetic resonance spectroscopy (1H-MRS) study. Br J Nutr. 2014;111(7):1272–4.
25. Burke DG, Chilibeck PD, Parise G, Candow DG, Mahoney D, Tarnopolsky M. Effect of creatine and weight training on muscle creatine and performance in vegetarians. Med Sci Sports Exerc. 2003;35(11):1946–55.
26. Turner CE, Byblow WD, Gant N. Creatine supplementation enhances corticomotor excitability and cognitive performance during oxygen deprivation. J Neurosci. 2015;35(4):1773–80.
27. Hultman E, Söderlund K, Timmons JA, Cederblad G, Greenhaff PL. Muscle creatine loading in men. J Appl Physiol. 1996;81(1):232–7.
28. McKenna MJ, Morton J, Selig SE, Snow RJ. Creatine supplementation increases muscle total creatine but not maximal intermittent exercise performance. J Appl Physiol. 1999;87(6):2244–52.
29. Preen D, Dawson B, Goodman C, Beilby J, Ching S. Creatine supplementation: a comparison of loading and maintenance protocols on

creatine uptake by human skeletal muscle. Int J Sport Nutr Exerc Metab. 2003;13(1):97–111.
30. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. Int J Surg. 2011;9(8):672–7.
31. Wechsler D. Manual for the Wechsler Adult Intelligence Scale. 1955;110. Available from:https://psycnet.apa.org/fulltext/1955-07334-000.pdf.
32. Brickenkamp R. Test d2: Aufmerksamkeits-Belastungs-Test Hogrefe. Testzentrale: Göttingen, Germany; 2002.
33. Reitan RM. Validity of the trail making test as an indicator of organic brain damage. Percept Mot Skills. 1958;8(3):271–6.
34. Schellig D. Block-tapping-test. Swets Test Services Frankfurt; 1997.
35. Lux S, Helmstaedter C, Lendt M. Verbaler Lern- und Merkfähigkeitstest: VLMT ; Manual. Beltz-Test; 2001. 81 p.
36. Benedict RHB, Schretlen D, Groninger L, Dobraski M, Shpritz B. Revision of the brief visuospatial memory test: studies of normal performance, reliability, and validity. Psychol Assess. 1996;8(2):145.
37. Bäumler G, Stroop JR. Farbe-Wort-Interferenztest nach JR Stroop (FWIT). Verlag für Psychologie: Hogrefe; 1985.
38. Aschenbrenner S, Tucha O, Lange KW. Regensburger Wortflüssigkeits-Test: RWT. Verlag für Psychologie: Hogrefe; 2000.
39. Lehrl S. Manual zum MWT-B. Balingen: Spitta-Verlag; 2005. p. 43.
40. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods. 2007;39(2):175–91.
41. Stoet G. PsyToolkit: a software package for programming psychological experiments using Linux. Behav Res Methods. 2010;42(4):1096–104.
42. Stoet G. PsyToolkit: a novel web-based method for running online questionnaires and reaction-time experiments. Teach Psychol. 2017;44(1):24–31.
43. Dienes Z. Using Bayes to get the most out of non-significant results. Front Psychol. 2014;29(5):781.
44. Wilcox RR. Introduction to robust estimation and hypothesis testing. London: Academic Press; 2011. p. 608.
45. Field A. Discovering statistics using IBM SPSS Statistics. London: SAGE; 2013. p. 952.
46. Colling L. ljcolling/go-bayesfactor. 2021. Available from: https://zenodo.org/record/4642331.
47. Monaco M, Costa A, Caltagirone C, Carlesimo GA. Forward and backward span for verbal and visuo-spatial data: standardization and normative data from an Italian adult population. Neurol Sci. 2013;34(5):749–54.
48. van Doorn J, van den Bergh D, Böhm U, Dablander F, Derks K, Draws T, et al. The JASP guidelines for conducting and reporting a Bayesian analysis. Psychon Bull Rev. 2021;28(3):813–26.
49. Forbes SC, Cordingley DM, Cornish SM, Gualano B, Roschel H, Ostojic SM, et al. Effects of creatine supplementation on brain function and health. Nutrients. 2022;14(5).https://doi.org/10.3390/nu14050921.
50. Dolan E, Gualano B, Rawson ES. Beyond muscle: the effects of creatine supplementation on brain creatine, cognitive processing, and traumatic brain injury. EJSS. 2018;7:1–14.
51. Roschel H, Gualano B, Ostojic SM, Rawson ES. Creatine supplementation and brain health. Nutrients. 2021;13(2).https://doi.org/10.3390/nu13020586.
52. Borchio L, Machek SB, Machado M. Supplemental creatine monohydrate loading improves cognitive function in experienced mountain bikers. J Sports Med Phys Fitness. 2020;60(8):1168–70.
53. Pires LAM, Forbes SC, Candow DG, Machado M. Creatine supplementation on cognitive performance following exercise in female Muay Thai athletes. NeuroSports. 2021;1(1):6.
54. Ling J, Kritikos M, Tiplady B. Cognitive effects of creatine ethyl ester supplementation. Behav Pharmacol. 2009;20(8):673–9.
55. McMorris T, Mielcarz G, Harris RC, Swain JP, Howard A. Creatine supplementation and cognitive performance in elderly individuals. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn. 2007;14(5):517–28.
56. Hammett ST, Wall MB, Edwards TC, Smith AT. Dietary supplementation of creatine monohydrate reduces the human fMRI BOLD signal. Neurosci Lett. 2010;479(3):201–5.
57. VAN Cutsem J, Roelands B, Pluym B, Tassignon B, Verschueren JO, De Pauw K, et al. Can creatine combat the mental fatigue-associated decrease in visuomotor skills? Med Sci Sports Exerc. 2020;52(1):120–30.
58. Watanabe A, Kato N, Kato T. Effects of creatine on mental fatigue and cerebral hemoglobin oxygenation. Neurosci Res. 2002;42(4):279–85.
59. Samadi M, Askarian A, Shirvani H, Shamsoddini A, Shakibaee A, Forbes SC, et al. Effects of four weeks of beta-alanine supplementation combined with one week of creatine loading on physical and cognitive performance in military personnel. Int J Environ Res Public Health. 2022;19(13).https://doi.org/10.3390/ijerph19137992.
60. Rawson ES, Lieberman HR, Walsh TM, Zuber SM, Harhart JM, Matthews TC. Creatine supplementation does not improve cognitive function in young adults. Physiol Behav. 2008;95(1–2):130–4.
61. Alves CRR, Merege Filho CAA, Benatti FB, Brucki S, Pereira RMR, de Sá Pinto AL, et al. Creatine supplementation associated or not with strength training upon emotional and cognitive measures in older women: a randomized double-blind study. PLoS One. 2013;8(10):e76301.
62. Ostojic SM, Korovljev D, Stajer V. Dietary creatine and cognitive function in U.S. adults aged 60 years and over. Aging Clin Exp Res. 2021;33(12):3269–74.
63. Giem P, Beeson WL, Fraser GE. The incidence of dementia and intake of animal products: preliminary findings from the Adventist Health Study. Neuroepidemiology. 1993;12(1):28–36.
64. McMorris T, Harris RC, Swain J, Corbett J, Collard K, Dyson RJ, et al. Effect of creatine supplementation and sleep deprivation, with mild exercise, on cognitive and psychomotor performance, mood state, and plasma concentrations of catecholamines and cortisol. Psychopharmacology. 2006;185(1):93–103.
65. Postle BR, Pasternak T. Short term and working memory. Encyclopedia of Neuroscience San Diego, CA: Elsevier; 2009. p. 783–9.
66. Gignac GE. Raven's is not a pure measure of general intelligence: Implications for g factor theory and the brief measurement of g. Intelligence. 2015;1(52):71–9.
67. Roth B, Becker N, Romeyke S, Schäfer S, Domnick F, Spinath FM. Intelligence and school grades: a meta-analysis. Intelligence. 2015;1(53):118–37.
68. Prokopidis K, Giannos P, Triantafyllidis KK, Kechagias KS, Forbes SC, Candow DG. Effects of creatine supplementation on memory in healthy individuals: a systematic review and meta-analysis of randomized controlled trials. Nutr Rev. 2023;81(4):416–27.
69. Prokopidis K, Giannos P, Triantafyllidis KK, Kechagias KS, Forbes SC, Candow DG. Author's reply: Letter to the Editor: Double counting due to inadequate statistics leads to false-positive findings in "Effects of creatine supplementation on memory in healthy individuals: a systematic review and meta-analysis of randomized controlled trials." Nutr Rev. 2023.https://doi.org/10.1093/nutrit/nuac111.
70. Forsberg AM, Nilsson E, Werneman J, Bergström J, Hultman E. Muscle composition in relation to age and sex. Clin Sci. 1991;81(2):249–56.
71. McCully KK, Forciea MA, Hack LM, Donlon E, Wheatley RW, Oatis CA, et al. Muscle metabolism in older subjects using 31P magnetic resonance spectroscopy. Can J Physiol Pharmacol. 1991;69(5):576–80.
72. McCully KK, Fielding RA, Evans WJ, Leigh JS Jr, Posner JD. Relationships between in vivo and in vitro measurements of metabolism in young and old human calf muscles. J Appl Physiol. 1993;75(2):813–9.
73. Smith SA, Montain SJ, Matott RP, Zientara GP, Jolesz FA, Fielding RA. Creatine supplementation and age influence muscle metabolism during exercise. J Appl Physiol. 1998;85(4):1349–56.
74. Kent-Braun JA, Ng AV, Young K. Skeletal muscle contractile and noncontractile components in young and older women and men. J Appl Physiol. 2000;88(2):662–8.
75. Rawson ES, Clarkson PM, Price TB, Miles MP. Differential response of muscle phosphocreatine to creatine supplementation in young and old subjects. Acta Physiol Scand. 2002;174(1):57–65.
76. Smith-Ryan AE, Cabre HE, Eckerson JM, Candow DG. Creatine Supplementation in Women's Health: A Lifespan Perspective. Nutrients. 2021;13(3). Available from: https://doi.org/10.3390/nu13030877.
77. Harris RC, Nevill M, Harris DB, Fallowfield JL, Bogdanis GC, Wise JA. Absorption of creatine supplied as a drink, in meat or in solid form. J Sports Sci. 2002;20(2):147–51.
78. Bodner-Montville J, Ahuja JKC, Ingwersen LA, Haggerty ES, Enns CW, Perloff BP. USDA food and nutrient database for dietary studies: released on the web. J Food Compost Anal. 2006;1(19):S100–7.

79. Steenge GR, Simpson EJ, Greenhaff PL. Protein- and carbohydrate-induced augmentation of whole body creatine retention in humans. J Appl Physiol. 2000;89(3):1165–71.
80. Ott A, Hugi A, Baumgartner M, Chaintreau A. Sensory investigation of yogurt flavor perception: mutual influence of volatiles and acidity. J Agric Food Chem. 2000;48(2):441–50.
81. Hennighausen RH. Die neue Trinkwasserverordnung - TrinkwV 2001. Gesundheitswesen. 2001;63(11):724–30.
82. Jäger R, Purpura M, Shao A, Inoue T, Kreider RB. Analysis of the efficacy, safety, and regulatory status of novel forms of creatine. Amino Acids. 2011;40(5):1369–83.

Publisher's Note

APPENDIX E: Publication Study 3

Sandkühler, J. F.*, Kahl, F.*, Sadurska, M. Z., Brietbart, P., Greenberg, S., Brauner, J. (2025). The Immediate Impact of App-Based Psychotherapeutic Exercises on Anxiety: An RCT. ***Depression and Anxiety.*** https://doi.org/10.1155/da/5586831

Wiley
Depression and Anxiety
Volume 2025, Article ID 5586831, 12 pages
https://doi.org/10.1155/da/5586831

WILEY

Research Article

The Immediate Impact of App-Based Psychotherapeutic Exercises on Anxiety: An RCT

Julia Fabienne Sandkühler,[1] **Fabian Kahl**,[2] **Magda Zena Sadurska**,[3,4] **Peter Brietbart**,[5] **Spencer Greenberg**,[4,5] **and Jan Brauner**[6]

[1]*Department of Psychology, University of Bonn, Bonn, North Rhine Westphalia, Germany*
[2]*Department of Child and Adolescent Psychiatry and Psychotherapy, University Medical Center Hamburg-Eppendorf, Hamburg, Germany*
[3]*Insthea Ltd., London, Greater London, UK*
[4]*Spark Wave, New York City, New York, USA*
[5]*Mind Ease Labs Ltd., Brighton, East Sussex, UK*
[6]*Department of Computer Science, University of Oxford, Oxford, Oxfordshire, UK*

Correspondence should be addressed to Julia Fabienne Sandkühler; jf.sandkuehler@gmail.com

Received 29 May 2024; Accepted 30 January 2025

Academic Editor: Richard Xu

Background: Despite the growing integrative trend in psychotherapy, few studies have examined the potential for immediate anxiety relief of many different psychotherapeutic exercises side by side under the same conditions. This information might be important to enhance engagement and self-efficacy, stop negative feedback loops, and prevent avoidant or destructive behavior during crises. Technology-based psychotherapeutic exercises are of particular interest because they are accessible and scalable.
Methods: This parallel, double-blind, randomized trial ($N = 1092$) compared 12 psychotherapeutic exercises of the Mind Ease app against a reading control and a measurement-only control. Efficacy was measured with a custom scale validated against the state subscale of the State–Trait Anxiety Inventory.
Results: Each of the 12 exercises significantly reduced anxiety more than controls ($p = 0.018$ to <0.001, $\eta^2_p = 0.06$ to 0.37, $d =$ 0.5 to 1.5, d [95% CI] for all exercises together vs. reading control = 0.8 [0.6; 1.0], and vs. measurement-only control = 0.8 [0.6; 1.0]). Exercises employing cognitive restructuring had effect sizes d [95% CI] of 0.5 [0.2; 0.8], 0.7 [0.3; 1.0], and 0.9 [0.6; 1.2], diaphragmatic breathing of 0.6 [0.3; 0.9], gratitude practice of 0.8 [0.5; 1.1], positive expressive writing of 1.1 [0.7; 1.4], progressive muscle relaxation of 1.3 [0.9; 1.6], guided imagery of 1.3 [1.0; 1.6], and mindfulness of 0.9 [0.6; 1.2], 1.0 [0.7; 1.3], 1.2 [0.9; 1.5], and 1.5 [1.2; 1.9]. Twenty-eight comparisons between exercises (42%) had $p < 0.05$, nine met the Bonferroni-adjusted threshold of $p < 0.0008$.
Conclusions: The 12 psychotherapeutic exercises proved effective at immediately mitigating anxiety. Differences between exercises were substantial even within categories. Mindfulness tended to have a larger effect than cognitive restructuring.

Trial Registration: ClinicalTrials.gov identifier: NCT05850975

Keywords: acceptance and commitment therapy; access; anxiety; app; cognitive behavioral therapy; eHealth; mHealth; mindfulness; psychotherapy; RCT

1. Introduction

Anxiety is one of the most prevalent mental health concerns worldwide, with high personal, social, and economic costs [1]. Even in the subclinical range, anxiety poses considerable distress and is often accompanied by unhealthy coping mechanisms, such as alcohol or other substance misuse. Learning healthy coping strategies is essential for preventing anxiety disorders and substance misuse disorders. Although numerous evidence-based therapeutic approaches exist [2–4], many individuals face barriers to accessing in-person treatment, including availability of therapists, cost, stigma associated with mental health treatment, and logistical challenges, such as transportation and time [5]. Consequently, digital mental health interventions such as mobile apps have gained increasing attention as a way to deliver evidence-based strategies widely and cost-effectively [6].

Within the realm of therapeutic strategies for anxiety, a broad range of approaches have demonstrated efficacy in both research and clinical practice [7–11]. Cognitive restructuring techniques focus on identifying and challenging maladaptive thought patterns, modifying how individuals interpret and respond to anxiety-provoking situations. Mindfulness-based interventions involve cultivating nonjudgmental awareness of the present moment and promoting acceptance of distressing thoughts and sensations. Relaxation techniques, such as progressive muscle relaxation or diaphragmatic breathing, aim to reduce physiological arousal and tension. Positive expressive writing and gratitude practice encourage deeper reflection on emotions and positive experiences. Although these methods often share the common goal of reducing anxiety, they differ in their underlying mechanisms and immediate versus long-term aims.

Despite the growing integrative trend in psychotherapy [12], few studies compare many exercises side by side with the same population, duration and setting, making a direct comparison between them difficult [11, 13–19]. Understanding the immediate effects of these interventions is important because many people seeking help are motivated by the prospect of quick relief, and immediate benefits may bolster ongoing engagement. Engagement strongly predicts positive psychotherapy outcomes [20] and is a concern of both mental health apps and in-person psychotherapy [21, 22]. Immediate effects might also be important for self-efficacy, to stop negative feedback loops, and to prevent avoidant or destructive behavior during crises.

The present study addresses these gaps by comparing the immediate effects of 12 app-based exercises in a single digital platform, with the same population, duration, and setting. The techniques employed in the exercises include several variations of cognitive restructuring, several variations of mindfulness, guided imagery, progressive muscle relaxation, diaphragmatic breathing, positive expressive writing, and gratitude practice.

We did not have hypotheses about which exercises would be most effective. The focus of this paper—the comparison between exercises—is therefore exploratory. Our preregistered hypotheses were that (1) each exercise would be more effective at reducing immediate anxiety than the control conditions and (2) the average exercise would be more effective at reducing immediate anxiety than the control conditions.

2. Methods

2.1. Trial Design. We conducted a parallel, double-blind, randomized, controlled trial. Participants were randomly allocated to either one of the 12 exercises offered by the Mind Ease app or one of two control conditions, namely reading an informational text about anxiety or carrying on with regular activities. The allocation ratio was 1 for each of the 12 exercises and 1.5 for each of the two control groups. The trial design and participant flow are summarized in Figure 1. We follow the CONSORT reporting guidelines for parallel-group randomized trials [23]. Before publication, the study was disseminated as a preprint [24].

2.2. Recruitment. Ethical approval was obtained from the Medical Sciences Interdivisional Research Ethics

Committee (MS IDREC) of the University of Oxford (reference number R82884/RE001). Participants gave informed consent to participate before being enrolled (see Appendix S1 for the recruitment text and consent form).

Participants were recruited online on the Positly platform, which uses Amazon Mechanical Turk, and tested on GuidedTrack between March 11, 2023 and July 19, 2023. Positly automatically used seven methods to help ensure that the study participants it recruited were high quality and not spammers: (1) It monitored the IP address of workers to prevent multiple submissions from the same IP address. (2) It kept track of whether web traffic was coming from a source that was previously known to have had high rates of spam. (3) It periodically checked that participants were paying attention. (4) It checked that participants' prior work had a low rejection rate. (5) It used CAPTCHAs to block automated bots. (6) It periodically checked that participants were able and willing to follow simple instructions. (7) It checked that participants spoke English.

Only participants aged 18 or above with the United States as their registered location on Amazon Mechanical Turk were able to take part. Participants completed a brief screening survey, which assessed their anxiety levels. Only participants who responded to at least two of the three slider questions (see Section 2.5) by saying they felt "quite bad/worried/tense" (score of 67 points) or worse were invited to take part in the study. For participant characteristics, see Section 3. The screening, interventions, and assessments were done automatically on GuidedTrack.

2.3. Interventions. There were 12 exercises and two control interventions. In the app, the exercises have names different from those used here. For a list of correspondence see Appendix S1.

2.4. Exercises

2.4.1. Cognitive Distortion List. This exercise consisted of cognitive restructuring with the aid of a cognitive distortion list. Participants were encouraged to identify and challenge distressing thoughts or beliefs that trigger anxiety. A list of common cognitive distortions (e.g., all-or-nothing thinking, overgeneralization) was provided to participants. Participants were guided to recognize which distortions they engaged in, to challenge these distorted thoughts, and to replace them with more accurate, balanced, and helpful thoughts.

2.4.2. Silver Lining. This exercise employed cognitive restructuring, reinterpreting a situation by identifying positive aspects or results.

2.4.3. Positive Expressive Writing. In this exercise, participants were guided to recall and write about a positive, joy-filled experience from their past, using as much detail (smell, feelings, thoughts, etc.) as possible.

2.4.4. Gratitude Practice. In this exercise, participants were guided to reflect on three things they were grateful for.

2.4.5. Guided Imagery. In this exercise, participants were guided to visualize themselves in a relaxing environment, while smiling (including their eyes) for the last minute.

2.4.6. Anxiety-Excitement Reappraisal. This exercise initiated with participants selecting a distressing concern they thought they might be able to feel better about. Participants then learned about the vicious cycle of emotion, bodily sensations, and thoughts driving panic attacks and were reminded that their feelings are not harmful. Subsequently, participants were guided to recognize the similarities between physical symptoms of anxiety and excitement and to reframe their understanding of their physical symptoms from anxiety ("I'm feeling anxious") to excitement ("I'm feeling excited"). Anxiety-excitement reappraisal is based on the understanding that anxiety and excitement are both states of high arousal but differ primarily in their appraisal as negative or positive, respectively. The technique is a form of cognitive restructuring. Finally, participants were asked to write down and execute an action based on their values.

2.4.7. Mindful Breathing. Participants were directed to focus their attention on their breath, observing each inhalation and exhalation without trying to alter the breathing pattern. When the mind wanders, which it often does, the instruction was to gently bring the attention back to the breath, without judgement.

2.4.8. Dropping Anchor. In this exercise commonly known as dropping anchor [25], participants were led through a series of mindfulness exercises, beginning with sound awareness, progressing to body sensations, and then to visual details. The aim was not to identify or analyze these sensations, but simply to notice them as they arise and fade, anchoring the individual in the present moment.

2.4.9. Body Scan. This exercise consisted mostly of mindful confrontation with bodily symptoms of anxiety combined with a "body scan" (shifting attention gradually through the body; [25]). Participants were guided to identify the specific physical sensations associated with their anxiety, to locate these sensations in their bodies, and to observe these sensations without trying to change them. The idea is that by approaching these symptoms with curiosity rather than fear, individuals can develop a different relationship with their anxiety, one characterized by acceptance rather than avoidance [26]. The exercise also includes a brief appearance of diaphragmatic breathing (see "diaphragmatic breathing" exercise) and cognitive restructuring (at the end of the exercise, participants were asked to reflect on how their bodily symptoms of anxiety might be informative and trying to help them).

2.4.10. Leaves on a Stream. This exercise included the Leaves on a Stream exercise, as well as guidance to mindfully observe thoughts and feelings without this visualization. Leaves on a Stream is a mindfulness-based exercise used to promote detachment (also called defusion) from one's thoughts and foster acceptance [27]. The idea is to see thoughts as temporary mental events, not reality-defining facts, reducing overidentification with thoughts. Participants visualized themselves by a stream, placing each arising thought on a leaf and observing it come and go. The cognitive restructuring method of challenging thoughts ("Is it actually true?") briefly appeared too.

2.4.11. Diaphragmatic Breathing. The exercise involved inhaling slowly and deeply through the nose, allowing the diaphragm to rise and the lungs to fully inflate. This was followed by a slow, controlled exhalation through the mouth, during which the diaphragm fell. The process was repeated many times.

2.4.12. Progressive Muscle Relaxation. Progressive muscle relaxation involves systematically tensing and then releasing different muscle groups throughout the body, leading to a state of overall physical relaxation.

2.4.13. Control Conditions. There were two control conditions: One control group was instructed to do what they would ordinarily do for 7 minutes (the expected average duration of the interventions), until a bell chimed (measurement-only control). Participants were not able to continue before the instructed time of 7 minutes had been reached. The other control group was given an informational text about anxiety to read [28] (reading control). See Appendix S1 for the text and exact instructions given to participants.

2.4.14. Similarity of Interventions and Treatment Expectations. The reading control was superficially similar to the exercises because the exercises also included reading texts. Measurement-only participants had the two anxiety measurements as well as the delay between them in common with the other groups. The measures, texts, and exercises were on the same platform and in the same visual style for all participants. The exercises and control conditions had similar durations (in minutes: exercises $M = 9.8$, $SD = 4.1$, reading control $M = 8.9$ min, $SD = 4.1$, measurement-only control $M = 12.6$, $SD = 3.5$). Participants were not told there would be control groups, but they were informed before taking part that the study was about online interventions to reduce negative feelings. It is therefore likely that the measurement-only participants realized that they were in a control group. In contrast, the reading control group was less obviously a control group—participants were given a long text about anxiety to read (Appendix S1), which was likely perceived as a psychoeducational intervention. Nevertheless, both control conditions likely produced lower treatment expectations than the exercises. The control groups were important to control for the time passed between the two measures. Participants knew they would be compensated regardless of how they

responded, so there was no financial incentive to lie. The fact that the study was conducted fully online and automatically without staff being present likely minimized the pressure to provide expected or socially desirable responses.

2.4.15. Changes After the Trial Commenced. Originally, participants could also be randomized into a group that received an exercise selected by Mind Ease's machine learning algorithm (rather than random allocation to an exercise). However, after the first 82 participants had been assigned to this group, we found a bug in the data processing for the ML algorithm. This meant that participants had not received the right recommendations. This bug only affected the ML algorithm in the study, and not the algorithm used in the public version of the Mind Ease app. We discarded the data of these 82 participants. As recruiting was slower and more expensive than expected (see Section 2.6), we would not have been able to power this group to an acceptable extent anymore, so we stopped the allocation to this group.

2.5. Outcomes. Our primary outcome was the anxiety score calculated as the average response to three slider questions (picture in Appendix S1). The slider questions were used because they took a very short time, were already part of the app, and we wished to capture the experience of real users as much as possible. These questions were asked before and after the interventions. In these three questions, participants were asked to report how they were feeling at that moment by moving a continuous slider that was accompanied by seven labels appearing depending on the position of the slider. The labels for the tense-relaxed slider were "very tense," "quite tense," "somewhat tense,""neither tense nor relaxed," "somewhat relaxed," "quite relaxed," and "very relaxed." "Neither tense nor relaxed" corresponded to the exact middle point of the scale, while the other six labels corresponded to an area covering one-sixth of the scale respectively. In the same fashion, the other two sliders ranged from "very worried" to "very calm" and "very bad" to "very good," respectively. The range of possible scores for the primary outcome was 0 (no anxiety) to 100 (high anxiety). Participants did not see these numbers on their interface.

We conducted a separate prestudy with 199 adult US participants to determine (1) the convergent validity with the state subscale of the State-Trait Anxiety Inventory (STAI) and (2) the internal consistency of the scale. The STAI state subscale consists of 20 self-report items (e.g., "I feel calm") on a four-point Likert scale from "not at all" to "very much so." Participants in the prestudy were recruited on Positly December 4–15, 2022 (see Appendix S1 for consent form and recruitment text). Their mean age was 39.6 years (SD = 11.7), 44% were female, and their mean STAI state subscale score was 37.3 (SD = 13.5). The three-slider scale correlated positively and highly significantly with the STAI state subscale, $r\,(197) = 0.872$, $p < 0.001$, without excluding any outliers. After excluding one obvious outlier based on visual inspection of the scatterplot (three-slider scale score of 27 and STAI state subscale score of 76), the result was very similar, $r\,(196) = 0.893$, $p < 0.001$. See Appendix S1 for a scatterplot of the data including the outlier. The results of the linear regression (excluding the outlier) are STAI state subscale score = three-slider scale score × 0.55 + 20.66. We used this equation to report the results of the main study in STAI state subscale score equivalent, so that they are in a unit that is familiar to readers. Cronbach's alpha, a measure of internal consistency, was 0.94 in the prestudy, with interitem correlations between 0.81 and 0.90. In the main study, Cronbach's alpha was 0.70 (pre) and 0.90 (post). These high levels of internal consistency are to be expected when measuring interdependent aspects of state anxiety and are comparable to those found for other state anxiety scales [29, 30]. Participant feedback on ease of use, clarity, and content validity indicated high satisfaction in all of these areas. For a comparison with similar scales and details on the participant feedback survey, see Appendix S1.

2.6. Sample Size. We originally preregistered a total sample size of 5550 completed participants (370 participants per exercise group, 555 participants per control group). However, when we had recruited 582 participants (39 per exercise group and 57 per control group), we noticed that recruitment was substantially slower and more expensive than expected, so we would not be able to recruit the preregistered number of participants. An interim analysis and new sample size calculation showed that the effect sizes were larger than expected, so we would nevertheless be able to power the study adequately. We therefore updated the preregistration to a recruitment goal of 1126 participants (75 per exercise group and 113 per control group), that is, about twice the number of participants we had recruited at that point. In the end, 1108 participants completed the study. As preregistered, we excluded those with a baseline anxiety score below 50, leaving us with 1092 participants (on average 73 per active treatment and 112 per control group). Finally, after also excluding participants based on a time criterion (see Section 2.8), there were 1054 participants, on average 106 in each control group and 70 in each active treatment group.

2.7. Randomization and Blinding. GuidedTrack was used to randomly assign the interventions to participants using simple randomization. Staff did not interact with the participants. Participants were not told whether they were in an intervention or in a control group. To make participants think, they were not in a control group despite not receiving any intervention, measurement-only participants were told that "we would like to test changes in mood over short periods of time."

2.8. Statistical Methods. Following the methods outlined in our preregistration, we conducted mixed ANOVAs with time (pre- vs. post-intervention) as the within-subjects variable, intervention (exercise vs. control) as the between-subjects variable and anxiety score as the dependent variable. We applied the Greenhouse–Geisser correction to all our analyses but the correction did not change any value. In addition, we performed independent *t*-tests of the improvement scores, which resulted in the same *p*-value (two-tailed) as the interaction of the ANOVA. It is a different way of doing the same analysis and was conducted to aid interpretation. See Appendix S1 for SE and SD of the pre and post scores.

In addition, we checked the robustness of our normality-assuming ANOVAs by performing 5%- and 20%-winsorised ANOVAs as well as a robust ANOVA which uses trimming

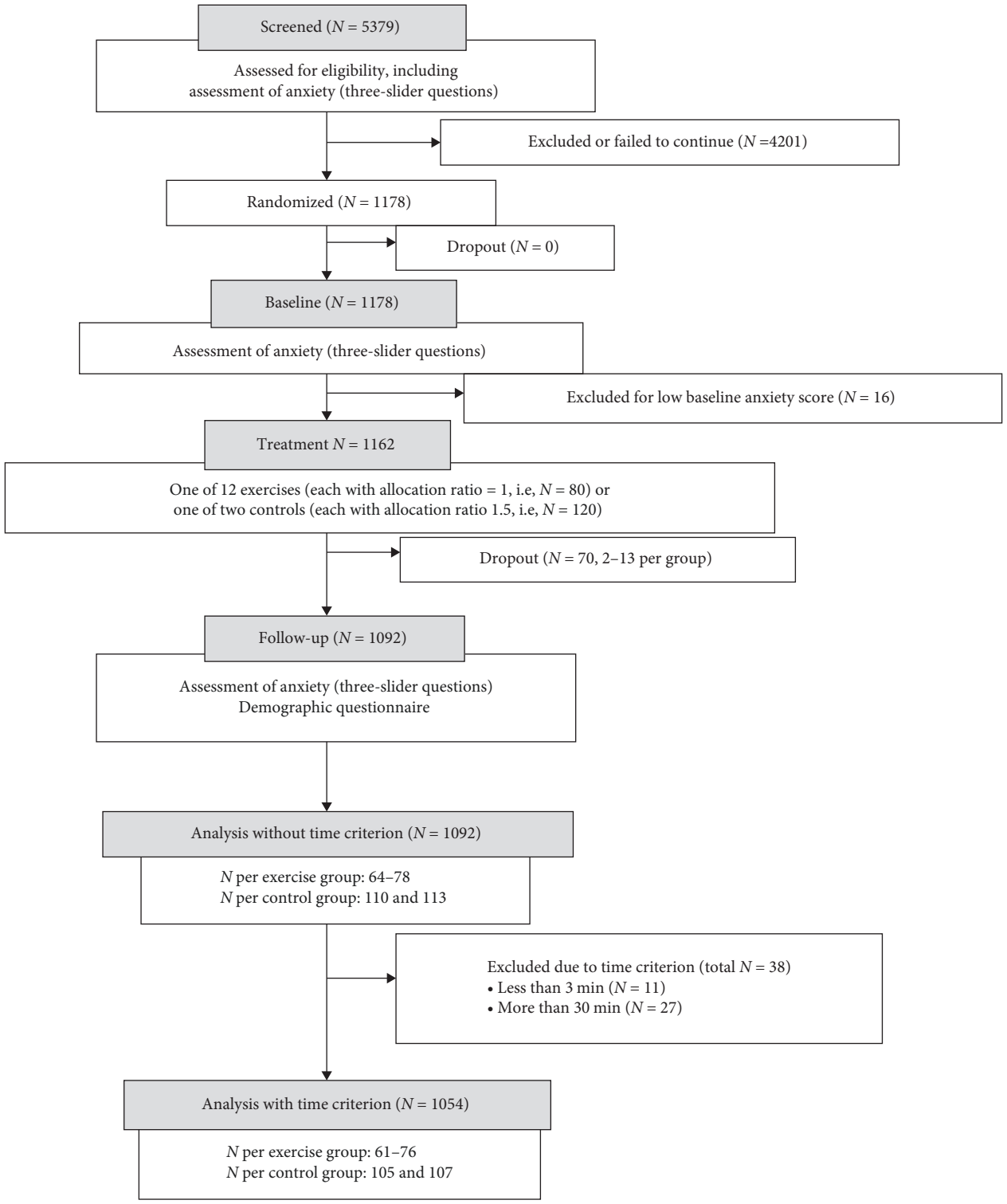


Figure 1: Participant flow through the study. For dropouts per group see Appendix S1.

and bootstrapping (performed with the sppi functions in the WRS2 R package) [31, 32].

As preregistered, we excluded participants with a baseline anxiety level below 50 on the three-slider scale (equivalent to a score of 48 on the STAI state subscale). We performed analyses including participants who took less than 3 and more than 30 minutes, as well as without these participants (time criterion). We implemented this time criterion because participants who spent too little time on the exercise cannot have performed it properly. Participants who took more than 30 minutes (in one case, over 7000 minutes) likely did other things than just the treatment between the two anxiety assessments. In

addition, anxiety fluctuates over time, so a large variation in time taken brings in noise. The results of the analyses with and without the time criterion are essentially the same. We did not think of preregistering a time criterion, but because it makes sense to have it, we focus on the analysis with the time criterion in this paper and report the results without the time criterion in Appendix S1. Participants were included in the analysis irrespective of the quality of their write-in responses, which was generally high (see Section 2.4).

There was no way for participants to receive any other intervention than the originally assigned one, therefore the issue of how to analyze participants where this happened did not arise.

3. Results

3.1. Participant Flow. The participant flow through the study, including the numbers allocated to each intervention, is presented in Figure 1.

3.2. Participant Characteristics. For participant characteristics at baseline, see Table 1.

Almost half of the participants (46%, 480 participants) took either mental health medication or used psychotherapeutic help. The mental health medication participants reported taking consisted mostly of antidepressants (81%), followed by sleeping pills or minor tranquillizers (27%), mood stabilizers (20%), and antipsychotics (6%). The psychotherapeutic help participants reported using consisted mostly of seeing a mental health clinician such as a counselor or a psychologist (72%), followed by books or blogs (17%), apps (9%), and support groups (1%). Participants were asked to write in an open text box what type of therapy they received. In total, 4.9% of participants reported CBT and 0.6% reported ACT as the type of therapy. However, most participants only wrote "talk therapy," "psychotherapy," or similar, so these percentages likely largely underestimate participants' exposure to these types of therapy. A fourth (25.3%) of participants reported meditating for 5 minutes a day at least 1 day a week. Including participants with no meditation days, on average participants meditated for at least 5 minutes 1 day a week ($M = 0.87$, SD = 1.77). Participants' baseline anxiety was 73.7 on the three-slider scale (possible range: 0–100), corresponding to quite "tense/worried/bad." Based on the regression in our prestudy, this corresponds to a STAI state score of 61.2 (possible range: 20–80). It has been suggested that clinically significant anxiety starts at a STAI state subscale score of 40 [33, 34], and possibly at a score of 55 for older adults [35].

3.3. Adherence and Side Effects. Exercises involved writing into text boxes in the app and guided meditations. Participants were not able to continue before the instructed time for the meditations had been reached. The following exercises included textboxes to be filled out by participants: "identifying cognitive distortions," "anxiety-excitement reappraisal," "Leaves on a Stream," "gratitude practice," "positive expressive writing," "silver lining," and "body scan." It was not possible for participants to continue the study without entering something into the textbox, apart from the exercise "body scan." The answers entered into the textboxes showed an appropriate engagement with and execution of the exercise for all participants except one for the exercise Leaves on a Stream and one for the exercise gratitude practice In addition, 18 "body scan" participants finished the study without entering any text.

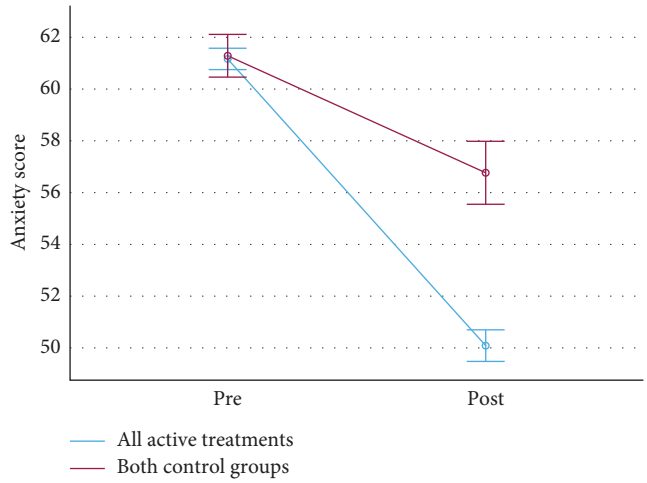


FIGURE 2: Pre- and post-intervention means of the anxiety scores for all exercises together and the two control groups together. The anxiety score is the STAI state subscale equivalent. The error bars represent 95% confidence intervals. STAI, State-Trait Anxiety Inventory.

Participants had the opportunity to give feedback on the study and many used this opportunity. No adverse effects were reported.

3.4. Effects of the Psychotherapeutic Exercises. In line with our hypotheses, the average intervention was significantly more effective than the control conditions at reducing anxiety (vs. reading control: $p < 0.001$, $\eta^2_p = 0.063$, $d = 0.8$; vs. measurement-only control: $p < 0.001$, $\eta^2_p = 0.059$, $d = 0.8$, Figure 2). In addition, in line with our hypotheses, each individual exercise was significantly more effective than the measurement-only control (Figure 3 and Table 2) as well as the reading control (see Appendix S1). This is true for the preregistered significance threshold unadjusted for multiple comparisons ($p < 0.05$), as well as for the significance threshold adjusted for multiple comparisons —$p < 0.004$ if adjusting only for the confirmatory comparisons between exercises and the measurement-only control, $p < 0.002$ if adjusting also for the exploratory comparisons between exercises and the reading control. This is very different from the 5% comparisons expected to have $p < 0.05$ by chance and thus strongly suggests that the exercises are effective. The skew for the pre- and post-distributions of interventions ranged from −0.4 to 0.4, with kurtosis ranging from −1.1 to 1.6 (see Appendix S1). The results were similar for winsorized as well as bootstrap and trimmed ANOVAs done as robustness checks (see Appendix S1). An exploratory analysis revealed that higher baseline anxiety was associated with a greater treatment effect, $r\ (1052) = 0.178$, $p < 0.001$.

Exploratory analyses (Figure 4) showed that the effects of the exercises differed more than would be expected by chance.

Table 1: Participant characteristics at baseline.

Characteristic	Total	Active treatments	Measurement-only control	Reading control
N	1054	842	107	105
Age in years: range (*M*, SD)	21–92 (39, 11)	21–92 (39, 11)	23–72 (39, 11)	22–73 (39, 11)
Sex (female, %)	66%	65%	69%	71%
Ethnicity (White, %)	77%	76%	78%	83%
Mental health medication (yes, %)	37%	36%	39%	38%
Psychotherapeutic help (yes, %)	26%	26%	30%	25%
Meditation days/week (*M*, SD)	0.87 (1.77)	0.81 (1.71)	1.43 (2.30)	0.76 (1.54)
Meditation (yes, %)	25.3%	24.8%	31.8%	22.9%
Anxiety (STAI state subscale score equivalent)	61.2 (6.1)	61.1 (6.1)	61.1 (6.2)	61.5 (6.1)

Note: Data are given as mean (standard deviation) or as percentage. To facilitate interpretation, the STAI state subscale score equivalent of the three-slider scale score is given, calculated based on our prestudy as STAI state subscale score = three-slider scale score × 0.55 + 20.66.
Abbreviation: STAI, State-Trait Anxiety Inventory.

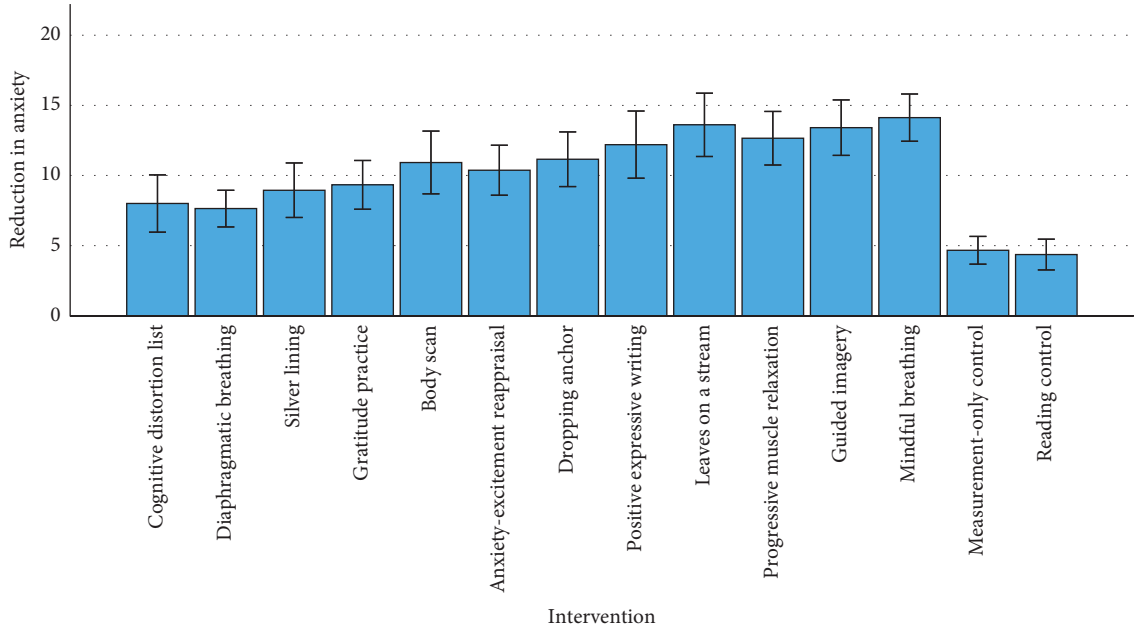


Figure 3: Reduction in anxiety, given in STAI state subscale equivalent, for the 12 exercises and two controls. Error bars represent 95% confidence intervals. For scatter plots and box plots of these data, see Appendix S1. STAI, State-Trait Anxiety Inventory.

One way of seeing that the differences between exercises seem to be different from what would be expected by chance is that Bonferroni-adjusting for 66 tests would give a significance threshold for p of $0.05/66 = 0.0008$, and there are nine comparisons that meet this adjusted significance threshold. Twenty-eight comparisons (42%) had $p < 0.05$. This is very different from the 5% comparisons expected to have $p < 0.05$ by chance and thus seems to suggest that there might be some real differences between the exercises.

4. Discussion

The present study examined and compared the immediate effects of 12 app-based exercises on anxiety levels. The exercises included several variations of cognitive restructuring, several variations of mindfulness, guided imagery, progressive muscle relaxation, diaphragmatic breathing, positive expressive writing, and gratitude practice.

In line with our hypotheses, our results revealed that the app interventions, when analyzed together as well as when analyzed separately, were significantly more effective than the control conditions at reducing immediate anxiety. The Cohen's d effect sizes for the different exercises ranged from 0.5 (traditionally considered "medium") to 1.5, with all but two of the 12 exercises over $d = 0.8$ (traditionally considered "large"). We consider the size of the improvements for all exercises in the app to be relevant. This interpretation is supported by a previous study which found $d = 0.5$ to be a clinically relevant reduction in anxiety [36], as well as roughly by a Cochrane review [37] using the rule of thumb of a 10% improvement being clinically relevant, which translates to $d = 0.67$ in our study. Our results align with the growing body of evidence indicating

Table 2: Comparing exercises to measurement-only control.

Exercise	N	ANOVA		*t*-test	Improv. diff. (*M*, SE)	Regression		Category
		p	η^2_p	*d* [95% CI]		beta	*p*	
All exercises	842	<0.001	0.059	0.8 [0.6; 1.0]	6.38 (0.55)	0.24	<0.001	NA
Cognitive distortion list	67	<0.001	0.058	0.5 [0.2; 0.8]	3.4 (1.2)	0.10	0.005	Cognitive restructuring
Diaphragmatic breathing	68	<0.001	0.072	0.6 [0.3; 0.9]	3.0 (0.8)	0.08	0.018	Diaphragmatic breathing
Silver lining	71	<0.001	0.094	0.7 [0.3; 1.0]	4.3 (1.1)	0.13	<0.001	Cognitive restructuring
Gratitude practice	71	<0.001	0.124	0.8 [0.5; 1.1]	4.7 (1.0)	0.14	<0.001	Gratitude practice
Body scan	68	<0.001	0.16	0.9 [0.6; 1.2]	6.3 (1.2)	0.17	<0.001	Mindfulness + exposure to bodily sensations of anxiety, (diaphragmatic breathing, cognitive restructuring)
Anxiety-excitement reappraisal	71	<0.001	0.17	0.9 [0.6; 1.2]	5.7 (1.0)	0.17	<0.001	Cognitive restructuring + exposure to bodily sensations of anxiety, (value-based action)
Dropping anchor	72	<0.001	0.191	1.0 [0.7; 1.3]	6.5 (1.1)	0.19	<0.001	Mindfulness
Positive expressive writing	61	<0.001	0.215	1.1 [0.7; 1.4]	7.5 (1.3)	0.22	<0.001	Positive expressive writing
Leaves on a stream	74	<0.001	0.264	1.2 [0.9; 1.5]	9.0 (1.3)	0.29	<0.001	Mindfulness + exposure to anxious thoughts, (cognitive restructuring)
Progressive muscle relaxation	69	<0.001	0.274	1.3 [0.9; 1.6]	8.0 (1.1)	0.25	<0.001	Progressive muscle relaxation
Guided imagery	76	<0.001	0.287	1.3 [1.0; 1.6]	8.7 (1.1)	0.27	<0.001	Imagery, (smiling)
Mindful breathing	74	<0.001	0.37	1.5 [1.2; 1.9]	9.5 (1.0)	0.30	<0.001	Mindfulness

Note: There were 107 participants in the measurement-only control group. The difference in improvement is given in STAI state subscale equivalent score (mean, standard error). Categories that played a major role are not in parentheses, categories that played only a minor role are in parenthesis. *P* and η^2_P are given of the interaction of the mixed ANOVAs with time (pre vs post intervention) as the within-subjects variable, intervention (exercise vs measurement-only control) as the between-subjects variable and anxiety score as the dependent variable. Cohen's *d* of the improvement score is given. The regression used STAI state subscale equivalent scores, with the improvement score as the outcome variable, and pre score and exercise as predictor (dummy variables with the measurement-only control group as the comparator). See the appendix for further details on the regression results, *F* and *df* of the ANOVA, and details on means, standard deviations and standard errors pre and post intervention separately.

	Cognitive distortion list	Diaphragmatic breathing	Silver lining	Gratitude practice	Body scan	Anxiety-excitement reappraisal	Dropping anchor	Positive expressive writing	Leaves on a Stream	Progressive muscle relaxation	Guided imagery	*d*	*M*
Cognitive distortion list	/	—	—	—	—	—	—	—	—	—	—	0.5	3.4
Diaphragmatic breathing	0.05, 0.766	/	—	—	—	—	—	—	—	—	—	0.6	3.0
Silver lining	0.11, 0.506	−0.19, 0.269	/	—	—	—	—	—	—	—	—	0.7	4.3
Gratitude practice	−0.17, 0.324	−0.26, 0.123	−0.05, 0.766	/		—	—	—	—	—	—	0.8	4.7
Body scan	−0.33, 0.057	−0.43, 0.013	−0.23, 0.185	−0.19, 0.265	/	—	—	—	—	—	—	0.9	6.3
Anxiety-excitement reappraisal	−0.30, 0.084	−0.42, 0.015	−0.18, 0.283	−0.14, 0.408	0.07, 0.700	/	—	—	—	—	—	0.9	5.7
Dropping anchor	−0.38, 0.027	−0.50, 0.003	−0.27, 0.111	−0.23, 0.166	−0.03, 0.877	−0.10, 0.554	/	—	—	—	—	1.0	6.5
Positive expressive writing	−0.47, 0.009	−0.61, 0.001	−0.37, 0.037	−0.34, 0.055	−0.14, 0.439	−0.22, 0.223	−0.12, 0.501	/	—	—	—	1.1	7.5
Leaves on a Stream	−0.62, 0.0003	−0.75, <0.0001	−0.52, 0.002	−0.49, 0.003	−0.28, 0.094	−0.37, 0.026	−0.27, 0.103	−0.15, 0.393	/	—	—	1.2	9.0
Progressive muscle relaxation	−0.57, 0.001	−0.74, <0.0001	−0.46, 0.007	−0.44, 0.011	−0.20, 0.241	−0.30, 0.083	−0.19, 0.273	−0.05, 0.764	0.11, 0.522	/	—	1.3	8.0
Guided imagery	−0.64, 0.0002	−0.79, <0.0001	−0.53, 0.002	−0.51, 0.002	−0.28, 0.099	−0.37, 0.024	−0.27, 0.107	−0.14, 0.437	0.02, 0.894	−0.09, 0.587	/	1.3	8.7
Mindful breathing	−0.78, <0.0001	−1.00, <0.0001	−0.67, <0.0001	−0.66, 0.0001	−0.39, 0.025	−0.51, 0.003	−0.38, 0.012	−0.23, 0.192	−0.06, 0.719	−0.19, 0.254	−0.09, 0.586	1.5	9.5

FIGURE 4: Comparing exercises to each other. *Note*: Cohen's *d* and two-sided p-value of Welch *t*-test for comparison between two exercises is given (unadjusted for multiple testing). Data with the time criterion is used. Light green: $p < 0.05$. Full green: $p < 0.0008$ (adjusted significance threshold). The last two columns show the effect size *d* and the mean *M* of each exercise when compared to the measurement-only control.

the potential of smartphone-based interventions [5] as well as self-guided and Internet-based therapies more generally [38–42] in managing mental health issues.

The effect sizes of the different exercises differed substantially even within categories. Cognitive restructuring with the aid of a cognitive distortion list had the smallest effect size ($d = 0.5$), while mindful breathing had the largest effect size ($d = 1.5$). In general, mindfulness exercises (d [95% CI] = 0.9 [0.6; 1.2], 1.0 [0.7; 1.3], 1.2 [0.9; 1.5], and 1.5 [1.2; 1.9]) tended to have larger effects than cog. restructuring exercises (d [95% CI] = 0.5 [0.2; 0.8], 0.7 [0.3; 1.0], and 0.9 [0.6; 1.2]). The mindfulness exercise with the smallest effect ("body scan") included exposure to the bodily sensations of anxiety, which is an important strategy for long-term improvement but might be challenging in the immediate term. It is encouraging that this exercise nevertheless had a large and positive effect size. A mindfulness exercise with a larger effect was Leaves on a Stream. It also involved exposure, but to thoughts instead of bodily symptoms, which might have been less challenging and the reason for the difference in effects. Interestingly, the cognitive restructuring exercise "anxiety-excitement reappraisal" had a larger effect size than other two cognitive restructuring exercises (cognitive distortion list, silver linings) despite being the only one to include attention to the bodily symptoms of anxiety. In contrast to the other two cognitive restructuring exercises, "anxiety-excitement reappraisal" did not involve making an argument, which might be the reason for this difference.

The "anxiety-excitement reappraisal" exercise was a cognitive restructuring exercise in which participants were encouraged to reinterpret their physiological symptoms of anxiety as excitement. Excitement was chosen as the target emotion because it is arousal-congruent to anxiety [43]. In a previous study, this form of cognitive restructuring improved performance and increased excitement but, in contrast to our study, did not reduce anxiety [43]. However, related approaches of reinterpreting stress as positive have been found to reduce anxiety [13]. In Brooks [43], guidance seems to have been limited to instructing participants to say "I am excited" out loud and to try to believe it. In the present study, the concept of anxiety-excitement reappraisal was explained to participants more comprehensively. It is possible that the discrepancy is due to this more detailed guidance. Alternatively, it is possible that the discrepancy is due to other ingredients that appeared briefly in the "anxiety-excitement reappraisal" exercise, namely another form of cognitive restructuring ("anxiety does not harm you") and the encouragement to perform a value-based action.

Diaphragmatic breathing had a much smaller effect size than mindful breathing (d [95% CI] = 0.6 [0.3; 0.9] vs. d [95% CI] = 1.5 [1.2; 1.9]). This seems to suggest that the mindfulness component in mindful breathing made this exercise so effective. This is in line with evidence on the beneficial behavioral and neurophysiological changes induced by mindfulness [15] and with a previous study that found that inducing an acceptance context (as mindfulness does) was more effective than diaphragmatic breathing at reducing state anxiety [44]. However, in contrast to that study, which found no positive effect of diaphragmatic breathing compared to a no-instruction control, our study found a significant positive effect of this exercise. The discrepancy might be due to the specific intervention (CO_2-enriched air inhalation) chosen to induce anxiety in that study. Diaphragmatic breathing (slow, deep, and diaphragmatic) does not allow hyperventilation, but hyperventilation increases the rate of removal of CO_2 from the blood and would be adaptive in this particular situation. In contrast to our finding, Hunt et al. [45] found the opposite pattern: diaphragmatic breathing had a more positive impact than mindful breathing on stress. This effect was moderated by how spiritual participants were, with more spiritual participants benefiting more

from mindful breathing than diaphragmatic breathing. It is possible that our participants were on average more spiritual than those of Hunt et al. [45] (we did not measure this), causing this discrepancy. Alternatively, it might be that the way Hunt et al. [45] introduced the mindful breathing exercise, emphasizing its thousands of years long tradition, increased engagement in their more spiritual participants, or decreased engagement in their less spiritual participants, or both. We did not emphasize tradition in the introduction of this exercise. Interestingly, the three exercises involving exposure to thoughts or bodily sensations of anxiety (d [95% CI] = 0.9 [0.6; 1.2], 0.9 [0.6; 1.2], 1.2 [0.9; 1.5]), had larger positive effects than diaphragmatic breathing (d [95% CI] = 0.6 [0.3; 0.9]), despite the former being more long-term oriented and the latter being a relaxation exercise focused on the immediate term. The size of the effect of diaphragmatic breathing on anxiety in the present study is in line with that of a recent meta-analysis [46].

It seems intuitive that progressive muscle relaxation and the guided imagery exercise were among the most effective exercises, as they are relaxation exercises. Similarly, it was to be expected that immediate positive effects would be achieved by the positive-focused exercises gratitude practice and positive expressive writing. Relaxation exercises and the positive-focused exercises have previously been found to have positive immediate effects (e.g., [18, 47]). It might seem less obvious that cognitive restructuring and even exercises involving exposure would reduce immediate anxiety, as these methods are more focused on long-term outcomes. It is encouraging that they did, seeing as there is strong evidence that they are important methods for long-term improvement in anxiety. Future research could explore the differences between these exercises further.

Our study is not without limitations. The control conditions were controlled mainly for time and measurement (measurement-only control) and attention (reading control). They likely produced lower treatment expectations than the psychotherapeutic exercises. Providing interventions that mimic psychotherapeutic exercises and produce high treatment expectations without including psychologically active ingredients other than treatment expectation is a challenge. This limitation affects the comparison between exercises and control conditions but not between exercises. Another limitation is that the study design was focused on immediate effects, thus limiting our understanding of the duration of the effects. We expect the exercises to have long-term positive effects, especially when performed regularly, because they employ techniques known to have long-term positive effects. However, to what extent this is true for these specific exercises and how exactly the effects evolve and compare over time needs more research. The participants in this study were recruited through the platform Positly (which uses Amazon Mechanical Turk adding additional quality control features on top to reduce inattentive participants and spammers—for details see the section on participants). These participants were likely less intrinsically interested in the exercises and thus might have put less effort into them than app users would, which might mean the effects in the real world are larger than in the study. The generalisability of the findings might be affected by limiting participants to the United States and by possible self-selection of participants being tech-savvy enough to be on Positly and being interested in taking part in the study.

In conclusion, our findings suggest that the 12 app-based exercises we examined had in some cases medium and in most cases large immediate positive effects on anxiety. Effect sizes differed substantially between exercises, including between exercises of the same category. Mindfulness exercises tended to have a larger positive effect than cognitive restructuring exercises. Typical relaxation exercises were not in all cases more effective than exercises that included exposure to anxiety-related thoughts and bodily sensations. Future research could compare the effectiveness of the exercises over time and across different populations.

Data Availability Statement

The data and code that support the findings of this study, as well as a comprehensive preregistration, are openly available on the Open Science Framework at https://osf.io, reference number 2wzyc.

Ethics Statement

Ethical approval was obtained from the Medical Sciences Interdivisional Research Ethics Committee (MS IDREC) of the University of Oxford (reference number R82884/RE001).

Consent

This manuscript contains only anonymized data. Participants gave consent for the publication of these data and also gave informed consent to participate before being enrolled.

Disclosure

A preprint has previously been published [24] as per the link [https://www.medrxiv.org/content/10.1101/2023.11.27.23299083v2]. The funder was involved in designing and planning the study and the technical implementation of the interventions but had no role in data collection, the analysis or interpretation of data, the writing of the report (apart from limited feedback), or the decision to submit the article for publication. Mind Ease had committed in written form to allowing publication irrespective of the results of the study. The study methods and analyses were preregistered. Jan Brauner had the final say over decisions and was not funded by Mind Ease or associated organizations.

Conflicts of Interest

Peter Brietbart was the CEO and Spencer Greenberg the founder of Mind Ease.

Author Contributions

Fabian Kahl, Julia Fabienne Sandkühler, Magda Zena Sadurska, Peter Brietbart, Spencer Greenberg, and Jan Brauner conceived of and designed the study. Fabian Kahl wrote the proposal for the ethics committee and ran the participant testing on Positly. Fabian Kahl and Julia Fabienne Sandkühler preregistered the

study. Julia Fabienne Sandkühler analyzed the data and wrote and revised the manuscript. Jan Brauner was the senior author of the study. All authors read and approved the final manuscript. Fabian Kahl and Julia Fabienne Sandkühler contributed equally to this study.

Funding

Funding for this study was provided by Mind Ease Labs Ltd. (Third Floor South, One Jubilee Street, Brighton, East Sussex, BN1 1GE, UK). Fabian Kahl and Julia Fabienne Sandkühler were funded by Mind Ease for this research project. Magda Zena Sadurska was funded by Spark Wave for this research project, which is a part owner of Mind Ease. Open Access funding enabled and organized by Projekt DEAL.

Acknowledgments

We thank Mind Ease for funding the study and for providing technical support and also thank the participants.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. (*Supporting Information*) Appendix S1: Appendix.

References

[1] World Health Organization, *Depression and Other Common Mental Disorders: Global Health Estimates* (World Health Organization, Geneva, 2017).

[2] S. G. Hofmann, A. T. Sawyer, A. A. Witt, and D. Oh, "The Effect of Mindfulness-Based Therapy on Anxiety and Depression: A Meta-Analytic Review," *Journal of Consulting and Clinical Psychology* 78, no. 2 (2010): 169–183.

[3] S. G. Hofmann, A. Asnaani, I. J. J. Vonk, A. T. Sawyer, and A. Fang, "The Efficacy of Cognitive Behavioral Therapy: A Review of Meta-Analyses," *Cognitive Therapy and Research* 36, no. 5 (2012): 427–440.

[4] J. G. L. A-tjak, M. L. Davis, N. Morina, M. B. Powers, J. A. J. Smits, and P. M. G. Emmelkamp, "A Meta-Analysis of the Efficacy of Acceptance and Commitment Therapy for Clinically Relevant Mental and Physical Health Problems," *Psychotherapy and Psychosomatics* 84, no. 1 (2015): 30–36, 2015, https://karger.com/pps/article-abstract/84/1/30/282854.

[5] J. Firth, J. Torous, J. Nicholas, et al., "The Efficacy of Smartphone-Based Mental Health Interventions for Depressive Symptoms: A Meta-Analysis of Randomized Controlled Trials," *World Psychiatry* 16, no. 3 (2017): 287–298.

[6] J. Firth, J. Torous, J. Nicholas, R. Carney, S. Rosenbaum, and J. Sarris, "Can Smartphone Mental Health Interventions Reduce Symptoms of Anxiety? A Meta-Analysis of Randomized Controlled Trials," *Journal of Affective Disorders* 218 (2017): 15–22.

[7] D. A. Clark, "Cognitive Restructuring," in *The Wiley Handbook of Cognitive Behavioral Therapy*, (John Wiley & Sons, Ltd., Chichester, UK, 2013): 1–22.

[8] A. Maddock and C. Blair, "How Do Mindfulness-Based Programmes Improve Anxiety, Depression and Psychological Distress? A Systematic Review," *Current Psychology* 42, no. 12 (2023): 10200–10222.

[9] Y.-F. Chen, X.-Y. Huang, C.-H. Chien, and J.-F. Cheng, "The Effectiveness of Diaphragmatic Breathing Relaxation Training for Reducing Anxiety," *Perspectives in Psychiatric Care* 53, no. 4 (2017): 329–336.

[10] A. Conrad and W. T. Roth, "Muscle Relaxation Therapy for Anxiety Disorders: It Works but How?" *Journal of Anxiety Disorders* 21, no. 3 (2007): 243–264.

[11] S. Kyeong, J. Kim, D. J. Kim, H. E. Kim, and J.-J. Kim, "Effects of Gratitude Meditation on Neural Network Functional Connectivity and Brain-Heart Coupling," *Scientific Reports* 7, no. 1 (2017): 5058.

[12] S. C. Hayes and S. G. Hofmann, "The Third Wave of Cognitive Behavioral Therapy and the Rise of Process-Based Care," *World Psychiatry* 16, no. 3 (2017): 245-246.

[13] J. P. Jamieson, B. J. Peters, E. J. Greenwood, and A. J. Altose, "Reappraising Stress Arousal Improves Performance and Reduces Evaluation Anxiety in Classroom Exam Situations," *Social Psychological and Personality Science* 7, no. 6 (2016): 579–587.

[14] S. G. Hofmann, S. Heering, A. T. Sawyer, and A. Asnaani, "How to Handle Anxiety: The Effects of Reappraisal, Acceptance, and Suppression Strategies on Anxious Arousal," *Behaviour Research and Therapy* 47, no. 5 (2009): 389–394.

[15] F. Zeidan, K. T. Martucci, R. A. Kraft, J. G. McHaffie, and R. C. Coghill, "Neural Correlates of Mindfulness Meditation-Related Anxiety Relief," *Social Cognitive and Affective Neuroscience* 9, no. 6 (2014): 751–759.

[16] J. Nguyen and E. Brymer, "Nature-Based Guided Imagery as an Intervention for State Anxiety," *Frontiers in Psychology* 9 (2018): 1858.

[17] K. H. Khng, "A Better State-of-Mind: Deep Breathing Reduces State Anxiety and Enhances Test Performance through Regulating Test Cognitions in Children," *Cognition and Emotion* 31, no. 7 (2017): 1502–1510.

[18] L. A. Pawlow and G. E. Jones, "The Impact of Abbreviated Progressive Muscle Relaxation on Salivary Cortisol," *Biological Psychology* 60, no. 1 (2002): 1–16.

[19] L. A. A. der Cleary, "Investigating the Impact of a Brief Cognitive Defusion Intervention on State Anxiety and Psychological Inflexibility/Avoidance," 2017, https://www.proquest.com/openview/a66a448066ee7e6ce250fa5926338a5d/1?pq-origsite=gscholar&cbl=2026366&diss=y.

[20] N. Kazantzis, C. Whittington, and F. Dattilio, "Meta-Analysis of Homework Effects in Cognitive and Behavioral Therapy: A Replication and Extension," *Clinical Psychology: Science and Practice* 17, no. 2 (2010): 144–156.

[21] M. J. Scheel, W. E. Hanson, and T. I. Razzhavaikina, "The Process of Recommending Homework in Psychotherapy: A Review of Therapist Delivery Methods, Client Acceptability, and Factors That Affect Compliance," *Psychotherapy: Theory, Research, Practice, Training* 41, no. 1 (2004): 38–55.

[22] T. Fleming, L. Bavin, M. Lucassen, K. Stasiak, S. Hopkins, and S. Merry, "Beyond the Trial: Systematic Review of Real-World Uptake and Engagement With Digital Self-Help Interventions for Depression, Low Mood, or Anxiety," *Journal of Medical Internet Research* 20, no. 6 (2018): e199.

[23] K. F. Schulz, D. G. Altman, D. Moher, and CONSORT Group, "CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomised Trials," *International Journal of Surgery* 9 (2011): 672–677.

[24] F. Kahl, J. F. Sandkühler, M. Z. Sadurska, et al., "Their Immediate Impact on Anxiety – A Randomised Controlled Trial," *MedRxiv* (2023).

[25] J. Kabat-Zinn, *Wherever You Go, There You Are: Mindfulness Meditation in Everyday Life* (Hachette, UK, 2009).

[26] B. K. Hölzel, S. W. Lazar, T. Gard, Z. Schuman-Olivier, D. R. Vago, and U. Ott, "How Does Mindfulness Meditation Work? Proposing Mechanisms of Action From a Conceptual and Neural Perspective," *Perspectives on Psychological Science* 6, no. 6 (2011): 537–559.

[27] S. C. Hayes, *A Liberated Mind: How to Pivot toward What Matters* (Penguin, 2020).

[28] N. A. Rector, *Anxiety Disorders an Information Guide* (Centre for Addiction and Mental Health, 2005).

[29] T. M. Marteau and H. Bekker, "The Development of a Six-Item Short-Form of the State Scale of the Spielberger State—Trait Anxiety Inventory (STAI)," *British Journal of Clinical Psychology* 31, no. 3 (1992): 301–306.

[30] A. Guillén-Riquelme and G. Buela-Casal, "Psychometric Revision and Differential Item Functioning in the State Trait Anxiety Inventory (STAI)," *Psicothema* 23, no. 3 (2011): 510–515.

[31] R. R. Wilcox, *Introduction to Robust Estimation and Hypothesis Testing* (Academic Press, 2011).

[32] A. Field, *Discovering Statistics Using IBM SPSS Statistics* (SAGE, 2013).

[33] R. G. Knight, H. J. Waal-Manning, and G. F. Spears, "Some Norms and Reliability Data for the State-Trait Anxiety Inventory and the Zung Self-Rating Depression Scale," *British Journal of Clinical Psychology* 22, no. 4 (1983): 245–249.

[34] G. Addolorato, C. Ancona, E. Capristo, et al., "State and Trait Anxiety in Women Affected by Allergic and Vasomotor Rhinitis," *Journal of Psychosomatic Research* 46, no. 3 (1999): 283–289.

[35] K. Kvaal, I. Ulstein, I. H. Nordhus, and K. Engedal, "The Spielberger State-Trait Anxiety Inventory (STAI): The State Scale in Detecting Mental Disorders in Geriatric Patients," *International Journal of Geriatric Psychiatry* 20, no. 7 (2005): 629–634.

[36] C. Bauer-Staeb, D.-Z. Kounali, N. J. Welton, et al., "Effective Dose 50 method as the Minimal Clinically Important Difference: Evidence from Depression Trials," *Journal of Clinical Epidemiology* 137 (2021): 200–208.

[37] B. K. Madsen, D. Zetner, A. M. Møller, and J. Rosenberg, "Melatonin for Preoperative and Postoperative Anxiety in Adults," *The Cochrane Database of Systematic Reviews* 12, no. 12 (2020): CD009861.

[38] R. Fischer, T. Bortolini, J. A. Karl, et al., "Rapid Review and Meta-Meta-Analysis of Self-Guided Interventions to Address Anxiety, Depression, and Stress During COVID-19 Social Distancing," *Frontiers in Psychology* 11 (2020): 563876.

[39] M. Hirai and G. A. Clum, "An Internet-Based Self-Change Program for Traumatic Event Related Fear, Distress, and Maladaptive Coping," *Journal of Traumatic Stress* 18, no. 6 (2005): 631–636.

[40] I. M. Marks, D. Mataix-Cols, M. Kenwright, R. Cameron, S. Hirsch, and L. Gega, "Pragmatic Evaluation of Computer-Aided Self-Help for Anxiety and Depression," *British Journal of Psychiatry* 183, no. 1 (2003): 57–65.

[41] A. Barak, L. Hen, M. Boniel-Nissim, and N. Shapira, "A Comprehensive Review and a Meta-Analysis of the Effectiveness of Internet-Based Psychotherapeutic Interventions," *Journal of Technology in Human Services* 26, no. 2–4 (2008): 109–160.

[42] A. B. Amstadter, J. Broman-Fulks, H. Zinzow, K. J. Ruggiero, and J. Cercone, "Internet-Based Interventions for Traumatic Stress-Related Mental Health Problems: A Review and Suggestion for Future Research," *Clinical Psychology Review* 29, no. 5 (2009): 410–420.

[43] A. W. Brooks, "Get Excited: Reappraising Pre-Performance Anxiety as Excitement," *Journal of Experimental Psychology-General* 143, no. 3 (2014): 1144–1158.

[44] G. H. Eifert and M. Heffner, "The Effects of Acceptance versus Control Contexts on Avoidance of Panic-Related Symptoms," *Journal of Behavior Therapy and Experimental Psychiatry* 34, no. 3-4 (2003): 293–312.

[45] M. Hunt, T. Rajagopal, F. Cerecino, and M. O'Neil, "Mindful versus Diaphragmatic Breathing: Spirituality Moderates the Impact on Heart Rate Variability," *Mindfulness (N Y)* 12, no. 11 (2021): 2743–2753.

[46] R. Shao, I. S. C. Man, and T. M. C. Lee, "The Effect of Slow-Paced Breathing on Cardiovascular and Emotion Functions: A Meta-Analysis and Systematic Review," *Mindfulness* 15, no. 1 (2024): 1–18.

[47] C. M. Burton and L. A. King, "Effects of (very) Brief Writing on Health: The 2-Minute Miracle," *British Journal of Health Psychology* 13, no. 1 (2008): 9–14.