Université Grenoble Alpes & Swansea University

A meta-scientific approach to understanding stress regulation

Dissertation

Submitted in satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY, Ph.D.

in cognitive sciences, psychology, and neurocognition
by

Alessandro Sparacio

March 7, 2023

Supervisors:

Dr. Hans IJzerman
Dr. Gabriela M. Jiga-Boy
Dr. Ivan Ropovik

Dissertation Committee:

Dr. Katy Tapper
Dr. Wolfgang Viechtbauer
Dr. Céline Baeyens
Dr. Pascal Hot
Dr. Christina Izura
Dr. Rebecca Shankland

Acknowledgments

From the bottom of my heart, I would like to thank my family who have supported me during this Ph.D. journey by helping me overcome challenging periods.

I want to thank my supervisors Hans, Gabriela, and Ivan, who have believed in me from the very beginning and who have guided me excellently in this PhD path. Their expertise, knowledge on the topic and timely feedback have helped me grow as a researcher and helped me master my Ph.D. topic.

I would also like to thank my friends; Stefano, Marcello, Pietro, Ludovica, Marco, Martina, Olivier, Kevin, Miguel, Ade, Yuna, and Agatha, who supported me during this journey.

I would also like to thank the MIBODA project which funded my research and made this work possible.

Finally, a heartfelt thanks to my better half, Arianna who has been close to me every day and who has made it easier to face complex periods and difficult challenges. She was a constant source of inspiration, and it is also thanks to her that I was able to complete this thesis and conclude this path.

Abstract

In the present Ph.D. project, I examined the efficacy of four strategies (i.e., self-administered mindfulness, biofeedback, being in nature, and emotional social support) on stress reduction of individuals. To examine the empirical robustness of stress regulation strategies, we conducted two meta-analyses of the literature on these strategies, and one multi-site study experimentally examining the effect of one of the strategies (self-administered mindfulness) on stress regulation.

In chapter 2, I present the result of a Registered Report meta-analysis in which we appraised the available evidence of being in nature and emotional social support on stress levels. After applying our publication bias correction techniques, we found an effect for the former strategy but not for the latter one. In chapter 3, I conducted a pre-registered meta-analysis that investigated whether two internal regulation strategies (i.e., self-administered mindfulness and biofeedback) could be considered reliable strategies for regulating stress levels. We found no evidence for the reduction of stress levels for these two strategies.

Finally, in chapter 4, I present the results of a highly powered multi-site study that investigated the efficacy of four single, brief stand-alone mindfulness exercises (versus three active control conditions) on self-reported experiences of stress in a population unfamiliar with mindfulness meditation. Results of the multi-site study provided evidence that participants had lower levels of self-reported levels of stress when they listened to one of the mindfulness exercises compared to those that were in the control conditions.

In both meta-analyses, the high heterogeneity, the low-quality of the included studies, and the presence of publication bias cast doubts on the strength of evidence behind the studied effects. For what concerns the multi-site project, even if we cannot exclude the presence of a demand effect, our results hold promise for the efficacy of self-administered mindfulness interventions for the downregulation of stress.

Taken together these studies suggest that stress levels can be reduced by relying on strategies that are non-pharmacological, non-invasive, and low-cost. Future studies (preferably in the form of Registered Reports) should try to make up for these shortcomings of the literature with well-powered studies and with the use of physiological measurements that could reduce the problem of the demand effect for studies on mindfulness.

Résumé

Dans le présent projet de doctorat, j'ai examiné l'efficacité de quatre stratégies (c'est-à-dire la pleine conscience auto-administrée, le biofeedback, le fait d'être dans la nature et le soutien social émotionnel) sur la réduction du stress des individus. Pour examiner la robustesse empirique des stratégies de régulation du stress, nous avons effectué deux méta-analyses de la littérature sur ces stratégies, et une étude multi-sites examinant expérimentalement l'effet d'une des stratégies (la pleine conscience auto-administrée) sur la régulation du stress.

Dans le chapitre 2, je présente le résultat d'une méta-analyse pré-enregistrée dans laquelle nous avons évalué les preuves disponibles sur le fait d'être dans la nature et sur le soutien social émotionnel vis à vis du stress. Après avoir appliqué nos techniques de correction du biais de publication, nous avons trouvé un effet pour la première stratégie mais pas pour la seconde. Dans le chapitre 3, j'ai effectué une méta-analyse pré-enregistrée qui a cherché à savoir si deux stratégies de régulation interne (c'est-à-dire la pleine conscience auto-administrée et le biofeedback) pouvaient être considérées comme des stratégies fiables pour réguler les niveaux de stress. Nous n'avons trouvé aucune preuve de la réduction des niveaux de stress pour ces deux stratégies.

Enfin, dans le chapitre 4, je présente les résultats d'une étude multisite (possédant une forte puissance statistique) qui a examiné l'efficacité de quatre brefs exercices de pleine conscience autonomes (par rapport à trois conditions de contrôle actif) sur les expériences de stress autodéclarées dans une population peu familière avec la méditation de pleine conscience. Les résultats de l'étude multisite ont démontré que les participants avaient des niveaux de stress autodéclarés plus faibles lorsqu'ils écoutaient l'un des exercices de pleine conscience par rapport à ceux qui étaient dans les conditions de contrôle.

Dans les deux méta-analyses, l'hétérogénéité élevée, la faible qualité des études incluses et la présence d'un biais de publication jettent des doutes sur la force des preuves à l'appui des effets étudiés. En ce qui concerne le projet multi-sites, même si nous ne pouvons pas exclure la présence d'un effet de demande, nos résultats sont prometteurs quant à l'efficacité des interventions de pleine conscience auto-administrées pour la régulation du stress.

Dans l'ensemble, ces études suggèrent que les niveaux de stress peuvent être réduits en s'appuyant sur des stratégies non pharmacologiques, non invasives et peu coûteuses. Les études futures (de préférence sous la forme de rapports enregistrés) devraient essayer de compenser ces lacunes de la littérature par de futures études et par l'utilisation de mesures physiologiques qui pourraient réduire le problème de l'effet de demande pour les études sur la pleine conscience.

Table of Contents

Chapter 1: Introduction 9

Chapter 2: Stress regulation via being in nature and social support in adults, a Registered Report meta-analysis 31

Method 38

Analyses 42

Results 49

Discussion 62

Conclusions 68

Preface of Study II 70

Chapters 3: Stress regulation via self-administered mindfulness and biofeedback interventions in adults: A pre-registered meta-analysis 71

Method 79

Analyses 83

Results 89

Discussion 100

Conclusions 108

Preface of Study III 111

Chapter 4: A large, multi-site test of self-administered mindfulness among English speakers 112

Study 1: Survey on self-administered mindfulness 117

- **Procedure** 117
- **Results** 118

Study 2: The multi-site project 120

- **Hypothesis** 120

Methods 120

- **Procedure** 124

Analyses 128

Results 132

Discussion 141

Conclusions 144

Chapter 5: Discussion 147

- **External regulations strategies (included in the Registered Report meta-analysis)** 148
- **Internal regulations strategies (included in the pre-registered meta-analysis)** 149
- **Multi-site study on self-administered mindfulness** 151
- **Four strategies, but a plethora of protocols** 151

Towards the building of a stress management intervention 152
Limits of stress regulation strategies 154
Conclusions 155
References 156
Appendices 183

Chapter 1: Introduction

Facing daily challenges is a normal part of life. However, when we are overwhelmed by them, we may become stressed. Let's take the case of a student who fails an exam because what she could do (e.g., time spent preparing for the exam, knowledge of the subject) proved to be insufficient to pass her test. If the student makes another attempt and fails the exam again, twice or three times, her stress response may well become chronic. Perhaps the student even falls into chronic anxiety or depression.

Stress is rising and seems to become an ever-greater problem in richer countries. As one example from the United States, the American Psychological Association (APA) in August 2017 conducted its annual "Stress in America" survey to investigate stress levels of people in the US. Results of this survey pointed to an overall increase in stress levels: People in the US were more likely to report symptoms of stress such as tension, anxiety, and nervousness than in the previous year.

Moreover, with the outbreak of the COVID-19 pandemic, the situation seemed to have worsened considerably. An APA survey conducted between August 4th and August 26th, 2020, revealed that people from the US were deeply affected by the COVID-19 outbreak. Of their total sample of 3,409 adults and 1,026 teens, 67% indicated that their stress level increased throughout the pandemic and 19% reported that their mental health worsened during 2020. The pernicious effects of the pandemic not only affected people from the US, but populations across the globe, causing a substantial increase in the burden of anxiety and depressive disorders in the richest countries, in particular of the western Europe and of north America (COVID-19 Mental Disorder Collaborators, 2021). As stress increased considerably from 2017 to the present day, it is vital to find stress regulation strategies that people can implement, to shield themselves from the development of mental health problems such as anxiety or depression.

There is a plethora of strategies to deal with stress, however not all of them are widely accessible. Some strategies may be too expensive or unaffordable for most people. In the present Ph.D. project, I focused on interventions that do not require medication use, are low-cost, and are within reach for most people, including 1) a type of meditation (i.e., self-administered mindfulness), 2) self-regulation based on feedback on biological mechanisms (i.e., biofeedback), 3) walking in natural environments (i.e., being in nature), and 4) seeking the support of loved ones (i.e., emotional social support). I examined which of these stress regulation strategies have a robust effect on stress regulation with a combined approach: The project includes two meta-analyses on the above-mentioned stress regulation strategies with a workflow relying on multiple bias-correction techniques (including multilevel regression-based models and permutation-based selection models) to provide the most up-to-date publication bias-corrected estimates, and one multi-site study that examined the effect of one of the strategies (i.e., self-administered mindfulness) on stress regulation.

Before describing these three projects in detail, in this introduction, I will focus on two theoretical models of modern psychology: The allostasis model (Sterling & Eyer, 1988) and the process model of emotions (Gross, 1998a, 1998b), to explain how stress regulation mechanisms work. The allostasis model describes the reactions of the body when we are under stress, and it is especially helpful for understanding how a persistent stress response can cause the individual to develop affective consequences of stress, such as chronic depression and anxiety.

The process model of emotions, on the other hand, provides a list of emotion regulation responses that are exploited by the stress regulation strategies we included in the present project. I will then provide an historical background to this thesis, with examples which give the reader a glimpse of how the concept of stress and stress regulation have evolved in the years before arriving at the aforementioned models. Finally, in the last section of this introduction, I will provide an overview of the methodological principles of this thesis. I will then describe in depth the three empirical studies that I included in the present Ph.D. project followed by the final discussion.

The process of the unfolding of stress - and how it remains

The allostasis model

The allostasis model developed by Sterling & Eyer, 1988 is central for the understanding of the stress response and of the reaction of the organism when an individual is exposed to stress for a prolonged period of time. The word allostasis derives from the Greek prefix ἄλλος (állos), translated into "other," "different" + the suffix στάσις (stasis), translated into "standing still"; as such, “allostasis” translates literally into "remaining stable by being variable" (Sterling & Eyer, 1988). The concept of “*allostasis*” foresees a set of mechanisms whose aim is to bring back the organism to a situation of equilibrium after that a stressor is encountered.

The concept of allostasis takes into account that there is an array of steady-state conditions and different stress situations activate different pathways of the stress response involving the hypothalamic-pituitary-adrenal axis (HPA) and the sympathetic nervous system (Goldstein, & McEwen, 2002). The condition of balance is achieved via a system of mechanisms that considers the different environmental conditions in which the individual lives (Sterling & Peter, 2020). Thus, for example, a student who has to take an exam will have high levels of blood pressure and a high concentration of cortisol in the blood (Ramsay, & Woods, 2014). Blood pressure and cortisol levels will remain high until the student takes the exam, returning to normal levels thereafter. Thus, in the current example, the level of a physiological parameter (e.g., blood pressure) varied in function of the challenge posed by the environmental threat.

Another notion that is crucial in the model of Sterling & Eyer (1988) is the “allostatic load” that occurs when multiple effectors are activated intermittently or continuously to restore a condition of balance disrupted by some environmental conditions (Goldstein, & McEwen, 2002). Allostatic load is characterized by the wear and tear of the hippocampus and of the immune and muscular systems, caused by a persistent presence of cortisol in the body (McEwen 1988). A long-lasting allostatic load is the culprit of some of the affective consequences of stress such as depression and chronic anxiety (Goldstein, & McEwen, 2002). Indeed, the atrophy of the hippocampus and of the amygdala caused by the allostatic load (McEwen 1988) has been associated with post-traumatic and depressive disorders because it affects how the brain processes emotional information and impairs cognitive function (McEwen, & Magarinos, 1997; McEwen, 1998; McEwen, 2000; Sapolsky, 1996).

The process model of emotions

If the allostasis model can help us understand how our body responds to stress from a mainly physiological point of view, Gross's process model of emotions is useful for understanding through which mechanisms individuals can regulate stress levels (Gross, 1998a, 1998b; Gross & Thompson, 2006). The process model of emotion regulation claims that specific strategies can be differentiated along the timeline of the unfolding emotional response (Gross, 1998a, 1998b; Gross & Thompson, 2007). We can distinguish between antecedent- and response-focused emotion regulation strategies at the most general level. The former occurs prior to emotion generation, while the latter refer to things we do once an emotion is already under way after the response tendencies have been generated (Gross, 1998a). Gross describes five sets of emotion regulation processes: Situation selection, situation modification, attention deployment, cognitive change, and response modulation (Gross, 1998, b).

Situation selection is the first of the antecedent-focus strategies that allow us to filter the situations that could be a source of negative emotions and stress. For example, an individual who is particularly sensitive to the fear of becoming infected with COVID-19 may prefer to go out for a drink in a bar that is not crowded. The second strategy, *situation modification*, leads to changing a situation to reduce the emotional impact that derives from it. Going back to the example above, an individual taking a drink in a bar may decide to leave when it becomes crowded. *Attentional deployment* involves diverting attention to another aspect of the situation to influence one's emotions. For example, an individual can be distracted from the fear of getting infected by COVID-19 by engaging in a conversation with a friend. *Cognitive change* allows an individual to change how a situation is assessed to alter its possible emotional impact, for example by thinking that "I am young and have been vaccinated, so I have a low risk of suffering severe consequences after a COVID-19 infection". Finally, *response modulation* occurs at the end of the emotion generation process and is based on directly influencing the physiological and behavioral components of the emotional response. For example, an individual might hide an expression of concern in a place that is too crowded to avoid being seen as a hypochondriac in the eyes of their friends.

The allostasis model and the process model of emotions are born after the term stress and the concept of stress regulation have undergone significant evolutions throughout history. In the next sections we will go over the historical development of the term "stress" and "stress regulation". The reader will gain an understanding of the conceptualizations of stress that gave rise to the allostasis model and to the process model of emotions as well as the historical development of the term stress from this summary. It is not meant to be a thorough overview of the history of stress or the history of stress regulation; rather, it is meant to serve as background information for the current Ph.D. dissertation.

History of the understanding of stress

Nowadays, the concept of stress is commonly associated with small unforeseen events in everyday life: Getting stuck in traffic, having a fight with a friend, and taking an exam are all potential sources of stress. Sometimes we even worry about what might happen in the future, even if the chances that something bad might happen are meager. As Epictetus said, “Man is not worried by real problems so much as by his imagined anxieties about real problems” (Epictetus et al., 1990). Mark Twain expressed a very similar concept many years later: "I've lived through some terrible things in my life, some of which actually happened" to express that only a small proportion of that we worry about are real problems to which we must find a solution.

People of the most developed countries live in relatively quiet times and are not generally faced with threats to their lives (although the COVID-19 pandemic and the Russian invasion of Ukraine have brought existential threats closer to our lives). Indeed, throughout history, people have faced much more severe threats that have significantly reduced their quality and life expectancy. For example, in the book *The Representation of External Threats, From the Middle Ages to the Modern World*, Crailsheim and Elizalde (2019) described some of the threats that people have faced throughout history. To cite an example the invasions of the barbarians who, for about a thousand years, besieged the capital of the Byzantine Empire (which fell in 1453 A.D.) putting the life of the population of the Eastern Roman Empire under constant stress. In different historical periods there have been heterogeneous sources of threats of various kinds (e.g., the bubonic plague of the 14^{th} century, political and economic instabilities) that have significantly impacted people's lives.

However, in this thesis we narrow the focus on military threats to describe how the concept of stress evolved. The choice is only seemingly arbitrary, because wars have been one of the most persistent events that impacted people’s lives throughout history, and because the outbreak of military conflicts have given a considerable boost to the understanding of the concept of stress (See Fueshko, 2016; Grinket & Spiegel, 1945; Lazarus & Folkman, 1984; Ustinova et al., 2014). This boost was given by situations of extreme stress that wars caused to the lives of people, a condition that later would have been called Post-Traumatic Stress Disorder (PTSD), a syndrome characterized by far more extreme consequences than other forms of stress.

Stress in the Classical Era and during the Middle Ages

The first historical account of the detrimental psychological effects of the horrors of battle on the lives of soldiers can be found in Gorgias' Encomium of Helen (485 BC - 375 BC). Here, the Greek philosopher noted some signs that resemble those of Post-Traumatic Stress Disorder: Agitated and anxious states that lasted in warriors' minds even after the end of the war (Gorgias et al., 2005; Ustinova et al., 2014). Gorgias claimed that the war's experience had destroyed a sense of equilibrium and left the survivors of the battlefield in a continual state of upheaval. The Greek tradition placed a lot of emphasis on the idea of equilibrium. Indeed, harmony and balance were fundamental aspects of existence, according to Hellenic thinkers (Lippman, 1963). To emphasize how changeability and adaptation are two principles that are inherent to all things, Heraclitus (540–480 BC) said that "everything flows like a river" (*Panta rei os potamòs*; Heraclitus & Marcovich, 1978).

In a similar vein, harmony was a necessary condition for the survival of organisms for Empedocle (Le Moal, 2007). This concept of equilibrium is taken up in modern times in studies on stress (and in the allostasis model as well): When one’s balance is disturbed, then the stress response occurs. Although some fundamental concepts relating to stress were indeed already introduced in ancient times, the term “stress” in fact had a completely different origin.

Indeed, the concept of stress was coined by the ancient Romans to refer to the physical pressure that a body exercises to another and, originally, the term was far from having a psychological connotation. Using this system of physical pressure, the ancient Romans were able to build arched bridges by exploiting compressive stresses (Baratta & Colletta, 1998). In fact, it is no coincidence that the term stress is a word derived from the Latin verb *strictus* which means tight, closed, or compressed (Strictus, n.d.; Robinson, 2018). The term strictus then entered the English dictionary with the meaning “distress” to indicate a form of physical injury and gave rise to the term "estresse" to mean narrowness and constriction in the French dictionary (Stress, n.d.; Robinson, 2018).

Although the term stress was coined in Latin times and became part of some languages derived from Latin, it was not frequently used during the Middle Ages (MacLehose, 2020). However, some 15th and 16th centuries authors described many symptoms that could be associated with a stressful condition. Shakespeare, for instance, recounted how King Henry IV struggled to sleep because of the weight of his duties as the ruler of the Kingdom of England, as opposed to the servants' easy sleep who felt no pressure (Shakespeare et al., 2007; Henry IV Part 2, act 3, scene 1). Henry IV's difficulty to sleep was undoubtedly a result of the pressure he felt from trying to run the country more effectively. Additionally, a number of phenomena that were originally documented in studies on medieval medicine are today recognized as signs of stress, including agony, worry, unease, and anxiety (MacLehose, 2020).

Stress in the modern era

When Hooke first introduced the term stress in his research on the elasticity of springs in the late 17th century, the term was first employed in the field of physics (Hinkle, 1974). This definition of stress referred to the force that another agent applied to a solid body that has the potential to cause the body to deform. Claude Bernard, a French physician who was regarded as the father of experimental medicine, made a significant contribution to research on stress later in the 19th century. Indeed, he described the idea of the "milieu intérieur" (also known as the environment within), which refers to an organism's ongoing effort to reestablish a condition of internal balance (Gross, 1998). According to this mechanism for example the body reduces the size of the arteries in cold weather so as not to disperse heat promoting the vasodilation of blood vessels when it is warm for the opposite reason.

Bernard's work led the term stress to be used extensively in the medical field to refer to strains and tensions of external agents that could negatively impact people's health (Holmes, 1986). For example, in the Statistical Account of Ireland, Williams Shaw Mason (1814 –1819) described how poor women could die prematurely because of "stresses", namely violent heats from hard work (Robinson 2018). Here, the term stress was used to identify a set of internal and external responses which greatly affected young women's life expectancy. Another example could be found in the work of Sir Williams Osler (1910), who defined the life of the Jewish businessman as characterized by "stress and strain," and this condition was directly linked to recurrent episodes of chest pain due to coronary heart disease, also called "*angina pectoris*" (Hinkle, 1974). Finally, two-and-a-half decades after Osler's description, Walter Bradford Cannon (1935) used the term stress to refer to a disturbance of balance occurring in a potentially threatening situation. This state of things elicits a reaction that Cannon called "*fight-or-flight*" (Cannon, 1915). The term stress would then be used extensively for years to come, and still to the present days, to describe the stress response itself.

Stress in contemporary society

The wars of the 20th century, as in centuries past, brought a renewed focus on stress. For instance, during the first World War, some veterans reported severe psychopathologies, but the cause was not immediately attributable to having participated in a war (Jones, Fear, &

Wessely, 2007). The problem was initially underestimated: In 1917, the French doctor Joseph Babinski attributed these symptoms to hysteria, while some thought that these behaviors were fiction designed ad hoc to not return to the battlefields (Philippon & Poirier, 2009). Only in 1916 the psychologist Charles Myers, in the medical journal *The Lancet,* stated that the brain injuries reported by the survivors were determined by the proximity to the bombings and caused by the excessive noise of the explosions. This syndrome was later called "shell shock" syndrome (Myers, 2016).

The Second World War gave an even greater impetus to the study of stress. The term stress gained so much importance that Grinker and Spiegel's (1945) book on war was titled *Men under stress*. In this book, Grinker and Spiegel (1945) described a variety of symptoms experienced by war veterans: Paranoia, dizziness, constant tension, and inability to relax. These symptoms were also often reported by soldiers who had not sustained any injuries on the battlefield, leading some doctors to recognize the importance of psychological factors in stress studies.

During the second half of the 20th century many authors tried to define the concept of stress, thanks also to research stimulated by the war. However, as compared to the previous definitions in which the term revolved around a concept of threatened equilibrium mostly from a physiological point of view, for the first-time psychological factors entered consistently in the definition of the term. Thus, for Lazarus and Folkman (1984), stress is a relationship between a person and the environment that is perceived as harmful when the individual is overwhelmed by environmental demands. For Shalev, Yehuda, and McFarlane (2000), stress is a psychophysiological response to events that evoke a sense of threat and unbalance in individuals. For McEwen (2007) instead the term stress means any physical, psychological, and emotional factor that is responsible for a physical and psychological tension. Finally, according to the American Psychiatric Association (2014), stress is a sense of being overwhelmed, of worry and pressure that can affect people of all ages, sexes, and races, impacting both their physical and mental health.

The term stress has therefore had a discontinuous evolution over the course of history. The concept of balance was already identified by the ancient Greeks in the works of Empedocles and Heraclitus. However, the term stress at the beginning was far from having a psychological connotation: Indeed, the term was associated with the concept of physical pressure exerted by one body on another. In the Middle Ages, the word stress was never used, although there were descriptions of individuals who suffered from symptoms attributable to the stress response.

Finally, after the two World Wars, the term started to be consistently associated with psychological factors, and in the second half of the 20th century we saw a flourishing of multiple definitions of stress. If stress has been so present throughout history, what strategies have people used from ancient times to reduce the negative impact of stress?

History of understanding of stress regulation

Early stress regulation strategies in the ancient era and Middle Ages

Throughout history, people have tried to regulate stress in many ways that are reminiscent of how we regulate stress in modern times and via strategies that for some aspects are similar to the ones included in the present Ph.D. project. For example, in ancient times, in his famous poem *Bucolicas* Virgile (70 BC - 18 BC) proposed a remedy to escape from the horrors of the war and from the worries of the everyday life: Being in touch with nature to relieve stress and to find *ataraxia*, a Greek term used to designate the perfect peace of the soul that comes from the liberation of passions (Virgile et al., 1990). Finding perfect peace, according to the ancient poet, was achievable by following the lifestyle of the shepherds who are the protagonists of Virgilian bucolicas. Indeed, the shepherds described by Virgile did not perform wearing or degrading jobs but composed sylvan songs in a world that was far from the excessive complexity of life in the city (Virgile et al., 1990). The restoring effect of being in nature was thus already acknowledged in ancient times. This conception of Virgil recalls for some aspects the modern theory of Ulrich (1983) based on which nature provides a restorative influence helping individuals recover from stress. That is because when a threat is faced, an unthreatening green environment might evoke feelings of stillness, reduce rumination, and increase physiological restoration (see also Ulrich et al., 1979).

Seneca (4 BC - 65 AD), instead, suggested avoiding stress altogether by not worrying about what has not happened yet and by focusing on the present moment. Seneca was part of the philosophical current of Stoicism according to which individuals must be pragmatic and should try to change only the situations that they can control and not get overwhelmed by excessive anxiety about situations that were beyond their control (Vogt, 2020). A concrete way for Seneca to minimize stress responsiveness was to identify with the worst-case scenario. In fact, the Roman philosopher was terrified of losing privileges, not to be in good health, and of becoming poor. To eliminate the anguish that this thought caused him, he decided to experience poverty from time to time. Indeed, according to Seneca, experiencing the worst-case scenario would have led to being prepared for every situation and this would have automatically reduced stress (Irvine, 2009). This strategy suggested by Seneca is somewhat reminiscent of a form of emotional regulation response that Gross introduced sometime later; the reappraisal (Gross, 1998a). Visualizing the worst possible outcome indeed allowed a reinterpretation of the current situation (that it will never be as bad as the worst-case scenario) dampening its emotional impact. Reappraisal in a way, is also central in some mindfulness exercises, in which individuals are invited to try alternative appraisals of life events (Garland et al., 2009).

Sometime later, the Roman emperor Marcus Aurelius (121 AD - 180 AD) described a technique called ‘journaling’, which consisted of recording daily events and noting down what happened to the mind and body when one experienced specific experiences. Throughout the practice of journaling, individuals could change behaviors to improve their physical and mental states (Aurelius & Gardini, 2005). The practice of journaling survived also to the present days, for instance noting down stressful situations has been linked to improvements in symptoms of depression (Lepore, 1997) and in reducing stress (Donnelly, & Murray, 1991). Finally, Cicero (106 AC - 43 AC) described another way to cope with stress in his book "De amicitia"; for Cicero, the sincere friend was the one who offered his help to overcome obstacles that were a source of stress (Cicero et al., 1949). Here, Cicero clearly refers to social support, one of the stress regulation strategies that nowadays has probably garnered the most empirical support (e.g., Cohen, 2004; Lakey & Cronin, 2008).

In the Middle Ages, a variety of strategies were also used to reduce stress levels. For instance, people utilized pharmacological remedies, including a selection of opiates, to calm and enable them to fight stress. Men of the Middle Ages also ingested/drank infusions of chamomile, or henbane, a poisonous plant that was used as a narcotic to relax and calm down. In addition to using these substances, those suffering from stress and insomnia were advised to listen to soothing music that came from musical instruments or natural elements (i.e., the sound of trees swaying in a light breeze; MacLehose, 2020). Music has also been used in the present days to successfully stifle the build-up of stress as it has been highlighted by a recent meta-analysis (Pelletier, 2004).

The concept of homeostasis

From the twentieth century onward, research focused not only on stress, but also on its regulation. In his book *The Wisdom of the Body* (1932), Cannon described the functioning of homeostasis: A steady state that is maintained because any tendency to change would be thwarted by some factors that resist change (Cannon, 1932). The concept of homeostasis to a certain extent takes up the concept of equilibrium already described in ancient times by

Empedocles and Heraclitus and expanded the idea of constancy of Bernard's internal environment. However, the concept of homeostasis specifically refers to a set of physiological mechanisms that have the goals of bringing back the body to a state of normal functioning (Cannon, 1929). Thus, the stress response occurs when the organism senses a disruption of homeostasis (Goldstein, & McEwen, 2002).

For example, an increase in blood sugar makes the individual thirsty, prompting him to drink water to dilute the concentration of sugar in the extracellular fluid. In addition to this, other bodily feedback systems ensure that blood sugar levels are not too high; for instance, when this happens, the action of insulin comes into play, which lowers blood glucose levels. In this case, the source of stress (i.e., elevated concentration of sugar in the extracellular fluid) has been extinguished by a set of homeostatic mechanisms. Homeostasis is therefore the process that allows feedback-based controls to maintain constant parameters; fluctuations in blood glucose are constantly monitored, and their concentration is regulated by antagonistic hormones in a negative feedback loop.

Further evolution of the studies on stress regulation came due to the work of Hans Selye, an Austrian doctor who started to study stress on rats. In 1936, Selye was able to identify a set of universal non-specific physiological reactions that a living being employed because of external demands coming from the environment. An exhaustive scheme of all responses enacted by "stressed" individuals was then developed in one of the first formal theories on stress: The General Adaptation Syndrome (GAS; Selye, 1946). According to this theory, three stages are encountered by the individual: 1) Alarm, whereby the individual perceives the environmental threat, and the central nervous system is activated to counter the menace and manifests as increased heart rate, blood pressure, and enhanced arousal; 2) resistance, whereby the organism tries to counter the negative effect of a prolonged stressor by keeping an immunological and hormonal answer to the perceived threat; and 3) exhaustion or recovery, whereby some permanent damages can arise (e.g., mental illnesses) if the stressors are maintained and the organism is unable to counter its adverse effects. Instead, if the individual is able to counter the harmful effects of the stressor, the organism returns to its normal functioning and the crisis is solved.

The works of Cannon (1932) and Selye (1936) inspired the conceptualization of the allostasis model. Indeed, the notion of "homeostasis" was taken up in the concept of allostasis. The difference however lies in the fact that while the concept of homeostasis responds to fixed parameters, the concept of allostasis provides for an adaptation of these parameters which vary with respect to environmental factors (hence the meaning of allostasis; remaining stable by being variable; Sterling & Eyer, 1988). The concept of allostatic load instead draws inspiration from the exhaustion phase of the General Adaptation Syndrome of Selye (1946) that foresees the insurgence of damages to the body if the organism is unable to counter the stressful situation.

Stress regulation redefined according to the appraisal theory

Cannon and Selye's studies were focused mostly on bodily physiological responses that are put into action when a noxious threat is perceived. However, what was missing until that time was the conceptualization of a set of responses that individuals could implement to deal with perceived threats. This set of responses, called "coping," was first described in 1966 by Richard Lazarus in his book "*Psychological Stress and the Coping Processes",* whose work shed light on stress regulation for the first time.

Lazarus proposed that people's reactions to a potentially stressful event not only depend on the event itself, but also on how people react to that event. In collaboration with Folkman, he elaborated the *"Transactional Model of Stress and Coping*" in 1984 when, for the first time, he provided a definition of stress centered on a psychological aspect. Indeed, Lazarus defined stress as "a particular relationship between the person and the environment that is appraised by the person as taxing or exceeding his or her resources and endangering his or her well-being" (Lazarus & Folkman, 1984, p. 19).

According to Lazarus and Folkman's model, stress and emotions depend on how an individual evaluates (appraises) their transactions with the environment. If people encounter something that overcomes their personal resources, they tend to feel high-level of stress (Lazarus & Folkman 1984). Key in Lazarus and Folkman's theory is the concept of "appraisal", which denotes how individuals determine why a particular transaction with the environment is stressful (Lazarus & Folkman, 1984). The theory of Lazarus and Folkman distinguishes two forms of appraisal: Primary and secondary appraisals. The former refers to assessments on the situation encountered (e.g, Is the exam I will have to take feasible?), while the second refers to assessments of the personal resources that can be deployed to favorably resolve the potential threatening situation (e.g., Have I studied enough to pass the exam?). Combining these two types of appraisals determines the intensity of the emotional response and the extent to which the situation stresses the individual (Chang, 1998; Dewe, 1991).

Stress and emotions

Due to Lazarus's studies on appraisal, the concept of emotional response following a stressful event began to emerge. Indeed, according to Lazarus (1991), "psychological stress theory is tantamount to a theory of emotion, and because the two literatures share overlapping ideas, the two fields might usefully be conjoined as the field of emotion theory" (p.10).

If we can consider the experience of stress as falling into the literature on emotion, then we can pick from this larger field some mechanisms that might be useful in stress regulation processes. In the next section, we will use the allostasis model to explain what the stress response consists of, and what are the affective consequences that a persistent stress response can generate. Finally, we will apply the process model of emotions to explain the mechanisms through which people can regulate stress with the four strategies that we selected for the current Ph.D. project.

The stress response and its consequences explained through the allostasis model

The allostasis model explains how our body responds to a stressful situation through the simultaneous activation of the autonomic nervous system (ANS) and of the HPA axis (Stephens & Wand, 2012). When a stressor is encountered, the information is sent to the amygdala, a part of the brain that is involved in processing emotions. If the environmental demand is perceived as dangerous, a signal is sent to the hypothalamus that is connected to the rest of the body via the ANS. The autonomic nervous system is divided into two subsystems: The sympathetic system, which activates the "fight-or-flight" response, and the parasympathetic system, which exerts an opposite response, relaxation. Upon activation of the sympathetic system, there is an increase in heart-rate, respiratory rate, and muscle tone (stimulated by the production of epinephrine and adrenaline). This set of responses aims to allow the individual to react as quickly as possible to the threatening stimuli.

Simultaneously with the activation of the sympathetic nervous system, the hypothalamus also triggers the HPA axis, a system consisting of different organs: The hypothalamus, the pituitary glands, and the adrenal glands. The aim of the HPA axis is to maintain the activity of the sympathetic nervous system until the perceived threat is countered. A cascade of reactions is triggered and the first hormone to be released by the hypothalamus is the corticotropin-releasing hormone (CRH) that arrives in the pituitary gland. In the pituitary gland, the CRH stimulates the production of another hormone: The adrenocorticotropic hormone (ACTH). In turn, the ACTH enters in the adrenal glands and leads to the production of several hormones up to the production of the best-known hormone regarding the stress response: Cortisol, whose function is to distribute energy in the form of glucose to the parts of the body that are the most active (head, heart, and muscles associated with the fight or flight response; Hellhammer, Wüst & Kudielka, 2009). The allostasis mechanism works via a system of negative feedback: Increases in the monitored variable's levels cause changes in effector activity that resist and, in turn, "buffer" those changes. Thus, for example, when the perception of danger decreases, cortisol level drops, and the parasympathetic nervous system helps to shut down the stress response and restore allostasis (McEwen & Stellar, 1993).

Alongside these physiological changes (i.e., physiological component of stress), the stress response is characterized by a series of well-defined affective reactions: Feelings of nervousness, strain, and tensions (i.e., affective component of stress; Ratanasiripong, Ratanasiripong, & Kathalae, 2012). Additionally, individuals that are overrun by external demands may start ruminating and perseverating on negative thoughts (i.e., cognitive component of stress; Alloy et al., 2000). People more prone to rumination are also more exposed to episodes of major depression than people who do not engage in rumination following a stressful life event (Alloy et al., 2000).

The three components of the stress response (i.e., affective, physiological and cognitive) should not be intended as separated, but either interdependent or even completely interwoven (see e.g., Pessoa, 2008). Indeed, emotions and cognitions, according to some literature, have interconnections within the brain and this might explain their being interwoven (Erickson & Schulkin, 2003; Halgren, 1992; Lane & Nadel, 2000)" (p. 2; Storbeck & Clore, 2008). Furthermore, a highly powered multi-site study showed that smiling could both increase ongoing feelings of happiness and make it start feelings of happiness in neutral scenarios (Coles et al., 2022). Results of this study suggest that peripheral nervous system activity and emotional experience as independent components of an emotion response (Coles et al., 2022).

The lack of a proper separation between these components also means that stress does not stand on its own; when stress exceeds what one can handle, longer-term affective consequences may emerge. Indeed, when the stress response is prolonged in the time the organism enters in the allostatic load characterized by the wear and tear of the body (McEwen 1988). In particular, the allostatic load leads to damage of the hippocampus, associated with depressive and anxiety disorders (McEwen 1988; McEwen, & Magarinos, 1997; Sapolsky, 1996).

Because of the negative consequences of stress, individuals need to find a way to shield themselves and to prevent the consequences of a long-lasting allostatic load. In the next section, we will outline the different strategies that individuals can use to regulate stress by providing some examples, and we will detail why our choice fell on the strategies that we included in this Ph.D. project. Finally, we will explain why the process model of emotions of Gross can be relevant to explain why these strategies might work in decreasing stress levels of individuals.

Towards the choice of our regulation strategies

Individuals can counter the maladaptive consequences of stress in various ways, with a plethora of strategies. Some of them can give temporary relief but, in the long run, can cause pernicious damage to the body (i.e. maladaptive strategies; Corbin et al., 2013). Others minimize the psychological, physical, or emotional harm of a stressor without drawbacks (i.e., adaptive strategies; Lazarus & Folkman, 1984).

In the present Ph.D. project, we focused only on strategies that can grant positive effects without harming the individual. Besides this primary distinction, we can further divide these two strategies into internal regulation strategies and external regulation strategies. The difference between these two families is intuitive: The former is based on the self-regulation abilities of individuals and only requires some devices (e.g., smartphones). On the other hand, the latter requires the intervention of an external source to reduce stress levels.

Examples of internal regulation strategies can be: Exercising, listening to music, practicing yoga, meditating, and biofeedback. On the other hand, external regulation strategies can be: Meeting some friends, going on vacation, or walking in a park. We choose arbitrarily four strategies from this large pool: Two of them are based on internal mechanisms (self-administered mindfulness and biofeedback), and two of them can be categorized as external regulations strategies (being in nature and emotional social support).

The reasons we decided to focus the Ph.D. project on these strategies are compelling: To begin with, they are non-invasive, relatively effortless, and virtually accessible to anyone because they are remarkably inexpensive. These categories are becoming more and more mainstream, and people are practicing them more than they used to. For instance, self-administered mindfulness meditation interventions have grown in accessibility and popularity over the past few years, possibly because of their seeming simplicity (Cavanagh et al., 2018).

The importance of studying more than one strategy relies on the fact that some strategies might work on some people and not on others; the difference may be due to the role of individual differences. For instance, De Vibe et al. (2015) found that participants higher in neuroticism showed more significant improvement in psychological distress and well-being compared to those in a control group after a Mindfulness-Based Stress Reduction (MBSR; Jon Kabat-Zinn, 1993). However, this strategy may not be particularly effective for high-scoring participants in extroversion, who may need to interact with people. For these same people, emotional social support might be a better strategy for reducing stress levels.

It is essential to understand which strategy has real effectiveness and which has not when it comes to regulating stress. Choosing the wrong strategy could have a null effect or adverse effects in the worst-case scenario (IJzerman et al., 2020). Additionally, it is important to understand for which type of individual a particular strategy works, and for which type of individual a particular strategy resonates in a different way.

Stress regulation through the process model of emotions

The strategies we have included in this Ph.D. project exploit different mechanisms to reduce individuals' stress levels and some of these mechanisms directly link to stress regulation processes of Gross's process model of emotions. For self-administered mindfulness, there is some evidence that practicing this form of meditation can improve the control of executive functions associated with better downregulation of stress (Moore & Malinowski, 2009; Raffone & Srinivasan, 2017; Teper & Inzlicht, 2013). Furthermore, some studies highlight how mindfulness can improve emotional regulation skills (e.g., Chambers et al., 2009). In fact, during the meditative practice, the individual can exploit a range of emotional regulation strategies (Gross, 1998). Indeed, with mindfulness, the individual can ward off the regret of having failed an exam by paying attention to their breathing (in this way, using attention deployment), or they can accept the outcome of an exam by not ruminating on it (in this case using cognitive change).

Biofeedback instead, can be considered a response modulation strategy because it directly influences the physiological response after the response tendencies have been already initiated. It helps recruit the parasympathetic branch of the nervous system and inhibits sympathetic action, leading to a regulated state and, therefore, to a significant reduction in the level of stress (Gummidela, Silva & Gutierrez-Osuna, 2021).

The first of the external regulation strategy: Being in nature can be used either as a situation selection or a situation modification strategy. Individuals seek nature to experience positive emotions and, at the same time, to avoid particular people, places, or objects to regulate emotions and reduce stress. In a similar vein, individuals can put an active effort to directly modify the situation to alter its emotional impact. For instance, people can relocate a discussion to a park to reduce tension (situation modification).

Finally, social support may refer to more than one stage of Gross' Process Model: An individual may decide to go out with friends who are more supportive than others (situation selection); after that he might decide to change the topic of a conversation to expose an issue he wants support on (situation modification); by discussing with his friends he might give different interpretation to that particular situation (cognitive deployment); and finally, spending time with a friend could be a distraction in itself from the stressful stimulus or situation (attentional deployment). In the chapters that follow we outline in detail the three empirical studies that are part of the current Ph.D. project.

Chapter 2: Stress regulation via being in nature and social support in adults, a Registered Report meta-analysis

Author Note: The preparation of this work was partly funded by a French National Research Agency ”Investissements d’avenir” program grant (ANR-15-IDEX-02) awarded to Hans IJzerman;PRIMUS/20/HUM/009, APVV-17-0418, and NPO Systemic Risk Institute (LX22NPO5101) grants awarded to Ivan Ropovik**.** Our OSF page can be found at https://osf.io/6wpav/. The funding sources had no role in the study design, collection, analysis or interpretation of the data, writing the manuscript, or the decision to submit the paper for publication. The authors that contributed to this manuscript are in order of contribution: Alessandro Sparacio, Ivan Ropovik, Gabriela M. Jiga-Boy, Adar Cem Lağap & Hans IJzerman

Abstract

In this meta-analysis, the authors investigated whether being in nature and emotional social support are reliable strategies to downregulate stress. We retrieved all the relevant articles that investigated a connection between one of these two strategies and stress. For being in nature we found 54 effects reported in 16 papers (total N=1,697, Mdn_N= 52.5), while for emotional social support we found 18 effects reported in 13 papers (total N=3,787, Mdn_N=186). Although we initially found an effect for being in nature and emotional social support on stress (Hedges' g = -.42; Hedges' g = -.14, respectively), the effect only held for being in nature after applying our main publication bias correction technique (Hedges' g = -.60). The emotional social support literature also had a high risk of bias. Although the being-in-nature literature was moderately powered (.72) to detect effects of Cohen's d = .50 or larger, the risk of bias was considerable, and the reporting contained numerous statistical reporting errors.

Keywords: Stress regulation, being in nature, social support, meta-analysis, Registered Report

Chapter 2: Stress regulation via being in nature and social support in adults, a Registered Report meta-analysis

How can we live in a fast-paced world where every unexpected challenge is just around the corner? Sometimes the obstacles are low or easy to get around; others may seem insurmountable. Life's obstacles can trigger a stress response that can be understood as, "a particular relationship between the person and the environment that is appraised by the person as taxing or exceeding his or her resources and endangering his or her well-being" (Lazarus & Folkman, 1984, p. 19). Stress experienced daily has an impact on health and on well-being of individuals (Bolger et al., 1989).

Thus, stifling the build-up of excessive stress is of paramount importance. In a previous meta-analysis, we synthesized empirical research on two stress regulation strategies (self-administered mindfulness and heart rate variability biofeedback; Sparacio et al., 2022). As we aim to build a comprehensive database in which different stress regulation strategies are evaluated based on their efficacy, here we add the synthesis of two other strategies: Being in nature and emotional social support. The reason why we chose these two strategies is like what guided the choice in our previous work: The decision was partly based on the fact that we were interested in analyzing scalable, non-invasive and cheap strategies that could be used by an extended number of individuals and partly arbitrary as to where we start with our approach. To check whether the named strategies have an effective role in reducing stress levels we conducted a meta-analysis with the following objectives: 1) To assess the evidential value of identified studies in both literatures, 2) for both being in nature and emotional social support, to calculate mean effect sizes for the stress response, for the different components of stress, as well as for the affective consequences of stress, 3) to apply publication bias correction techniques to have more realistic estimates of the efficacy of either regulation strategy 4) and to determine whether personality traits were used as moderators in stress regulation studies.

We intend to shed light on whether being in nature and emotional social support has stress reducing effects or not through our meta-analysis and how big the effect - if any - is. Our combination of publication bias-correction techniques can provide a less biased estimate of the effects of interest (Cf., IJzerman et al., 2022; Sparacio et al., 2022).

Stress Regulation

Stress is usually defined as a state of strain and tension that occurs when we are overwhelmed by external demands with the impossibility of dealing with them for lack of resources (Lazarus & Folkman, 1984). In our previous meta-analysis, we classified the stress response based on three components: Affective, physiological, and cognitive (see Du, Huang, An, & Xu, 2018; Schneiderman, Ironson, & Siegel, 2005; Sparacio et al., 2022; Watson & Clark, 1988). As we noted there, these different components are not truly conceptually separate (Pessoa, 2006; Phelps, 2008), but we apply them as useful categories for application. Because stress can have long-term consequences if not kept under control, we also included an assessment of the affective consequences of stress (such as depression and chronic anxiety). We decided to pick depression and chronic anxiety as relatively arbitrary starting points for constraints of time and resource and because those are traditionally the most investigated outcomes for these interventions.

The first strategy we focused on here, being in nature, we restricted to interventions like walking in a natural environment and/or watching it (Antonelli et al., 2019). According to the "stress recovery theory" (Ulrich, 1983), nature provides a restorative influence helping individuals recover from stress. Ulrich's (1983) theory relies on a psycho-evolutionary theorizing: Humans evolved in the course of centuries in natural places adapting both psychologically and physiologically to these types of environments. The argument is that when a stressor is encountered, an unthreatening natural environment might evoke feelings of pleasantness, decrease stressful thoughts, and promote physiological restoration (see also Ulrich et al., 1979).

In the empirical literature, being in nature has been found to have a positive influence on the different components of stress. For the affective component, one study found that participants that walked in a natural setting (as compared to when they walked in a built environment) had a greater reduction of levels of self-reported stress (Beil & Hanes, 2013). For what concerns the physiological component, in one study of coronary artery disease participants, who were randomly allocated to a seven-day walking in a park (vs. a seven days walking-in-an-urban-environment) condition had lower cortisol levels and lower heart rates (Grazuleviciene et al., 2015). As it pertains to the cognitive component, a brief walk in a natural setting (vs. 90 a minute walk in an urban setting) reduced self-reported levels of rumination (Bratman, Hamilton, Hahn, Daily, & Gross 2015). Finally, for what concerns the affective consequences of stress, one study found that a walk in a green area (as compared to a group of non-walkers) reduced symptoms of depression (Marselle et al, 2014).

The other strategy, emotional social support, has probably garnered the most empirical support out of the two (e.g., Cohen, 2004; Lakey & Cronin, 2008). Cohen and Wills (1985) suggested that social support can act as a shield protecting the individual from negative consequences of stress. There are two main models that explain the relationship between stress and close relationships. The first, the stress-buffering hypothesis, states that social support is connected to wellbeing by reducing stress appraisals or weakening the association between stress and negative health outcomes. The second, the main effect hypothesis, posits that social support has a beneficial effect, decreasing the level of distress, regardless of whether people are under stress (Cohen & Wills, 1985). The buffering effect has been thought to be associated with a dampened hypothalamic–pituitary–adrenal (HPA) axis activity and a decrease in the response of the autonomic nervous system (ANS; Carter, 1998).

One particular theory, "social baseline theory" (e.g., Beckes & Coan, 2011) offers an account that can provide a mechanism for the stress-buffering hypothesis, as it suggests that social support and proximity to others reduces the perceived threat of a stressor and people can thus exert less effort in regulating stress (Coan & Sbarra, 2015; Ein-Dor et al., 2015). Stress according to the theory, is reduced because individuals can distribute the efforts needed to achieve goals with other people (e.g., partner, friends, family members, or even strangers), a phenomenon known as "load sharing". In one particular study illustrating this phenomenon, people held hands with a partner or a stranger and were confronted with the threat of a (mild) electric shock. When people held hands with someone, areas related to stress were less activated when confronted with the electric shock and the reduction of stress was greater the more familiar the partner (Coan, Schaefer, & Davidson, 2006; Coan et al., 2017).

For the current Registered Report meta-analysis, we take a more narrow view on social support, as we restrict ourselves to emotional social support that is defined at a global level as the act of talking, listening, and being empathetic with a distressed individual (Zellars & Perrewé, 2001). Emotional social support can be achieved through verbal expressions (talking to or listening to the partner) or via physical contact (e.g., holding a partner's hand or talking with the partner; Coan, Schaefer, & Davidson, 2006; Ditzen et al. 2007). e leave out other forms of social support (informational, instrumental, and appraisal) as emotional social support is thought to be associated with well-being and consequently lower mortality and lower levels of stress (Reblin & Uchino, 2008).

For what concerns the affective component of stress, in one study participants' state anxiety decreased when emotional support was provided by a friend (compared to participants that did not receive any kind of support; Bowers & Gesten, 1986). In a study focused on the physiological component, participants that were assigned to a physical contact condition (as compared to the no social support condition) exhibited lower heart rate activation and cortisol response (Ditzen et al., 2007). For what concerns the cognitive component, one study found that participants with high levels of emotional social support responded to daily stressors with less ruminative behaviors (as compared to participants with low levels of emotional social support; Puterman, DeLongis, & Pomaki, 2010). Finally, as regards to the affective consequences of stress, studies have found that low levels of social support predict depression both in a non-clinical and clinical populations (Revenson et al. 1991; Brugha et al. 1987).

Bias in estimating the efficacy of stress regulation

How can we assess whether there is solid evidence on the efficacy of these strategies? Many fields of science, including psychology, have been confronted with a replication crisis (the fact that replication studies have failed to find the same results as original studies; Klein et al. 2018; Maxwell, Lau, & Howard, 2015; Open Science Collaboration, 2015). Publication bias (the likelihood that positive results have a higher probability of getting published; Rosenthal, 1979; Sutton, Duval, Tweedie, Abrams, & Jones, 2000) and questionable research practices (which is generally used as a term to encompass various scientific misconducts such as excluding data on the basis of post-hoc criteria; John et al., 2012) are often seen as two of the main culprits for low replicability rates.

The psychological literature therefore contains an unknown proportion of unreliable and false positive findings that may also characterize the field of stress regulation. For instance, in our previous meta-analysis, we analyzed whether self-administered mindfulness and biofeedback were effective strategies to decrease stress. We detected an effect for both strategies. However, when we applied the same publication bias techniques as we intend to apply here, we found no more evidence that self-administered mindfulness and biofeedback were successful in reducing stress. Indeed, our analyses suggest that the originally detected effect may have largely been due to publication bias (Sparacio et al., 2022). Thus, a thorough systematic assessment of the empirical evidence contained in the literature is needed (IJzerman et al., 2020).

At present, we have no way of knowing whether the two strategies are reliably effective interventions against stress. There are no current meta-analyses specifically on emotional social support and stress. Some meta-analyses do exist on the topic, but they need necessary improvements[1]. For being in nature, only one meta-analysis exists on a very specific type of being in nature, "forest-breathing" (Antonelli et al., 2019), which did not account for publication bias at all. We tried to improve upon these prior approaches by synthesizing up-to-date available evidence, as well as by applying state-of-the-art bias correction techniques. In so doing, we followed a workflow similar to our previous meta-analysis on stress regulation (Sparacio et al., 2022).

Method

To ensure methodological rigor and transparency, our materials and analysis code are available on the Open Science Framework (https://osf.io/6wpav/). As our goal is to build a database of data on different stress regulation strategies, we also added the data to PsychOpenCAMA, an existing public repository in which data from other meta-analyses are stored (Burgard, 2021). We already submitted data of our first pre-registered meta-analysis (Sparacio et al., 2022) to this platform on 24/09/2021. Our meta-analysis was pre-registered on the OSF (https://osf.io/c25qw). Any changes to the pre-registration were fully disclosed on our OSF page using the template provided by Moreau and Gamble (2020; Appendix A). This

[1] In our Stage 1 Registered Report, we had mistakenly referenced Schwarzer and Leppin (1989) and Harandi, Taghinasab, and Nayeri (2017) as focusing on emotional social support and stress. Schwarzer and Leppin (1989) focused on self-reported general health as outcome, whereas Harandi et al. (2017) focus on mental health.

research was conducted in line with the CO-RE Lab Lab Philosophy v5 (Silan, Bellemin, Dujols, Sparacio, Adetula, & IJzerman, 2021).

Inclusion criteria and search strategy

To frame the eligibility criteria in a structured way, we followed the Participants, Intervention, Comparator, Outcome, and Study design (PICOS) Framework (Schardt et al., 2007). We chose to only include studies on participants that are adults (people aged 18 years or older). For the current meta-analysis, we selected two interventions (being in nature and emotional social support). In case of designs comparing groups, for being in nature, we included effects based on a comparison to a control group in which participants performed the same activities (e.g., walking or viewing the surroundings) in an urban environment, or to a passive control condition (participants are in an untreated comparison group, e.g., waitlist control). For emotional social support, we included effects based on a comparison to an active control condition (in that participants were involved in tasks that were not related to stress regulation) and/or to a passive control condition. In case there were more sources of emotional social support for each study, we included the effect based on the closest connection with the participant (e.g., partners over friends, friends over strangers).

If there was more than one comparator in the same study (i.e., presence of both an active and a passive control group), we chose the contrast with the active control group. We measured the affective, the cognitive, and the physiological component of stress taken at post-test of both the experimental group and the control group. For the affective and cognitive components as well as the affective consequences, we relied on self-report measures. For the physiological component, we relied on physiological biomarkers of the stress response (e.g., heart rate, cortisol levels).

To ensure a search strategy that was reproducible, we documented 1) the exact search strategy 2) the dates on which the research was conducted 3) the exact search string. Our search strategy followed the recommendations provided by Maggio et al. (2011). The following databases were searched: ProQuest, (an online platform which covers research indexed in APA PsycArticles, APA Psycinfo, ProQuest Dissertations & Theses Global), PubMED, and Scopus. We searched the titles and abstracts of the articles.

The first author (AS) performed the literature search and excluded articles that did not match the inclusion criteria. Screening by title and abstract was carried out using Rayyan QCRI (Ouzzani, Hammady, Fedorowicz, & Elmagarmid, 2016), a web and mobile app for systematic reviews and meta-analyses. The first author then manually searched reference lists of the included studies for relevant citations and unpublished reports. Finally, we used social networks (Facebook groups and Twitter) and mailing lists (Society for Personality and Social Psychology; SPSP, European Association of Social Psychology; EASP, European Society for Cognitive and Affective Neuroscience; ESCAN; Environmental Psychology; ENVPSY) to request unpublished data. To ensure that we did not miss relevant articles, we also searched references of past meta-analyses related to being in nature and emotional social support. We included studies of existing meta-analyses that satisfied our inclusion criteria. Finally, we contacted authors that published studies on the topic to inquire whether they had any unpublished research, in-progress manuscripts, or in-press manuscripts (see our templates in Appendices B and C).

Following the inclusion criteria of our meta-analytical approach: 1) We included published articles, preprint articles, working papers, dissertations, and books (we excluded studies that were not published in English), 2) we included any type of study (randomized control trials and observational studies) that estimated the effect of (or exposure to) being in nature or emotional social support, 3) we included studies that measured at least one of the three components of the stress response or at least that measured the affective consequences of stress, 4) For being in nature we included studies with participants who performed any type of physical activity as long as the same activity was performed in the same way by the corresponding control group in a non-natural setting 5) the participants of the study had to be humans. A study was excluded 1) if it was a review (either narrative or systematic), 2) if the sampling frame of the study explicitly involved participants below 18 years of age, 3) if the data necessary to compute our analyses were missing (and not obtainable after having requested them to the authors of the paper), or 4) if other active treatments (e.g., mindfulness) were combined with the stress regulation strategies of interest (being in nature or emotional social support). We then added a sub-exclusion criterion related to emotional social support, excluding studies with types of support that were not emotional (i.e., informational or instrumental social support or social support via appraisal). A PRISMA flow chart of the overall literature search and inclusion procedure is shown in Appendices D and E.

Coding and Data Preparation

Two coders independently coded the data. We cross-checked the coding process for systematic coding errors twice – after the first 10% and 20% of the data – both for social support and being in nature separately. In case of systematic coding discrepancies, the two coders discussed, refined the coding scheme, thereby resolving discrepancies (in case this did not lead to convergence, the two coders consulted the second author). We used Cohen's Kappa as a measure of inter-rater agreement. Following the guidelines of Landis and Koch (1977), we considered an agreement of $\kappa > 0.60$ for metric or multinomial variables acceptable. For binary variables (e.g., published), we assessed the coding agreement using the percentage agreement.

We extracted data for the following variables: Publication year, the number of citations of the paper by Google Scholar at date of extraction, journal name, reported overall N, gender ratio, publication status, reported effect sizes, total *N*, cell means, standard deviations and *N*s, test statistic, degrees of freedom, the type of effect (e.g., bivariate effects, covariate-adjusted effects), whether the effect was considered focal (reported in the abstract), the design of the study, the type of population, the category of stress-regulation strategy (being in nature, emotional social support), the type of control group (no control group, active, passive, being in an urban environment, different source of emotional social support), whether it was on one of the components of stress (affective, cognitive, or physiological) or on the affective consequences of stress, and the instrument employed to assess stress levels. We converted all the relevant effect sizes (ES) to Hedges' *g*, a standardized mean difference corrected for small samples (Hedges & Olkin, 1985). To convert the reported effect sizes to Hedges' *g,* we primarily used the group posttest means, *SD*s (or *SE*s), and *N*s. If these were not available, we computed the Hedges' *g* effect size from the reported test statistics or converted from other types of reported effect sizes. The computation and conversion of all effect sizes were carried out in R, using formulas laid out in Borenstein et al. (2009; analysis code available at: https://github.com/alessandro992/Registered-report-meta-analysis).

To mitigate the effect of undisclosed participant exclusions, we checked whether the sum of group *N*s approximately matched the total sample size (*N* +/-2). If they matched, we used the reported group *N*s. If the sum of group *Ns* did not match the total sample size, we computed group *N*s based on the reported degrees of freedom, assuming a balanced design. If only the total sample size was reported, we also assumed a balanced design and divided the total *N* by the number of conditions. We applied by default a correlation of .50 for within-participants designs.

Analyses

Analysis Strategy

Our analysis strategy closely mirrors the workflow of IJzerman et al. (2022) and Sparacio et al. (2022). Prior to conducting our analyses, we screened for influential outliers using a Baujat plot (Baujat et al., 2002) and influence diagnostics indices (Viechtbauer & Cheung, 2010). Outliers with an excessive influence on the meta-analytic model (standardized residual > 2.58) were then excluded in a sensitivity analysis. By default, we used a multilevel random-effects model using the restricted maximum-likelihood estimation with Satterthwaite's small-sample adjustment (Pustejovsky & Tipton, 2021)[2]. We included all the relevant outcomes from each included study. We handled dependencies among effects by using robust variance estimation, assuming correlated and hierarchical effects (Pustejovsky & Tipton, 2021). By relying on robust variance estimation, we could simultaneously account for both types of dependencies among the effects (if effects were nested within studies, this technique allowed us to estimate effects based on the same participants). Because the data on sampling correlations among effects tend to be unavailable in the individual studies, we assumed a constant sampling correlation between the effects of .5 (see also, Kolek et al., 2022, Sparacio et al., 2022). We used a robust HTZ-type Wald test to test the equality of effect sizes across the levels of the studied moderators (Pustejovsky & Tipton, 2021).

To estimate the range of effect sizes that can be expected in similar future studies, we calculated the 95% prediction intervals. For each analysis we conducted, when the included effects (k) were less than 10, we did not interpret the estimates[3]. This threshold is somewhat arbitrary, but a threshold needs to be chosen. After all, large samples have large expected sampling variability, leading to imprecise results (see also Sparacio et al., 2022; IJzerman et al., 2022).

[2] We switched to ordinary two-level random-effects model if (1) the multilevel model failed to converge in the overall model or in any of the subgroups; or (2) if the variance components of the model were not well identifiable (specifically, if the log-likelihood did not peak at the variance estimates for both variance components).

[3] 10 is to be considered for the total effects by type of subgroup analysis, not by category. For instance, if there are 5 studies on physical social support and 15 on verbal social support we will conduct the relative subgroup analyses, as the total number of effects is 20. However, if the total number of effects is below 10, we will not run that subgroup analysis.

To investigate the heterogeneity caused by variations in population characteristics or conceptual aspects of utilized study designs, we pre-registered a set of subgroup analyses for both being in nature and emotional social support: Proportion of females (versus males) in the sample, type of comparison group, and type of population (student non-clinical, non-student non-clinical, and clinical). For being in nature, we tested the type of exposure as a possible source of heterogeneity (nature walking, nature viewing, mixed). For emotional social support, we conducted two additional subgroup analyses: The type of social support (0=not specified, 1=physical, 2=verbal, 3=mixed, 4=other) and the source of social support (0=not specified, 1= stranger, 2=known person; see for more details our coding sheet; https://osf.io/4cjux/). Although we believed a priori that this coding to be exhaustive, if we realized that our coding sheet was inadequate throughout the coding process, we refined our coding scheme and we would document these changes in Appendix A: Protocols and deviations sheet. This happened for being in nature for which in studies where participants were exposed to natural settings through images, videos, or virtual reality, we introduced the term "virtual seeing." Finally, we ran two moderation analyses to assess whether studies with high risk of bias and mathematically inconsistent means or *SDs* showed inflated effect sizes. In case of additional non-pre-registered subgroup analyses, we disclosed them on our OSF page using the template provided by Moreau and Gamble (2020; Appendix A).

The R code also allows the reader to easily change numerous arbitrary values (e.g., the assumed constant sampling correlation, the within-subjects correlation, etc.) to explore the impact on the results. All models were fitted using restricted maximum-likelihood estimation using R packages metafor, version 2.5 (Viechtbauer, 2010) and clubSandwich, version 0.4.2. (Pustejovsky, 2020). The data analysis was carried out in R also using the following packages: Esc (Lüdecke, 2017), tidyverse (Wickham et al., 2019), lme4 (Bates, Maechler, Bolker, Walker, 2015), dmetar (Harrer et al., 2021), and psych (Revelle, 2018).

Correction for publication bias

Null or negative results are typically less likely to be published, leading to a biased sample of included studies. Such publication bias tends to lead to an inflation of the observed mean effect sizes and the Type I error rate (Carter et al., 2019; Hong & Reed, 2020; Ioannidis, 2008). In an effort to adjust the meta-analytic estimates for publication bias, we primarily used a selection modeling approach (McShane, Böckenholt, & Hansen, 2016).

We employed a 3- or 4-parameter selection model (4PSM; McShane, Böckenholt, & Hansen, 2016) and used it as the primary inferential and estimation bias-adjustment method. Selection models are a statistically principled family of models that directly model the publication selection process. The 4PSM implementation has two components: A data model of two parameters that describes how data are generated in absence of publication bias (effect size and heterogeneity parameters) and a selection model mimicking the publication process, represented by a weight parameter–likelihood that a study with non-significant results is published compared to a study with significant findings and a parameter reflecting the likelihood of the result being in the opposite direction (McShane et al., 2016). If a given set of results yielded less than four focal *p*-values per interval, the model dropped the fourth parameter to provide for a more stable estimation. To deal with dependencies in the data and avoid arbitrariness in the selection of effects within studies, we applied a permutation-based procedure, iteratively selecting only a single focal effect size from each independent study, estimating the model in 5,000 iterations, and then picking the model yielding the median ES estimate (where both the interpretation and inference was based on that median model)[4].

To further explore the results of publication bias-adjustment, we did the following: First, we assessed the variability in adjusted estimates under different assumptions of the

[4] That is, we picked the median estimate from the parameter distribution and, with it, the corresponding model that the estimate was originating from. The goal of this procedure was to preserve the mutual consistency between the estimate, *z*-value, *CI*s, and *p*-value.

publication selection process using Vevea and Woods' (2005) step-function models with a priori defined selection weights (instead assessing them via estimates of maximum likelihood). These step-function models allowed us to explore the results by varying the assumed severity of bias, modeling moderate, severe, and extreme selection.

Second, we employed a multi-level RVE-based implementation of the PET-PEESE model (see IJzerman et al., 2022; Sparacio et al., 2022), having the same hierarchical structure as the random-effects models. PET-PEESE regresses the effect size on a measure of precision. Because larger studies are less likely to stay the unpublished, model slope is assumed to indicate the presence of small-study effects (this includes publication bias). On the other hand, the model intercept can then be interpreted as an average ES for a hypothetical, infinitely precise study (Stanley & Doucouliagos, 2014). To use a measure of precision that is uncorrelated with the effect size, we used $\sqrt{(2/N)}$ and a $2/N$ terms instead of standard error and variance for PET and PEESE, respectively[5].

Third, we used a robust Bayesian model-averaging approach to integrate the selection modeling and regression-based approaches and let the data determine the contribution of each model by its relative predictive accuracy to fit the observed data (Bartoš et al., 2021). This approach effectively dodges the need of choosing among competing approaches – and commits us to only a single set of assumptions about the nature of the true biasing selection process. Substantive interpretations were guided by the estimates and inferential results of the 4PSM solely. The other exploratory bias-adjustment methods served a descriptive purpose, to provide the reader with a more comprehensive view on bias adjustment under quantitatively

[5] As the 4PSM tends to have more favorable error rates under many conditions than PET, the reader can also define the 4PSM as a conditional estimator for PET-PEESE instead of the traditional PET in the R code, to explore the effect of such decision on the resulting inference.

and qualitatively different assumptions (Vevea & Woods models and PET-PEESE, respectively) and using a more general, Bayesian model-averaging approach (RoBMA[6]).

The detailed specification of the employed models can be found in code in the supplementary materials. There, we also report the results for the following bias-adjustment methods: *P*-uniform* (Van Aert & Van Assen, 2021) and the Weighted Average of the Adequately Powered studies (WAAP-WLS; Stanley et al., 2016). A summary of the workflow employed to account for publication bias can be found in Appendix F.

The quality-of-evidence assessment

As one of the main objectives of every meta-analysis should be to appraise the quality and integrity of the underlying reported evidence, we assessed the risk of bias at the study level, assessed the evidential value by looking for indications of *p*-hacking, looked for numerical inconsistencies in the reported data, and estimated the average power in the literature to detect various magnitudes of effects.

First, we evaluated the study quality using the Revised Cochrane risk of bias tool for randomized trials (RoB 2; Sterne et al., 2019). This tool assessed the risk of bias in five predetermined domains related to the experimental design and methodology of the study in question (e.g., the randomization process or the measurement of the outcome). Based on the judgment for each individual domain, an overall algorithmic-based judgment on the risk of bias was drawn up (i.e., “high risk-of-bias”, “some concern”, or “low risk-of-bias”). The rater had the right to override the suggested risk of bias judgments when justified only by downgrading the judgment.

Second, we assessed the evidential value in a set of significant findings, using the *p*-curve method (Simonsohn, Nelson, & Simmons, 2014). A right-skewed distribution of

[6] Apart from reporting the results of these bias adjustments, we examined whether the primary 4/3-PSM estimate fell within the 95% credible interval of the RoBMA estimate (being based on a more general model).

significant *p*-values indicates evidential value, i.e., that selective reporting is not the sole explanation of the observed findings. Conversely, a left-skewed *p*-curve points to a substantial prevalence of selective reporting or other forms of questionable research practices. To handle the dependencies among the *p*-values derived from the same sample, a permutation-based procedure was employed. We recomputed all focal *p*-values from the reported descriptive or test statistics, randomly extracted only a single effect size for each set of interdependent effects, estimated the *p*-curve in 200 iterations, and averaged over the set by interpreting the model having the median *z*-score for the right-skew of full *p*-distribution.

Third, we checked for numerical inconsistencies in the reported means and *SD*s using the GRIM (Brown & Heathers, 2016) and GRIMMER (Anaya, 2016) tests, respectively, and *p*-values. In case of discrete variables (e.g., Likert scales), decimals in means and *SD*s follow a granular pattern for each combination of *N* and the number of items, which makes it possible to identify instances where a given mean or *SD* is mathematically impossible given the reported *N* (Anaya, 2016; Brown & Heathers, 2016). We also screened the entire included papers for inconsistencies in the reported *p*-values using the statcheck package (Epskamp & Nuijten, 2018). The method works as follows: (1) Article pdf files are converted to plain text, (2) they are scanned for statistical results reported in APA style, (3) test statistics and degrees of freedom are extracted to recompute the *p*-value, (4) which then gets compared to the reported *p*-value. We examined in which proportion of primary studies were the *p*-values inconsistent with the reported test statistics and how many of those inconsistencies led to an inferential decision error.

Fourth, we computed mean statistical power in the literature to detect various hypothetical effect sizes (.20, .50, and .70). In the supplementary materials, we also report median power to detect the bias-corrected estimates based on the 4PSM and PET-PEESE models.

Results

The final meta-analytic dataset comprised 54 effects reported in 16 papers on being in nature (total N = 1,697, Mdn_N = 52.5, M_{Age} = 29.54, SD_{Age} = 7.75) and of 18 effects reported in 13 papers on emotional social support (total N = 3,787, Mdn_N = 186, M_{Age} = 47.72, SD_{Age} = 14.92). For being in nature, the included studies were published between 1993 and 2021. Studies on emotional social support were published between 1997 and 2021.

Do being in nature and emotional social support reduce stress?

Table 1 lists the goals and conclusions of our analyses meta-analysis.

Objective	Conclusion
1. To assess the overall empirical evidence of a) being in nature and b) emotional social support	We found an overall effect for a) being in nature and for b) emotional social support.
2. To assess the mean effect sizes for a) being in nature and b) emotional social support on the stress response, on the different components, and on the affective consequences of stress.	We found a) a naïve meta-analytic estimate for being in nature on stress (g = -.42), on the physiological (g = -.31), and affective (g = -.49) components and b) a naïve meta-analytic estimate for emotional social support (g = -.14) on the physiological (g = -.26), and affective (g = -.11) components. We were unable to calculate an estimate for both strategies on the affective consequences of stress as we had too few effects (respectively k = 6 for being in nature and k = 4 for emotional social support).
3. To apply publication bias correction techniques to have more realistic estimates of the efficacy of a) being in nature on the stress response and of b) emotional social support on the stress response.	Once we applied the 4PSM, we still found an effect for a) being in nature (g = - .60), but b) not for emotional social support (g = -.01).

4. To determine whether personality traits were used as moderators in a) being in nature and b) emotional social support.	Personality traits were almost never used as moderators in the studies of stress for a) being in nature (n = 0) or b) emotional social support (n = 1).

Table 1: Goals and conclusions of the Registered Report meta-analysis. Effect sizes are reported in Hedges' *g*.

Overall, we found an effect for being in nature that held after excluding improbably large effect sizes and after applying the correction for publication bias. Being in nature also reduced stress for both the affective and the physiological components. For emotional social support, we also found a significant overall meta-analytic estimate, but the effect disappeared after correcting for publication bias. Personality traits were almost never examined. In what follows, we present the results in greater detail separated by strategy (being in nature and emotional social support). We also present our pre-registered subgroup analyses as part of our auxiliary goals.

Does being in nature reduce stress? (Goals 1-3)

First, we investigated whether excluding outliers had a material effect on our main conclusions. For being in nature, there were five excessive effects above Cohen's d = 2, reported in three studies. One of the studies reported a single effect size of Cohen's d = 4.82, while a second study reported three effects, all of them being improbably high (-2.64, -2.60, and -2.07), and the third study reported a (highly influential, based on large N) effect size of 2.43[7]. We therefore decided to deviate from our pre-registration and excluded these outliers, given that they were so unrealistically large. We then proceeded with our pre-registered

[7] The given study also reported another three large effects (1.28, 1.12, and 0.87). Although not improbable on their own, they exerted considerable influence on the models, as they were based on samples with large N. These effects were not excluded in this sensitivity analysis.

overall, bias-corrected, and component-specific effects, as well as our quality assessment of the literature.

Overall effects of being in nature on stress

After excluding outliers, the naive meta-analytic estimate was Hedges' g = -.42, 95% CI [-.64, -.21], p <.001, suggesting that this strategy may be effective in reducing stress levels. Forty-four percent of the coded effects were statistically significant. The 95% prediction interval (i.e., the effect size expectation for a newly conducted study) was quite wide, [-1.4, .56]. This was due to a high heterogeneity, τ = .45, with an I^2 = 85.75% meaning that more than half of the observed variance was due to true heterogeneity (48.83% due to between- and 36.92% due to within-cluster heterogeneity). Table 2 summarizes the results for the overall effect of being in nature and emotional social support. Contour-enhanced funnel plot and forest plot are displayed in Figure 1.

	k	*g* *[95% CI]*	*SE*	τ	I^2	4PSM estimate	4PSM *p*-value	PET-PEESE estimate	PET-PEESE *p*-value
Being in nature	54	-.42 [-.64, -.21]	.11	.45	85.75%	-.60 [-1.02, -.18]	<.001	-.44 [-.87, -.02]	.04
Emotional social support	18	-.14 [-.24, -.04]	.05	.16	84%	-.01 [-.26, .05]	. .195	-.11 [-.41, .2]	.46

Table 2: Meta-analysis for being in nature and emotional social support. Values in brackets represent 95% CI.

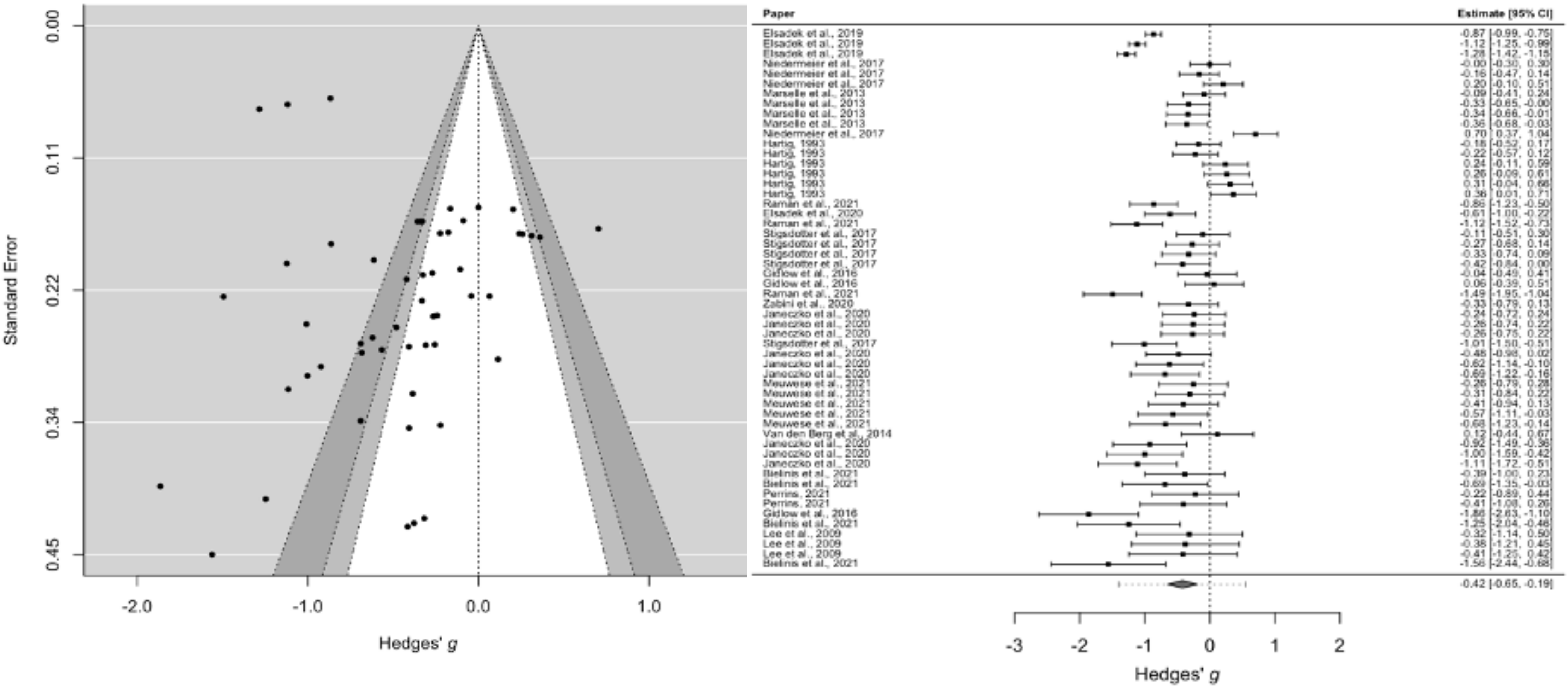


Figure 1: Contour-enhanced funnel plot and forest plot for being in nature after outlier exclusions.

Effects of being in nature on stress after publication-bias adjustment

Our primary publication bias-correction technique, the 4PSM, indicated an effect of being in nature on stress, with Hedges' g = -.60, 95% CI [-1.02, -.18], p = .006. According to our predetermined inferential criteria, we thus concluded that being in nature was effective in reducing participants' stress levels. We then used the Vevea and Woods step function models with a priori defined selection weights denoting moderate/severe/extreme selection to examine the variability in the bias-adjusted estimates. The results suggested that with a rising severity of the assumed selection bias, the effect of the intervention became smaller (and even reached an opposite direction under extreme selection), with estimates of -.31, -.19, and -.01 for moderate, severe, and extreme selection, respectively (all the estimates were rather imprecise and thus non-significant, with p-values equal to .01, .17, and .95, respectively). In other words, the higher the severity of publication bias the smaller the estimate of the effect gets, which implies that publication bias had a significant impact on the literature of being in nature[8].

Stress component-specific effects of being in nature

For being in nature, we categorized k = 28 effects as falling in the affective component of stress (with an effect size of Hedges' g = -.49), and k = 20 effects as failing into the physiological component (Hedges' g = -.31). We did not categorize any effect as falling into the cognitive component. The difference between components was not significant (Wald's test p = .47). Finally, we classified a set of k = 6 as being part of the affective consequences of stress; as this fell below k = 10, we did not analyze this set.

Quality-of-evidence assessment for being in nature and stress

---

[8] In addition, we performed exploratory analyses to determine how the outcome changed when we used other publication bias adjustment methods that we had previously registered. The PET-PEESE did not detect a signal Hedges' g = -.55 , 95% CI [-1.27, .17], p = .13. The RoBMA model instead suggested the presence of an effect Hedges' g = -.42, 95% CI [-.57, -.28].

Figure 2 displays the risk of bias for the included studies overall and for each of the five risks of bias dimensions.

Figure 2: Overall risk of bias and risk of bias assessment for each of the five dimensions for being in nature.

Judging the risk of bias using the RoB2 tool, 25% of the studies were rated to be at low risk of bias overall, the majority (50%) showed some concerns, and a smaller proportion (25%) were deemed to have a high risk of bias. We then proceeded with the assessment of the evidential value with a *p*-curve test. The full and half *p*-curve tests ($z = -6.68$ and $z = 5.69$, respectively) were significant (both $p < .001$) hinting at the presence of evidential value. However, because of the need to iteratively permute only a single focal effect from each study, the median model was based only on $k = 4$ effects.

After that, we screened the set of included papers with the statcheck package. Of the included papers, only $n = 8$ (47%) reported results in APA format. The 96% of the included effects were reported correctly, with only 4% having statcheck errors. Half of these errors changed the nature of the resulting statistical inference (results reported as significant with the actual recomputed $p > .05$). Regarding the presence of mathematically inconsistent means and *SD*s, $n = 30$ (51%) of the coded effects were derived from group means and *SD*s. Out of those effects, 63% were fully consistent with the reported cell sizes. In the remaining 37%, either the *SD* or both, mean and *SD* were mathematically inconsistent. The exclusion of these effects only led to a negligible change in the estimated *ES* ($\Delta g = .03$).

Furthermore, studies on being in nature were not well-powered to detect the full range of hypothetical, theoretically relevant effects sizes. More specifically, the average power in the set of included studies to detect small effects (Cohen's $d = .20$) was quite low (.17), but moderately powered (.72) to detect effects of Cohen's $d = .50$ or larger. Overall, we thus conclude that being in nature leads to a reduction of stress.

Is the reduction of stress by being in nature moderated by personality traits? (Goal 4)

We wanted to investigate whether personality traits (e.g., neuroticism) were used in studies on stress regulation for both strategies. For being in nature, none of the included studies assessed personality traits as a potential moderator. We were thus unable to assess whether personality moderated the effect of being in nature on stress.

Being in nature: Pre-registered moderator analyses (Auxiliary Goals)

Next, we examined several subgroup analyses, which will help us further understand the sources of heterogeneity in the literature. Please note that, due to the limited number of included studies, we were unable to apply publication-bias correction techniques and the results below should thus be interpreted with care.

Investigating heterogeneity of the being-in-nature effect: Characteristics of the population

We did not find that type of sample was related to the magnitude of the effect (Wald's test, $p = .5$). For being in nature, the mean proportion of female participants across the included studies was 50%. We found a slightly negative effect of gender on the efficacy of being in nature in reducing stress ($B = -.01$, $p = .02$). Although sampling is rarely representative, the effect of gender could at least partially explain the heterogeneity of our effects.

Characteristics of the being-in-nature intervention

For being in nature, we investigated whether the effect varied as a function of the type of exposure in the natural environment. For this subgroup analyses we modified our coding scheme by adding a category that could include some studies that were left out with the previous coding scheme[9]. The majority of the effects came from studies in which participants walked in a natural environment (k = 23), in a sizable portion (k = 18), participants were in a mixed condition (nature viewing and walking). In the remaining set of effects, participants were in a natural viewing condition through a virtual medium (k = 7) or they were physically present in a green environment (k = 6). The effect sizes of these groups were not significantly different, p = .70.

Does emotional social support reduce stress? (Goals 1-3)

For emotional social support, none of the effects were deemed an outlier. We therefore immediately proceeded with our pre-registered overall, bias-corrected, and component-specific effects, as well as our quality assessment of the literature.

Overall effects of emotional social support on stress

The naive meta-analytic estimate suggested the presence of an effect and 55% of the effects included in our synthesis were significant. The naive meta-analytic estimate was Hedges' g = -.14, 95% CI [-.24, -.04], p < .001, suggesting that emotional social support is effective in reducing stress. The 95% prediction interval was large, with the true effect in a new published study being expected in the range from -.51 to .23. The heterogeneity for emotional social support was considerable, τ = .16, I^2 = 84% (73.25% due to between and

[9] We added "virtual viewing" to indicate studies in which participants were exposed to natural environments via photos/videos /virtual reality. We documented this change in Appendix A: Protocols and deviations sheet.

10.82% due to within-cluster heterogeneity). Contour-enhanced funnel plot and forest plot are displayed in Figure 4.

Effects of emotional social support on stress after publication-bias adjustment

Our primary publication bias-correction technique, the 4PSM, failed to find an effect for emotional social support, Hedges' g = -.01, 95% CI [-.26, .05], p = .195. We thus concluded that there is a lack of evidence to support the efficacy of emotional social support in reducing stress. Furthermore, the sensitivity analysis employing the Vevea and Woods (2005) step-function model failed to find an effect of the emotional social support at -.09 in case of moderate selection (p = .08), at -.04 under severe selection (p = .46) and at .02 when assuming extreme selection (p = .69). That suggests that the adjusted lack of effect was thus very stable regardless of the assumed functional form of the selection mechanisms[10].

Stress component-specific effects of emotional social support

For emotional social support, almost all effects fell in the affective component (k = 13; Hedges' g = -.11), while only four effects fell into the physiological component (and we thus did not analyze the results). None of the coded effects for this strategy were categorized as being part of the cognitive component. That made any statistical comparison impossible. Finally, for emotional social support, a small proportion of the effects k = 4 was considered as being part of the affective consequences of stress (where we again did not analyze the results).

Quality-of-evidence assessment for emotional social support on stress

We then evaluated the risk of bias assessment for emotional social support. The majority of the studies (95%) were assessed as having a high risk of bias, due to the fact that

[10] We also ran exploratory analyses to see how the effect varied when applying other publication bias adjustment techniques that we pre-registered. The PET-PEESE failed to find an effect, Hedges' g = -.11, 95% CI [-.41, .20], p = .46; we reached a similar conclusion with the RoBMA model Hedges' g = -.21, 95% CI [-.41, .00].

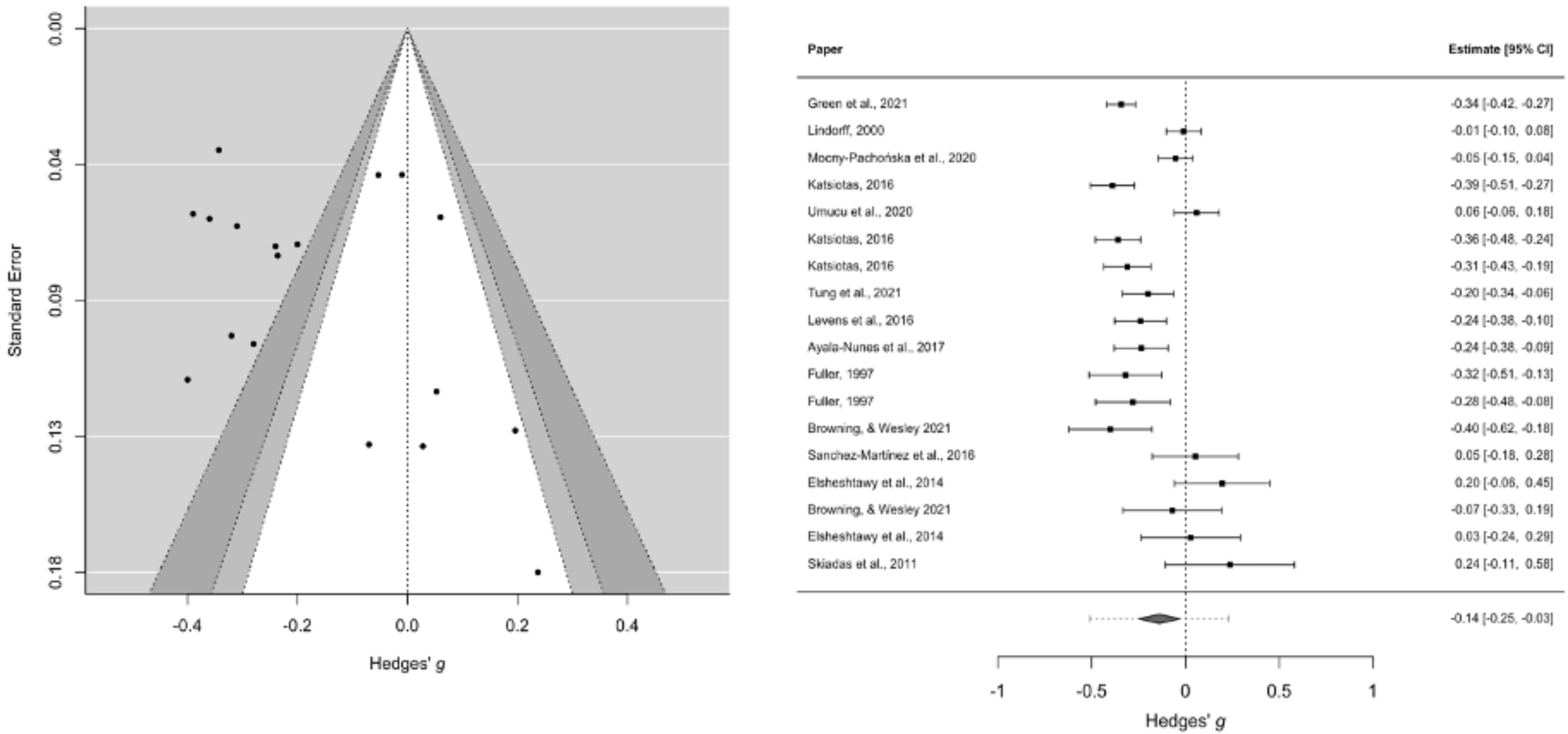


Figure 4: Contour-enhanced funnel plot and forest plot for emotional social support.

most of the studies were observational and therefore not randomized. We provide an overview of the risk-of-bias assessment for emotional social support in Figure 5.

P-values for both the full *p*-curve ($z = -8.35$, $p < .001$) and the half *p*-curve tests ($z = -7.67$, $p < .001$) were significant, suggesting the presence of evidential value in the given set of included significant focal effects (the *p*-curve test was based only on $k = 6$ independent effects). We then screened the included studies with the statcheck package for inconsistencies in reported test statistics and *p*-values. Only a minority of the included papers, $k = 3$, 23% of the total, were reported in a standard APA style, and two of the papers contained at least one reporting inconsistency.

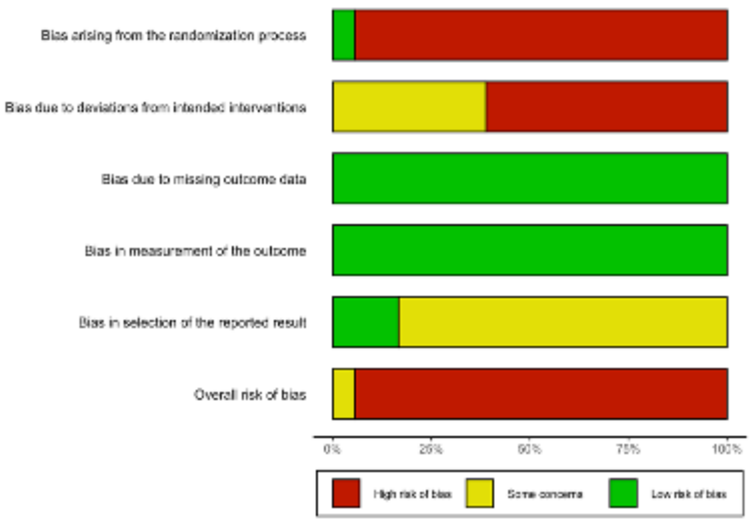


Figure 5: Overall risk of bias and risk of bias assessment for each of the five dimensions for emotional social support.

Eighty-five percent of the reported results were not flagged as statistical errors by the statcheck screening, while in the remaining 15%, the reported results were statistically inconsistent. For two-thirds of those errors, the reported *SD*s, means, or both, were mathematically inconsistent, which thus means a low-quality of statistical reporting. None of the synthesized means or *SD*s were inconsistent with the reported sample size.

The average power in the set of included studies for emotional social support to detect small effects (d = .20) was .49, but was more than adequate (.99) by conventional criteria to detect effects of d = .50 or larger. Overall, we conclude that we cannot find evidence in favor of emotional social support reducing stress.

Is the (lack of) reduction of stress through emotional social support moderated by personality traits? (Goal 4)

Despite this lack of support, we did examine whether personality differences could be of relevance. For emotional social support however, only one study (yielding two effects) studied trait self-esteem as the moderator of the relationship between emotional social support and mental health. Thus, based on the lack of evidence in the primary literature, it is not possible to know whether the lack of the effect of emotional social support is due to individual differences.

Emotional social support: Pre-registered moderator analyses (Auxiliary Goals)

Next, we will again examine several subgroup analyses, which will help us further understand the sources of heterogeneity in the literature. Please note again that, due to the limited number of included studies, we were unable to apply publication-bias correction techniques and the results below should thus be interpreted with care.

Investigating the lack of the emotional-social-support effect: Characteristics of the population

Less than half of the overall sample was female, on average (43.45%). There was no significant effect of the proportion of men versus women as a moderator on stress reduction (B = .001; p = .13). Next, we investigated whether the efficacy of emotional social support differed by the type of population. For this strategy, we had an equal number of effects for each different type of population (k = 6; student non-clinical, non-student non-clinical, clinical); we again did not find a significant moderating effect (Wald test p = .29). Neither gender or type of population could thus explain the heterogeneity (and thus the lack) of the effect.

Characteristics of the emotional-social-support intervention

We then investigated whether the types of support differed in their ability to decrease participants' stress levels. For most of the effects ($k = 14$), the source of emotional social support was not specified. In a negligible portion of the effects ($k = 2$), the emotional social support took the form of physical contact, and for $k = 1$, the type of social support was both verbal and mixed (being the set of effects less than 10, we did not analyze these results). Concerning the source of emotional social support, it was not specified for $k = 16$ of the effects, while for $k = 1$, the support came from a known person, and for $k = 1$ the support came from a stranger.

Overall quality checks of the being-in-nature and emotional-social-support literatures: Study designs and sensitivity analyses

Study design characteristics

We investigated whether the effect sizes of the included studies varied as a function of the characteristics of the experimental design. First, we verified whether effects related to the two strategies were different in relation to the type of control group employed (active vs. passive). For both strategies, the Wald's test was not significant, $p = .12$ and .17 for being in nature and emotional social support, respectively, suggesting that the type of comparison group was not associated with the magnitudes of the effect sizes.

Second, we tested whether the effects of being in nature and emotional social support varied as a function of the published status of the included articles. For being in nature the majority of the effects were extracted from published studies ($k = 51$), while a minor portion were extracted from non-published studies ($k = 8$). The difference between these two groups was not significant ($p = .23$). For emotional social support the 38.9% of the effects ($k = 7$)

came from the gray literature, while the 61.1% of the effects came from the published literature. The difference between these two groups was significant (Wald's test, p = .03) indicating that unpublished studies yielded a larger effect (Hedges' g = -.32) than published studies (Hedges' g = -.09), on average.

Sensitivity analyses

We finally conducted a set of sensitivity analyses related to our main results to assess how methodological factors were related to the magnitude and precision of the reported effects. First, we excluded effects based on inconsistent means and/or *SD*s. For being in nature 11 effects were excluded. However, the differences between effects that were inconsistent and those of were not, was not significant (Wald's test, p = .91). For emotional social support, we did not carry out this analysis as none of the coded effects had inconsistent means or *SD*s. Finally, we checked how the *ES* varied in relation to the risk of bias assessment.

The difference in terms of *ES* between studies judged to be at a high risk of bias (k = 4) was not different as compared to studies judged to have low/some concerns risk of bias (k = 13) as Wald's test was not significant, p = .85. For social support, we were unable to carry this sensitivity analysis as we had only one effect with a risk of bias that was acceptable.

Discussion

Through a Registered Report meta-analysis, we evaluated the efficacy of two stress regulation strategies: Being in nature and emotional social support. After applying our main publication-bias correction technique (and after excluding outliers with improbable effects), we only found an effect for being in nature on decrease of stress. In what follows, we outline our assessment of the quality of both literatures, interpretation of some (potentially contradictory results), limitations of our assessment and of the literatures, key areas of improvement, and some concluding remarks.

Quality assessment of the being-in-nature and emotional social support literatures

Being in nature and stress: Risk of bias. While we conclude there to be an overall effect of being-in-nature interventions on stress reduction, the being-in-nature literature is not entirely without challenges: 50% of the studies were at some risk of bias and 25% were at high risk of bias. For being in nature, the majority of high risk of bias studies were due to potential bias arising from the randomization process. Specifically, 4 studies (out of 16) were judged to have a high risk of bias, while 3 had a medium risk, suggesting that participants were not properly randomized to avoid the influence of known or unknown prognostic factors on the final results.

Many of these risks seem easily fixable. One prominent example, Marselle et al. (2013), surveyed groups walking in different environments on mental well-being, depression, and self-reported stress. In such a case, we recommend authors to use experimental methods to bolster the strength of causal inferences and/or to reduce potential confounds by adding predictors and using conditional random forests as an analysis method to identify the most important predictors of the projected outcome (cf., Wittmann et al., 2021; Szabelska et al., 2022).

Another (medium) source of bias for being in nature arose from the domain, “deviation from intended interventions” where 12 studies were judged to have a medium risk of bias mostly due to a lack of blinding of participants. In one of such studies, Lee et al. (2009), a partial solution was found: Participants visited forest and urban environments and were surveyed for how comfortable they felt and how soothed and refreshed they felt. While these variables are obvious candidates for being influenced due to demand characteristics, the researchers also collected salivary cortisol, diastolic blood pressure, and pulse rate, variables which are all much less likely to be influenced by demand characteristics.

Being in nature and stress: Statistical challenges. Further, while there were statistical problems in the being-in-nature literature (with 4% having statcheck errors and 37% had mathematically inconsistent results), these results did not materially change the outcome of our meta-analysis. Nevertheless, greater care should surely be taken to ensure test statistics are correctly reported, as these numbers are worryingly high. All of this together, however, with the fact that the literature was sufficiently powered, provides us with sufficient confidence that being in nature in fact reduces stress, at least for specific populations.

To provide a more conclusive response on the effectiveness of being in nature and just *how much* being in nature reduces stress we recommend the adoption of pre-registration, or, better Registered Reports across the literature. Registered Reports will provide a much more accurate estimate of the effect sizes in this literature, and therefore, the exact efficacy of being in nature. It will therefore allow for a much better comparison to other interventions (such as biofeedback or self-administered mindfulness; Sparacio et al., 2022a, b). Finally, posting results of stress management techniques in an open repository (e.g., PsychOpen CAMA; Burgard et al., 2021) would allow other researchers to re-analyze the data and verify the validity of the results.

Emotional social support and stress: Risk of bias. The literature on emotional social support showed no better results in terms of quality: The literature is at high risk of bias (95% of the studies were at high risk; only one of the studies was a randomized controlled trial). In 12 out of 13 studies, problems arose from the randomization process as most of the studies were observational, making almost all the studies to have an overall high risk of bias.

Of course, emotional social support is much harder to (ethically) manipulate than being in nature. It is not easy to change the nature of one's social network and to change the level of one's emotional social support. For instance, in the study of Levens et al., (2016), "one hundred and eight-one freshman undergraduate participants completed questionnaires assessing depressive symptoms, family and instrumental support, and perceived stress reactivity" (p. 342). How would a researcher reasonably assess the causal relationship between emotional social support and stress?

For this, we have two recommendations. First, like for being in nature, it is possible to improve the quality of causal inferences by surveying participants on a host of additional (potentially relevant) variables (such as neuroticism, attachment security, the quality of one's social network, and so forth), and thereafter applying analytic techniques that are less prone to overfitting and to problems with collinearity (such as conditional random forests; Szabelska et al., 2022). Second, one could attempt to improve the quality of emotional social support from a support-giver by letting dyads participate in relationship-focused therapy (see, for instance, Johnson et al., 2013) and comparing this with an active and passive control condition.

More generally, to better map how emotional social support (if at all) and being in nature affect stress (in experimental or cross-sectional studies), we think that a much better record is needed. We therefore strongly recommend assessing a number of different variables to study how these effects differ across situations and across different people (see Appendices G and H).

Emotional social support and stress: Statistical challenges. Furthermore, the literature seemed to contain a considerable number of problems with statistical reporting: in 15% of the cases, test statistics and p-values were mathematically inconsistent, while in 2/3 of those 15% cases means and standard deviations were inconsistent. It thus becomes very difficult to interpret the body of evidence based on a literature that has such a large number of errors. Overall, however, the literature was sufficiently powered to detect medium-size effects. The improvement of computational reproducibility is crucial here. We recommend doing so by building in transparency from the outset, by utilizing the OSF in one's workflow and by utilizing a set "research workflow" (see e.g., Silan et al., 2021).

Interpreting seemingly contradictory results

At the level within our meta-analysis, there were seemingly contradictory results between different analysis methods. For the being-in-nature literature, our main inference engine, the 4PSM, converged with a technique that has become more popular as of recent, the *p*-curve. For the emotional-social-support literature, this was not the case. How could this be? A first part of the explanation rests on the number of effects that went into the analysis models: The number of effects that entered into the *p*-curve analysis was small (k = 4 and k = 6 for being in nature and emotional social support, respectively), whereas for the 4PSM this was slightly larger (k = 16 and k = 13 for being in nature and emotional social support, respectively). The second part of the explanation rests on the fact that the 4PSM has a more acceptable false-positive rate than the *p*-curve in the presence of significant heterogeneity (Carter et al., 2019, Hong et al., 2020). Overall, we thus remain with our conclusion that there is sufficient evidential value for the effect of being in nature on stress, but not for emotional social support on stress, due to our greater confidence in the 4PSM as our inference engine.

At the level between our meta-analysis and conclusions from other reviews, there were also substantial differences. A priori, we had suspected that emotional social support would provide sufficient evidence, as the literature seemed to be reasonably powered. Nevertheless, despite a large body of empirical support (e.g., Cohen & Wills 1985; Cohen, 2004; Lakey & Cronin 2008) and a reasonably powered literature, we failed to find any evidence for the effect of emotional social support on stress. Perhaps emotional social support does lower stress, but as is clear from our analyses, the literature 1) is at too high risk of bias, 2) has too many statistical errors, 3) is too heterogeneous, and 4) contains too much publication bias to make a reasonable effect size estimate.

Limitations of our assessment: Constraints on generality (Simons et al., 2017)

The limitations of our assessment primarily apply to being in nature, as that is where we found an overall effect, even after applying our publication bias correction techniques. While there was an overall effect, there was a slight effect of gender. It is hard to infer whether this indeed will extend to other populations, given that sampling is hardly ever representative. Our recommendation is therefore to control for gender in future studies. In addition, effects of being in nature were quite consistent across healthy non-student populations and healthy student populations. Studies in this literature also did not vary across natural environments, being present in a green environment, viewing nature through a virtual medium, or a mixed condition. Effect sizes were similar across these different conditions and thus seemed to be quite generalizable (we don't know the effects for clinical populations, however).

Further, the average age of the studies for being in nature was 29.54 with an average deviation of 7.75, thus capturing a decent segment of the population. Nevertheless, it is unclear whether the effects will hold for minors and people above 40. The effects were studied in quite a wide variety of countries (3 in the United States, 1 in Poland, 1 in Malaysia, 4 in the United Kingdom, 1 in Japan, 1 in Finland, 1 in Germany, 1 in China, 1 in the Netherlands, 1 in South Korea, 1 in China, 1 in Italy, and 1 in Denmark) thus allowing for a reasonable generalizability across a few countries. However, all but two of the countries are defined by the OECD as higher income countries, and even the two on the list of lower- and middle-income countries where studies (China and Malaysia) were conducted are in the upper middle income countries. Whether the effects of being in nature on stress are even relevant for people in lower middle-income countries is thus unknown. Finally, we had little to no information about personality characteristics and can therefore not make a reasonable assessment on whether the effects of being in nature differ across different people.

Conclusions

Our Registered Report investigation, after applying publication bias techniques and driven by our inferential criteria, found that being in nature is an effective strategy to reduce stress. For emotional social support, we do not find the intervention to be efficacious. While the results for being in nature may be promising, the limited quality of the literature poses a potential threat to the validity of the findings. More rigorous studies on the topic - and thus the adoption of Registered Reports, pre-registration, and data sharing - will lead to less research waste, and ultimately, to better interventions.

Data availability

Underlying data

Data from this article can be found on the OSF page (https://osf.io/6wpav/).

Extended data

To ensure methodological rigor and transparency, we have made our data and the script available on the Open Science Framework (https://osf.io/6wpav/)

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0). At least some of the data/evidence that will be used to answer the research question already exists AND is accessible in principle to the authors (e.g., residing in a public dataset or with a colleague). The authors used the data to create a coding scheme BUT the authors certify that they have not yet accessed any part of summary statistics.

Contributors

Conceptualization: Alessandro Sparacio, Ivan Ropovik, Gabriela Jiga-Boy, and Hans IJzerman.

Data curation: Alessandro Sparacio and Ivan Ropovik.

Formal analysis: Alessandro Sparacio, Ivan Ropovik, and Hans IJzerman.

Funding acquisition: Hans IJzerman and Gabriela Jiga-Boy.

Investigation: Alessandro Sparacio.

Methodology: Alessandro Sparacio, Ivan Ropovik, and Hans IJzerman.

Project administration: Alessandro Sparacio, Ivan Ropovik, Gabriela Jiga-Boy, and Hans IJzerman.

Resources: Alessandro Sparacio, Ivan Ropovik, Gabriela Jiga-Boy, and Hans IJzerman.

Software: Alessandro Sparacio and Ivan Ropovik.

Supervision: Ivan Ropovik, Gabriela Jiga-Boy, and Hans IJzerman.

Validation: Ivan Ropovik, Gabriela Jiga-Boy, and Hans IJzerman.

Visualization: Alessandro Sparacio, Adar Cem Lağap and Ivan Ropovik.

Writing - original draft: Alessandro Sparacio, Ivan Ropovik, and Hans IJzerman.

Writing - review & editing: Alessandro Sparacio, Ivan Ropovik, Gabriela Jiga-Boy, and Hans IJzerman

Preface of Study II

The Registered Report meta-analysis found naive meta-analytic effects for being in nature and emotional social support. However, after applying publication bias correction techniques, the effect held only for being in nature. In the following chapter we investigated the efficacy of two other internal regulation strategies (i.e., self-administered mindfulness and biofeedback) in a pre-registered meta-analysis. The structure of this meta-analysis mimicked the previous one: We investigated the same outcome (i.e., the stress response and the affective consequences of stress) and we used a very similar analysis workflow (e.g., 4PSM, *p*-curve, PET-PEESE). However, this meta-analysis was conducted chronologically before the Registered Report, for this reason it may seem that some of the analyses techniques we used are obsolete. We decided to place it as a central chapter, because the meta-analysis on self-administered mindfulness ties into the multi-site study that constitutes the final chapter of this Ph.D.

Chapters 3: Stress regulation via self-administered mindfulness and biofeedback interventions in adults: A pre-registered meta-analysis

Author Note: Our OSF page can be found here (http://bit.ly/39MifmH). The authors that contributed to this manuscript are: Alessandro Sparacio, Ivan Ropovik, Gabriela M. Jiga-Boy, Patrick S. Forscher & Hans IJzerman

Abstract

We conducted a pre-registered meta-analysis to appraise available evidence on two stress regulation strategies: Self-administered mindfulness meditation and heart rate variability biofeedback. We used a combination of keywords to find as many experimental and observational studies as possible, all of which highlighted a link between the two strategies and different components of stress (physiological, affective, and cognitive) and affective consequences of stress. We found 35 effects for self-administered mindfulness reported in 14 papers (total N= 1141, Mdn_N= 67) and 31 effects for biofeedback reported in 13 papers (total N= 468, Mdn_N= 27). We found no evidence for the efficacy of self-administered mindfulness (Hedges's g = .25, p = .21) and biofeedback (Hedges's g = .40, p = .15). The lack of an effect was mostly due to the low quality of the literature, high heterogeneity, high risk of bias, and the lack of Registered Reports. We recommend against relying on these interventions prior to obtaining a more reliable evidence base.

Keywords: Stress regulation, self-administered mindfulness, biofeedback, meta-analysis, publication bias

Chapter 3: Stress regulation via self-administered mindfulness and biofeedback interventions in adults: A pre-registered meta-analysis

Stress can influence how we experience emotions, how our body functions, and how we think (Aldwin 2007; Lazarus & Folkman, 1984). When stress gets excessive, it can even compromise people's everyday functioning (Jiaxuan, Jiali, Yuanyuan, & Wei, 2018). In cases when excessive stress lasts for extended periods of time, it can become an important driver in developing depression and/or anxiety disorders (Yang et al., 2015). Thus, knowing how to regulate stress effectively is vital. But it is not always desirable to utilize pharmacological interventions, as these may come with side effects. Further, pharmacological interventions may simply not be the most cost-effective way to deal with stress. It is thus essential to develop cost-effective (and non-pharmacological) interventions in order to cope with stress (q.v., de Witte, Spruit, van Hooren, Moonen, & Stams, 2019).

Two very popular and non-pharmacological ways to regulate stress are self-administered mindfulness (a type of meditation that does not require an instructor) and "biofeedback" (a self-regulation training based on feedback on physiological mechanisms). Here, we plan to synthesize the evidence on these two strategies via a meta-analysis. Our primary reason to choose these interventions are because they are non-invasive and do not require the presence of other people. However, the choice is also partly arbitrary as the synthesis of these strategies will be a first step towards building a more comprehensive database to understand various stress-regulation strategies and their efficacy.

Overall, our motivation to conduct this meta-analysis is to assess the state-of-the-art in stress-regulation research and to provide directions on how to improve stress-regulation research moving forward. To accomplish this goal, we reviewed the existing literature and addressed some important questions: For which components (physiological, emotional, cognitive) underpinning biofeedback and self-administered mindfulness is there adequate empirical support? In addition, are individual differences taken into account when it comes to the efficacy of different stress regulation intervention? Is it possible to identify for whom certain strategies work and for whom they don't? We intend to shed light on the mechanisms underpinning stress regulation by employing a workflow incorporating various publication bias-detection techniques. To be as inclusive as possible, we also included in our meta-analysis affective states that are consequences of stress.

Stress regulation

Stress is generally understood as a non-specific response of the body, which occurs when external demands exceed internal resources (Lazarus & Folkman, 1984; Selye, 1956). The response to stress can be thought of as separated in three different components: Affective (Watson & Clark, 1988), physiological (Schneiderman, Ironson, & Siegel, 2005) and cognitive (Du, Huang, An, & Xu, 2018). We use these components as tools for our meta-analysis; we nevertheless agree with the theoretical position that these different types of response do not really present conceptual distinctions (Pessoa, 2008; Phelps, 2006) and that the systems underlying these three components influence each other during stress (De Witte, Spruit, van Hooren, Moonen, & Stams, 2019; McEwen & Gianaros, 2010). But yet, splitting them into three different categories allows us to better understand the potential applied value of certain stress-regulation strategies.

The first of the three components, the affective one, is characterized by feelings of nervousness, strain, and tensions that arise when individuals are overrun by external demands (Ratanasiripong, Ratanasiripong, & Kathalae, 2012). The second, the physiological component, is characterized by an activation of the hypothalamic–pituitary–adrenal axis (HPA axis; Stephens & Wand, 2012). Physiological responses include, but are not limited to, the activation of the autonomic nervous system, which can be assessed through changes in heart rate, heart rate variability, systolic and diastolic blood pressure, skin conductance, and cortisol (Bally, Campbell, Chesnick, & Tranmer, 2003). At the level of the cognitive response, stress has been found to be associated with changes in cognitive functions (which include processes like reflection and rumination; McFarland, Buehler, Von Rüti, Nguyen, & Alvaro, 2007). For the present meta-analysis, we decided to focus on measurements of perseverative thinking and rumination.

The lack of a true separation between these components also means that stress does not stand on its own; when stress exceeds what one can handle, longer-term affective consequences may emerge, like depression or chronic anxiety (Cohen, Kamarck, & Mermelstein, 1983). Regulation strategies for stress oftentimes include ways to shield oneself from such longer-term consequences of stress. In his model, Russel (1980) classifies affect (and longer-term affective consequences) into valence (positive vs. negative) and arousal (high vs. low). The former is related to the degree of pleasantness of the affective experience while the latter is related to the level of arousal of the affective experience (Feldman, 1995; Russel, 1980). Crossing these two foci lets us categorize the majority of affective experiences.

To downregulate stress people may rely on strategies, like self-administered mindfulness meditation and heart rate variability biofeedback. Self-administered mindfulness shares features, such as a non-judgmental attitude and an acceptance of inner experience, with other mindfulness protocols. In contrast with other protocols, however, self-administered mindfulness does not require the presence of an instructor, is available 24/7 to people, and tends to be one of the least costly ones (Spijkerman, Pots, & Bohlmeijer, 2016). Self-administered mindfulness can be administered via smartphone applications, audio files, and books which can guide the user through self-administered mindfulness exercises.

In the empirical literature, a two-week self-administered mindfulness meditation intervention (compared to a passive control group), has been found to influence the affective component by reducing self-reported stress (Cavanagh et al., 2013; Cavanagh et al., 2018). We have found no studies on self-administered mindfulness that address the physiological component (probably due to the fact that self-administered protocols are mostly administered online). For traditional mindfulness interventions, Sanada et al. (2016) found slightly lower cortisol levels after intervention. For what concerns the cognitive component, a brief protocol of self-administered mindfulness compared to a waitlist control, reduced maladaptive cognitions like perseverative thinking (Cavanagh et al., 2018). Finally, a daily guided self-administered mindfulness intervention via an audio CD (compared to a passive control group) has been found to lead to a significant reduction in depression (May Barry et al., 2018).

The other strategy, biofeedback, is a technique that allows people to have immediate feedback on a specific physiological function (such as heart rate) that is under the control of the autonomic nervous system. Biofeedback interventions are thought to recruit the parasympathetic branch of the nervous system, thereby inhibiting sympathetic activity (Frank et al., 2010). Based on the biofeedback received, people can learn how to change a behavior (e.g., their breathing rate) to improve the particular function they are monitoring. Biofeedback is typically delivered via computers or smartphones and is thought to improve self-regulatory capacities (Gross, 1998). For the present meta-analysis, we have decided to focus on one type of biofeedback called heart rate variability biofeedback. Heart rate variability biofeedback is, after all, the most used and thought to be the most efficacious in reducing stress (Lehrer & Gevirtz, 2014).

In one study, female nursing students participated in a 5-week intervention relying on three biofeedback sessions per day. The experimental (as compared to a passive control) group showed a significant reduction in self-reported levels of stress (Ratanasiripong, Ratanasiripong, & Kathalae, 2012). In another study, evidence was more mixed: Participants in a biofeedback (compared to an active control group) condition showed improvements in respiratory rate, but no changes in blood pressure were found (Prinsloo et al., 2011). For what concerns the cognitive component, several authors found a negative association between heart rate variability and perseverative thinking (Eddie et al., 2014; Thayer & Friedman, 2002). Finally, in one study, six sessions of heart rate variability biofeedback (compared with an active control group) interventions improved symptoms of depression at post testing and at 1 month follow-up (Lin et al., 2019).

Bias in estimating the efficacy of stress regulation

Nearly all of science, including psychology, is affected by publication bias (the likelihood that positive results have a higher probability of getting published; Rosenthal, 1979) and data contingent analyses (Gelman & Loken, 2013). Fanelli (2010), for example, estimated that psychology's and psychiatry's published findings contain over 90% significant results. Such a high positive ratio of findings is statistically highly unlikely, as the concerned literature is underpowered (Maxwell, 2004). The psychological literature thus likely contains a large number of false-positive and overestimated effects, as also hinted at by the growing body of replication studies in psychological science. When researchers repeat earlier found results with the same procedures, they often fail to find the same results (Klein et al. 2018; Maxwell, Lau, & Howard, 2015; Open Science Collaboration, 2015).

It is reasonable to assume that some proportion of the literature on stress – like any other literature in psychology – suffers from replicability issues. To be able to provide people with advice on appropriate stress-regulation strategies, a meta-analytic assessment is thus needed to assess the extent of how empirically robust the published findings are (cf., IJzerman et al., 2021). Past meta-analyses on self-administered mindfulness interventions (Cavanagh, et al., 2014; Taylor et al., 2021) and biofeedback (Goessl, Curtiss, & Hofmann, 2017) have tried to estimate the severity of publication bias in the literature. All three meta-analyses have found that biofeedback and self-administered mindfulness are moderately effective in reducing stress, depression, and anxiety. However, neither meta-analysis has dealt with publication bias adequately. Taylor et al. (2021) and Cavanagh et al. (2014), for example, assessed publication bias with funnel plots and fail-safe *N*, without carrying out formal adjustment. The former technique is a mere descriptive visualization requiring subjective judgment. The latter is also not a formal detection or adjustment method, resting on problematic assumptions, and is now widely considered to be outdated (see Ferguson & Heene, 2012). Goessl et al. (2017) instead addressed publication bias using the trim-and-fill, which itself too is known to be problematic. Trim-and-fill is known to have an excessive false-positive rate under even the most realistic conditions (Carter et al., 2019; Hong et al., 2020).

To deal with publication bias, we employed several state-of-the-art publication bias correction methods, assuming a more realistic data-generating process behind the published effects of stress-regulation strategies. In doing so, we followed the workflow of IJzerman et al. (2021), assessing the evidential value of the literature using a permutation-based *p*-curve analysis and by estimating a naive meta-analytic effect size (and its heterogeneity) using a RVE-based multi-level model, assuming correlated and hierarchical effects (CHE working model; Pustejovsky & Tipton, 2021).

Then, we employed a tandem procedure involving the 4-parameter selection model (McShane, Böckenholt, & Hansen, 2016) and the regression-based PET-PEESE method (Stanley & Doucouliagos, 2014) to try to correct for publication bias. We considered an effect to be present only if all the meta-analytic techniques detected one. This approach was highly conservative, but intentionally so to mitigate the risk that publication bias is a sole explanation of the target effect.

Research overview

To appraise the available evidence on the effects of self-administered mindfulness and biofeedback on stress in a more detailed manner, we have conducted a meta-analysis with the following objectives: 1) to assess the evidential value of identified studies in both literatures, 2) for either regulation strategy, to estimate mean effect sizes for the three components of stress (affective, physiological, cognitive), 3) for either regulation strategy, to estimate the mean effect sizes for the affective consequences of stress, 4) to adjust the target estimates for publication bias using various techniques, and 5) to determine whether personality traits were taken into account in these literatures.

Method

The present meta-analysis aimed to be reproducible and fully transparent. In order to accomplish this goal, we made analysis scripts and the data for this study publicly available on our project page on the OSF (http://bit.ly/39MifmH). We intended to map the methodological quality and possible biases in this field, carrying out several sensitivity analyses with respect to our main results, and controlling for prognostic methodological factors related to the magnitude and precision of the reported effects in subgroup analyses.

Our meta-analysis was pre-registered on PROSPERO with protocol number [CRD42020179810] and on the OSF (http://bit.ly/39MifmH). Any changes to the pre-registration are fully disclosed on our OSF page (using the template provided by Moreau and Gamble 2020; Appendix I. All appendices can be found in the material section on the OSF page http://bit.ly/39MifmH.). We intend to build a cumulative scientific knowledge base by depositing the data of our meta-analysis to PsychOpen CAMA, a platform whose aim is to construct a community-augmented meta-analysis (Burgard et al., 2021; Tsuji, Bergmann, & Cristia, 2014).

Inclusion criteria, and search strategy

To frame our research question in a structured way, we followed the Participants, Intervention, Comparator, and Outcome (PICO) Framework (Schardt et al., 2007). We chose to only include studies on participants that are adults (people aged 18 years or older). We only included two interventions (self-administered mindfulness and biofeedback based on heart rate variability). For designs comparing groups, we included effects based on a comparison to an active control condition, meaning that participants engage in some tasks during the intervention period or to a passive control condition (participants are in an untreated comparison group; e.g., waitlist control). If there was more than one comparator in the same study (i.e., presence of both an active and a passive control group), we chose the contrast with the active control group. The main outcome of this meta-analysis was the level of stress after intervention, divided into three components (affect, cognition, and physiology). We relied on self-report measures (for the affective and the cognitive component and for the longer-term affective consequences of stress) taken at post-test (immediate and/or delayed) of the two groups (experimental and control). For the physiological component, we relied on changes in the activation of the hypothalamic-pituitary adrenal axis (e.g., via cortisol levels or heart rate) taken at post-test (immediate and/or delayed) of the two groups (experimental and control).

After framing our search strategy, we conducted a preliminary search to pilot the planned strategy and coding scheme (on a dozen eligible articles randomly chosen from each strategy category). To guarantee the reproducibility of the literature search, we followed the recommendations by Maggio et al. (2011), who provided guidelines to specify 1) which databases we used 2) which search terms we used, 3) which Medical Subject Headings (MeSH) or thesaurus terms we used, 4) on which dates we searched, and 5) what our search limits were. We searched the literature using the following sources: ProQuest, (an online platform which included research coming from APA PsycArticles, APA Psycinfo, ProQuest Dissertations & Theses Global), PubMED, and Scopus.

The first author (AS) conducted the literature search and excluded articles which were not relevant to the aim of the present meta-analysis. Screening by title and abstract was carried out using Rayyan QCRI (Ouzzani, Hammady, Fedorowicz, & Elmagarmid, 2016). He then manually searched reference lists for more studies and unpublished reports. Finally, we asked for unpublished results via social networks (Facebook Groups and Twitter), via relevant mailing lists (SPSP, EASP, ESCAN), and via listservs of an association of clinical and health psychology (EACLIPT). Furthermore, to have a wider coverage of the literature, we contacted authors that have published studies on the topic to see if they have any unpublished research, in-progress manuscripts, or in-press manuscripts (see our templates in Appendices L and M). We also included articles from previous meta-analyses that we may have missed with other search methods (the articles that we obtained from these meta-analyses are included in Appendix M).

The inclusion criteria for our meta-analysis were as follows: 1) We included published primary studies, preprint articles, working papers, dissertations, and books (we included only studies published in English), 2) we included both randomized control trials and observational studies that estimated the effect of (or exposure to) one of the two stress regulation strategies, 3) we included studies that measured at least one of the three components of the stress response, which we included in the meta-analysis (physiological, affective, and cognitive), or at least that measured the affective consequences of stress, and 4) the participants of the study had to be humans. Studies were excluded if 1) the paper was a review or a meta-analysis, 2) participants were adolescents (under 18 years of age), 3) a common-metric effect size reflecting a bivariate effect or relationship could not be computed, or 4) the two types of intervention included in our meta-analysis (self-administered mindfulness & biofeedback based on heart rate variability) were combined with other active treatment (e.g., cognitive behavioral therapy or physical exercise).

We then added sub-exclusion criteria related to the two stress-regulation strategies. For self-administered mindfulness, we excluded studies 1) which involved the presence of an instructor, 2) which cannot be self-administered, and 3) which require meditation practices in groups with other people. For each study on self-administered mindfulness, we coded the source from which participants got the instructions for doing the mindfulness exercises (e.g., an audio source or via a website). For biofeedback, we excluded from the meta-analysis any kind of biofeedback that was not based on heart rate variability. A PRISMA flow chart of the overall literature search and inclusion procedure is shown in Appendices N and O.

Coding and data preparation

The coding was carried out by two independent coders, with the first author coding all the included literature and the second coder coding at least 20% of the data for cross-checking. The coding process was cross-checked after the first 10% and after the second 10%. To evaluate the inter-rater reliability, Cohen's Kappa and inter-rater agreement were calculated. We double-coded a subset of the data, including the part of the statistics and characteristics of the individual studies. Following the guidelines of Landis and Koch (1977), we considered an agreement of $\kappa > 0.60$ for metric or multinomial variables acceptable. Consensus between two coders was reached through discussion in case of coding disagreements.

We extracted data for the following variables: Publication year, the number of citations of the paper, country, journal name, reported overall N, gender ratio, publication status, reported effect sizes, total *N*, cell means, standard deviations and *N*s, test statistic, degrees of freedom, the type of effect (e.g., main omnibus, main bivariate, or interaction), the reported *p*-value, the design, the type of population, the nationality of participants, whether individual differences were taken into account or not, the name of stress test used in experiments, the duration of the intervention, the type of intervention, the category of stress-regulation strategy (self-administered mindfulness or biofeedback), the type of control group (active or passive), whether it was one of the components of stress (affective, cognitive, or physiological) or an affective consequences of stress, the number of measurements, the instrument employed to assess stress levels, the number, and the duration and the frequency of interventions. We also coded whether the effect was "focal" from the perspective of the author (if mentioned in the abstract, we coded it as focal).

We converted all the relevant bivariate effect sizes to a common metric (Hedges' *g*), which is a standardized mean difference corrected for small sample bias (Hedges & Olkin, 1985). We also extracted the reported degrees of freedom as a more accurate estimate of the sample size because dropping of participants may have not been properly disclosed. For designs with two or more experimental groups, we checked the accuracy of the sample size in two ways: First, we looked at the reported *N*s of each group and computed the total *N*. If the sum matched the overall sample size +/-2, the reported group cell sizes were used. If that was not the case, the cell *Ns* were computed from the degrees of freedom, assuming a balanced design. In the case that the authors provided information only for the total sample size, we also assumed equal cell sizes. For within-participants designs, a conservative correlation between the measurements of $r = .50$ was assumed. Lastly, we conducted a sensitivity analysis varying the assumed correlation from .10 to .90, in steps of .20 in order to determine the impact on the overall effect size estimate.

Analyses

Analysis strategy

Following the workflow of IJzerman et al. (2022) we utilized a multilevel random-effects meta-analysis model using the restricted maximum-likelihood estimation with the R package *metafor*, version 2.0 (Viechtbauer, 2010). We included all the important dependent variables in our models even if some of them came from a single study. To deal with dependencies in the data, we employed a robust sandwich-type variance estimation (RVE) with the CHE working model (Correlated and Hierarchical Effects; Pustejovsky & Tipton, 2021). The aim was to simultaneously account for both types of dependencies in the data, nesting of effects within studies and clustering due to the fact that some effects were estimated on the same participants. Here, we assumed a constant sampling correlation of .5, since sampling correlations among the effects are usually not reported. As a sensitivity analysis, we relaxed this assumption by varying the sampling correlation from 0 to .6 in steps of .2. We estimated the heterogeneity using τ (*SD* of the distribution of true effects) and I^2 (proportion of total variation in study estimates due to heterogeneity). We did not interpret any estimates when the number of included effects (*k*) were less than 10, because of the large expected sampling variability of such estimates, leading to estimates that are not precise enough to be useful.

In the analysis, we first excluded studies with high risk of bias and effects based on mathematically inconsistent means or standard deviations. We then carried out an in-depth diagnosis of the random-effects meta-analytic model, analyzing especially the presence of influential outliers. In case that there were excessively influential outliers (based on Beaujat plot and influence diagnostic indices), we examined their effect on the overall result in a sensitivity analysis. With this subset of studies, we first tested the overall effect sizes of the two different strategies (self-administered mindfulness and biofeedback) on the components of stress and the affective consequences of stress. Subsequently, we adjusted for publication bias to assess the level of empirical support for the specific intervention. We then proceeded with subgroup analyses to determine whether the intervention efficacy varied as a function of study or intervention characteristics.

For what concerns both strategies, five moderator variables were taken into account: Number of sessions (of self-administered mindfulness or biofeedback), intervention duration, number of females versus males, type of comparison group (active or passive control), and timing of the effect (after the intervention, after last follow-up). Lastly, we included a subgroup analysis related to the type of population (student non-clinical, non-student non-clinical, clinical): If we were to find considerable heterogeneity, we conducted subgroup analyses on these groups to check if it was the source of heterogeneity. Again, for all subgroup analyses, effect sets with less than 10 effects were not analyzed. To test for equality of effect sizes across levels of the moderators, we used the robust HTZ-type Wald test.

Lastly, we omitted observational studies in a sensitivity analysis. A figure of the analysis workflow can be found in Appendix P. If we decided that additional subgroup analyses would be necessary exploratorily, we disclosed them on our OSF page (again, see Appendix I).

Correction for publication bias

Under publication bias, meta-analytic effect size estimates tend to have a high false-positive rate (if H_0 is true) or they end up being overestimated (if H_0 is false; Carter, Schönbrodt, Gervais, & Hilgard, 2019; Ioannidis, 2008). A secondary consequence of publication bias is the fact that researchers might conduct data-contingent analyses (Simmons, Nelson, & Simonsohn, 2011) to obtain *p*-values less than .05.

To try to mitigate these problems and estimate an unbiased effect size of stress regulation strategies, we tried to account for publication bias using a variety of approaches. Using a similar procedure as IJzerman et al. (2021), we first estimated the evidential value with the *p*-curve method, applied on a set of significant results (Simonsohn, Nelson, & Simmons, 2014). Second, we tried to estimate an unbiased average effect using techniques that have only recently become available: The 4-parameter selection model (McShane, Böckenholt, & Hansen, 2016) and a mixed-effects implementation of PET-PEESE (Stanley & Doucouliagos, 2014).

P-curve

P-curve is a technique used to assess the evidential value in a set of significant findings (Simonsohn, Nelson, & Simmons, 2014). According to Simonsohn, Nelson, and Simmons (2014), we can infer the presence of bias by observing the shape of the *p*-curve. Under a null effect ($d = 0$) the distribution of *p*-values is uniform. When an effect is present ($d \neq 0$) the results of that experiment are more likely to be associated with small rather than high *p*-values. The greater the statistical power, the steeper the *p*-curve (leading to a higher degree of right-skew in *p*-values). A *p*-curve with a left-skewed distribution of significant *p*-values may instead suggest the presence of questionable research practices. To handle the dependencies among the effects, we subsetted only the focal effects, randomly selected only a single effect size for each dependent set of effects, permuted this procedure in 5000 iterations, and selected the model with the median z-score for the right-skew of full *p*-distribution.

4-parameter selection model (4PSM)

The 4PSM is an implementation of selection methods, which are techniques that estimate and correct for publication biases regarding the size, direction, and statistical significance of study results (McShane et al., 2016). The 4PSM is a statistical model that has two components: 1) A data model describing how the data are generated when there is no publication bias and 2) a selection model that emulates the publication process. Each of them consists of two parameters. The two data model parameters are: a) Effect size parameter, which models the population average effect size and b) heterogeneity parameter. The selection model is represented by c) a weight parameter, which models the likelihood that a study with non-significant results is published compared to a study with a *p*-value below .05

and d) the likelihood of a result being in the opposite direction. These parameters allow for an estimate of the effect size and degree of heterogeneity in a way that accounts for publication bias (McShane et al., 2016). If a given set of results yielded less than four *p*-values per interval, the model dropped the fourth parameter to provide for a more stable estimation. Just like the *p*-curve, the 4PSM also assumes the included effect sizes to be independent. Therefore, we implemented the same permutation-based procedure, iteratively selecting only independent effect sizes, estimating the models, and averaging over the estimates.

As a sensitivity analysis, we have also applied the Vevea and Woods approach (2005) with a priori defined selection weights. Using a series of fine-grained step function models, this allowed us to examine the results when varying the assumed severity of bias (modeling moderate, severe, and extreme selection). The dependencies among the effects were handled using a permutation approach. The exact specification of these step function models can be seen or easily customized by the user in the accompanying script.

PET-PEESE

PET-PEESE is a conditional regression-based meta-analytic model that aims to correct for publication bias. The difference between PET and PEESE is that in the former, the effect size is regressed on the standard error while in the latter, the effect size is regressed on the squared standard error (variance) instead. A slope of the regression line indicates a relationship between the standard error and the effect size. This pattern hints at the presence of publication bias or other small-study effects. Because the PET model has a greater accuracy when the effect is absent and the PEESE model is more accurate when the effect is present, Stanley and Doucouliagos (2014) suggested using the PET model first. If the estimate of the PET is significantly different from zero, PEESE is used, and its result is taken as the bias-corrected effect size of interest. If PET does not detect any effect, the estimate of the PET is retained instead.

However, in many realistic situations, PET tends to have an unfavorable combination of false-positive rate and statistical power. Here, we used the 4PSM as a conditional estimator for PET-PEESE, as it was shown to have a relatively adequate false-positive rate under most conditions (Carter et al., 2019, Hong et al., 2020). Dependencies among the effects were modeled by employing a RVE-based multi-level model, assuming correlated and hierarchical effects (CHE working model; Pustejovsky & Tipton, 2021). A chart of the techniques employed to account for publication bias can be found in Appendix Q.

Quality assessment

As a first step to assess study quality, we identified whether some study values were mathematically impossible. If in our studies the outcome was a discrete variable (as is the case for Likert scale items), means and standard deviations follow fixed granular pattern for each combination of *N* and number of items (Anaya, 2016; Brown & Heathers, 2016), rendering some values mathematically impossible.

Additionally, we used the Risk of Bias 2 (RoB 2) Tool by the Cochrane Foundation (2020). Using the RoB 2, we assessed the risk of bias corresponding to five predetermined domains: 1) risk of bias arising from the randomization process, 2) risk of bias due to deviations from the intended intervention (effect of assignment to intervention), 3) risk of bias due to missing outcome data, 4) risk of bias in the measurement of the outcome, and 5) risk of bias in the selection of the reported result. Within each domain, a series of signaling questions were answered to determine the relative risk of bias for each domain. Based on the answer to these questions, we used the algorithmic approach to assess the risk of bias corresponding to each domain as “low”, “some concerns”, and “high”. Then, based on the judgment for each individual domain, an overall rating on the risk of bias was determined. A study was categorized as having a Level 3 overall score (“high risk of bias”), if one of two conditions were met: A) the study scores having high risk of bias in at least one of the five domains or B) the study led to some concerns for at least three of the five domains.

Inter-rater agreement

We calculated the inter-rater agreement for the first 10% of the coded articles. For the variables for which Cohen's Kappa was lower than our acceptance threshold (< .60), the coding strategy was respecified. For the first 10% of articles there were 4 variables that had values lower than the acceptance threshold, indicating an underdeveloped coding scheme for these variables. The procedure was repeated after coding the second 10% of the data. The discrepancies in the rest of the data were resolved by discussion.

Results

The final meta-analytic data set consisted of 76 effects reported in 26 papers on self-administered mindfulness and of 43 effects reported in 20 papers on biofeedback. In line with the pre-registration, we applied strict selection criteria for the effects and we excluded studies that scored high in risk of bias[11] assessment and single effects with mathematically inconsistent means or standard deviations. We also screened the data set for outliers using the Baujat and Gosh plot and carried out influence diagnostics. None of the coded effects was found to exert an undue influence on the meta-analytic models. After the exclusions, 35 effects for self-administered mindfulness, reported in 14 papers (total N= 1141, Mdn_N= 67, M_{Age} = 31.06, SD_{Age} = 9.86), and 31 for biofeedback, reported in 13 papers (total N= 468, Mdn_N= 27, M_{Age} = 27.43, SD_{Age} = 9.74), met the inclusion criteria for a meta-analytic synthesis. In what follows, we will address the core questions of our meta-analysis and we will present the set of subgroup analyses that we performed.

[11] We pre-registered that we would have included observational studies at first, and that then, we would have excluded them via a sensitivity analysis. However, we excluded them before running any analyses.

Answering the four core questions of the meta-analysis

In Table 3, we provide a summary of the questions we asked in our pre-registration.

Question	Conclusion
Is there support for the reduction of stress through self-administered mindfulness and biofeedback?	No support.
For which components (physiological, emotional, cognitive) underpinning biofeedback and self-administered mindfulness is there adequate empirical support?	The status of any of these components is unclear. We found some empirical support for the affective component for self-administered mindfulness and for the physiological component for biofeedback, but publication bias adjustment could not be applied.
Are individual differences taken into account in primary studies when it comes to the efficacy of different stress regulation interventions?	Individual differences were not taken into account in the focal analyses as they were only used as exclusion criteria in the included studies.
Can we infer for whom certain strategies work and for whom they do not?	Women seem to benefit more from a biofeedback intervention as compared to men.

Table 3: Core questions of the meta-analysis

Overall, we did not find sufficient evidence for the overall efficacy of self-administered mindfulness and biofeedback. We also did not find evidence for the reduction of stress regulation for the different components of the stress response nor as they pertain to reducing the affective consequences of stress. For the different subgroup analyses we conducted, publication bias adjustments were not applied due to the low number of effects; these results should therefore be interpreted with caution.

Do self-administered mindfulness or biofeedback reduce stress?

We did not find evidence for the effect of self-administered mindfulness on the three components of stress (affective, physiological, cognitive), nor on the affective consequences of stress. Using a random-effects meta-analysis, we found that out of all eligible coded effects (k = 35), 46% were statistically significant. The overall meta-analytic estimate was at g = .37, 95% CI [.19, .56], p < .001, suggesting the presence of an effect[12] of self-administered mindfulness on reducing stress levels. The 95% prediction interval was broad however, with the true effect in a new published study being expected to fall between -0.29 and 1.03. In the examined set of effects, the amount of absolute heterogeneity was significant and substantial, τ = .29. Only I^2 = 59.8% of the total variance across observed effect estimates was systematic (20.0% due to between- and 39.8% due to within-cluster heterogeneity). Forest plot and the contour-enhanced funnel plot are displayed in Figure 6.

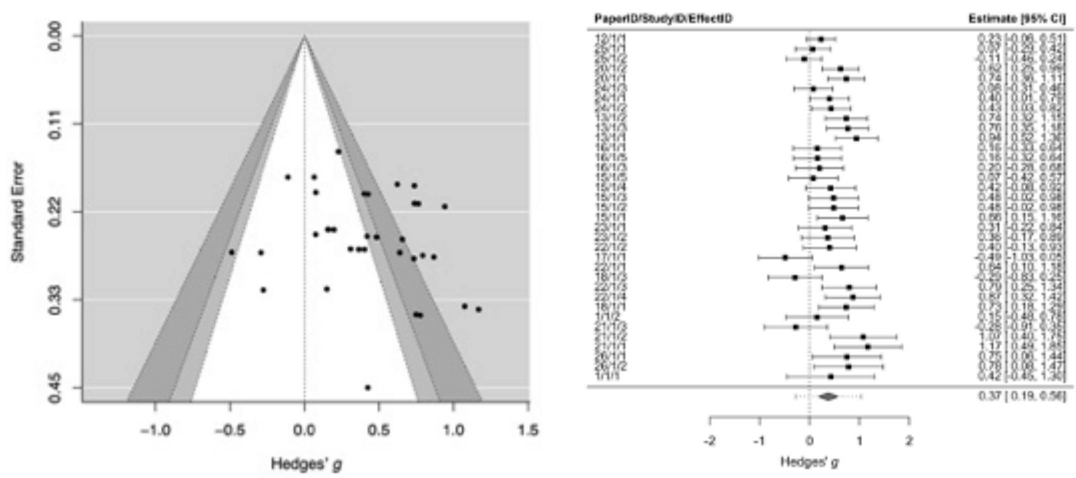


*Figure 6: C*ontour-enhanced funnel and forest plot for self-administered mindfulness.

[12] The effects in Figure 6 represent the positive influence of self-administered mindfulness intervention on reducing stress levels (i.e., higher scores correspond to lower stress levels).

In a second step, the full and half *p*-curve tests ($z = -3.34$ and $z = -3.66$, respectively) were significant (both $p < .001$). Thus, the *p*-curve suggested the presence of evidential value in the given set of included significant effects that were considered focal (however, the permuted model with a median *z*-score was based on only $k = 4$ independent effects). We then applied publication bias-correction techniques. The 4-parameter selection model did not detect a significant effect, with $g = .25$, 95% CI [-.14, .63], $p = .21$. When varying the assumption about the severity of the publication bias using the Vevea and Woods approach (2005), the estimates ranged from $g = .23$, $p = .06$ (for moderate selection) to $g = .02$, $p = .84$ (for extreme selection). PET (and also PEESE) model corroborated the result of the 4-parameter selection model, with PET at $g = .16$. This estimate was relatively imprecise, as is obvious from the 95% CI of [-.65, .97], associated with $p = .68$.

Overall, based on our inferential criteria that all of our analysis techniques should show an effect, we were not able to reject the null hypothesis of zero efficacy of self-administered mindfulness. Although the naive meta-analytic estimate and the *p*-curve suggested the presence of an effect, the 4-parameter selection model and subsequent PET model were not able to detect an effect.

Relatedly, the power of the literature was low. The average power in the set of included studies to detect small effects ($d = .20$) was .21, but was adequate (.82) by conventional criteria to detect effects of $d = .50$ or larger. With regards to the quality and integrity of the underlying evidence, 37% of the coded effect sizes came from designs judged to have low risk of bias, 39% from designs showing some concerns, and 24% originated from designs with high risk of bias. None of the coded effects were based on mathematically inconsistent means or standard deviations.

For biofeedback, we also did not find robust evidence as the techniques employed did not all converge on the presence of an effect. We first calculated the overall meta-analytic effect for biofeedback. Among the biofeedback effects reported in the literature, 35% were statistically significant. The overall meta-analytic estimate was at $g = .59$, 95% CI [.32, .86], $p < .001$. The 95% prediction interval was wide, with the true effect in a new published study being expected in the range from -0.32 to 1.51. We found high heterogeneity, $\tau = .40$, $I^2 = 61.66\%$ (18.9% due to between and 42.8% due to within-cluster heterogeneity). The forest plot and contour-enhanced funnel plot for this set of effects can be seen in Figure 7.

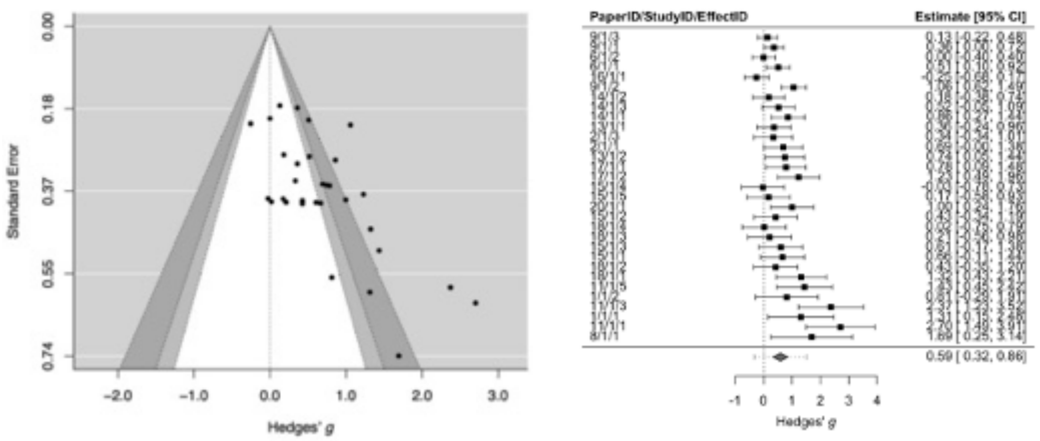


*Figure 7: C*ontour-enhanced funnel and forest plot for biofeedback.

P-values for both, the full *p*-curve ($z = -2.19$, $p = .01$) and the half *p*-curve tests ($z = -1.40$, $p = .08$) were below $\alpha = .10$[13], indicating the presence of evidential value for the *p*-curve in the given set of included significant focal effects (again, based only on $k = 4$ independent effects for the median model). For what concerns the correction for publication bias, the 4-parameter selection model was not significant $g = .40$, 95% CI [-.14, .95], $p = .15$. The estimates proved to be sensitive to the degree of assumed severity of publication bias, ranging from a significant $g = .47$, $p = .01$ for moderate selection, to $g = .18$, $p = .26$ for extreme selection. Subsequent PET (and also even PEESE) estimate was in the opposite direction $g = -.72$, 95% CI [-1.47, 0.04], $p = .061$, which can be regarded nil.

[13] If the half p-curve test is right-skewed with p < .05 or both the half and full test are right-skewed with p < .1, then p-curve analysis indicates the presence of evidential value (Simonsohn, Nelson, & Simmons 2014).

The average power across the literature to detect at least small effects (*d* = .20) and medium-sized effects (*d* = .50) was only .11 and .44, respectively. Lastly, with respect to the quality of evidence and reporting inconsistencies, 15% of the coded effects came from studies with low risk of bias, 76% from study designs having some concerns, and 9% of the effects came from studies with high risk of bias. A sizable proportion (26%, *k* = 5) of the coded effects were based on means or standard deviations that were inconsistent with the reported cell sizes.

Thus, similarly, to self-administered mindfulness, the present data did not provide evidence for the presence of an empirically robust effect. To formally compare the relative efficacy of the two strategies, we used a meta-regression model controlling for design-related factors that might be prognostic with respect to the unadjusted effect sizes (i.e., might vary between the two strategies). As these covariates, we included the study design, type of population, type of comparison group, published status, and the assessment of the risk of bias. Using a robust Wald's-type test, we did not find a significant difference in the efficacy of the two strategies, $F(1, 7.31) = 1.91$, $p = .21$. Nor did we find an effect in a model without covariates ($p = .24$). Table 4 summarizes the results for the overall effect of self-administered mindfulness and biofeedback.

	k	*g* [95% CI]	*SE*	τ	I^2	4PSM estimate	4PSM *p*-value	PET-PEESE estimate	PET-PEESE *p*-value
Self-administered mindfulness	35	.37 [.19, .56]	.08	.29	59.8%	.25 [-.14, .63]	.21	.16 [-.65, .97]	.68
Biofeedback	31	.59 [.32, .86]	.12	.40	61.66%	.40 [-.14, .95]	.15	-.72 [-1.47, 0.04]	.061

Table 4: Meta-analysis for self-administered mindfulness and biofeedback. Values in brackets represent 95% *CI*.

For which components is there adequate empirical support?

In order to establish the presence of adequate empirical support, each meta-analytic technique should have converged on the presence of an effect. However, it was not possible to answer this question exhaustively due to an insufficient number of studies which made it infeasible to employ publication bias adjustment techniques. Thus, the empirical support that we retrieved should be considered suggestive, not conclusive. For self-administered mindfulness, we categorized most of the effects as falling in the affective component (k = 26), We classified many fewer effects as falling into the physiological (k = 4, g = .09) or cognitive component (k = 2, g = .23). Finally, we classified k = 3 effects as being part of affective consequences of stress. The affective component was associated with a large effect size g = .39, 95% CI [.21, .58] (it did not make sense to analyze the others due to a low k). The differences between the components in terms of effect size was not statistically significant, Wald's test p = .18.

For biofeedback, we categorized the majority of the effects as falling in the physiological component (k = 11), followed by the affective component (k = 10) and the cognitive component (k = 1). Finally, we classified k = 7 effects as being part of the affective consequences of stress. The average effect size for the physiological component was large, with g = .51, 95% CI [.19, .82] and similar in size to the affective component with g = .61, 95% CI [.23, .99]. The difference between the components in terms of effect size was not statistically significant, Wald's test p = .78.

Are individual differences taken into account?

For self-administered mindfulness, 19 studies out of 26 reported some assessments of individual differences. For biofeedback, there were 17 studies that reported some assessments of individual differences (out of 20). In most of the cases, these assessments underlay exclusion criteria. For example, for self-administered mindfulness, participants were excluded if they had previous meditation experience (e.g., if they meditated 6 months before the experiment) or if they suffered from certain psychological symptoms (e.g., exceeding a certain threshold for trait anxiety).

For biofeedback, in some cases, participants were excluded if they did not meet certain criteria that made them suitable for that clinical trial (e.g., reporting subclinical psychotic symptoms over a certain threshold). However, for either strategy, all measurements of individual differences were used as exclusion criteria or as baseline measurements. In the included studies, individual differences were never taken into account in any of the analyses. Thus, it was not possible to make any kind of inference as to which strategy is most effective in reducing stress levels based on individual characteristics.

For whom do certain strategies work and for whom do they not?

Although we had little information to understand for whom these strategies work and for whom they don't, we can however indirectly examine whether the two strategies had a different efficacy, throughout all stress components, based on characteristics of the population. In line with the pre-registration, we performed subgroup analyses as we found considerable heterogeneity for both strategies.

For self-administered mindfulness, the mean proportion of female participants across the included studies was 66.5%. We did not find an effect of gender on the efficacy of self-administered mindfulness interventions in reducing stress (p = .61). Next, we conducted a subgroup analysis to examine whether students' samples yielded larger effect sizes as compared to other groups; the effect size coming from a population of students (k = 12), was not larger when compared to non-student healthy participants (k = 14), or to a clinical population (k = 9; Wald's test p = .87).

For biofeedback, most of the participants were also women (67.6%). However, in this case, stress reduction was more pronounced after a biofeedback intervention in women than in men (B = .008; p = .002). Next, we again conducted a subgroup analysis to examine whether the efficacy of biofeedback interventions varied as a function of the type of population. For this strategy, the effects were mostly coming from a non-student healthy participants (*k* = 13) and non-student (*k* = 15) clinical population. The difference between groups was not significant (Wald z-test $p = .72$). Based on the characteristics of the population, the only indication we can draw is that women may benefit more than men from biofeedback intervention.

Methodological moderators

In the next set of moderator analyses, we examined which characteristics of interventions and methodological characteristics of included studies were most effective in terms of stress reduction, taking together the full set of studies per domain.

Characteristics of the intervention

First, we tested whether the efficacy of self-administered mindfulness and biofeedback varied as a function of the number, frequency, and the duration of the interventions. The number of intervention sessions in the included studies ranged from 1 to 64 ($M = 29.3$; $Mdn = 27$). There was no evidence for the effect of the total number of intervention sessions ($b = .01$; $t = 1.93$; $p = .053$). However, the intensity/frequency of intervention (i.e., the number of sessions per week until the end of treatment) was significantly associated with a reduction in levels of stress ($b = .07$; $t = 4.7$; $p < .001$). Similarly, spending a greater number of hours engaging in self-administered mindfulness was significantly related to the reduction of stress ($b = .04$; $t = 4.2$; $p < .001$).

Next, we examined whether there was a moderation by the source of the mindfulness intervention. Interventions can namely be delivered via web-based platforms, smartphone applications, books, audio files, or a mix of these sources. The most frequent sources were: Web-based platforms ($k = 12$, $g = .40$) and books ($k = 9$, $g = .54$). The difference between sources was, however, not significant (Wald z-test $p = .77$).

The number of interventions in the included studies for biofeedback ranged from 1 to 40 interventions (*M* = 9.5, *Mdn* = 6*)*. The effect for the number of intervention sessions was not significant (*b* =.03, *t* = 1.02, *p* = .31*)*, just as the overall duration of the biofeedback interventions (*b* =.04, *t* = 1.28, *p* = .2). Having a greater number of biofeedback interventions per week was however associated with a significant reduction in stress (*b* = .15, *t* = 3.00, *p* = .003). Again, all the aforementioned analyses should be interpreted as only suggestive, given the low number of *k*.

Study design characteristics

We studied how several design characteristics of the included studies related to the effect sizes of those studies. First, we checked whether effects of the studied interventions were smaller when passive control groups were utilized compared to when active control groups were used. We did not detect such an effect for either of the intervention strategies, Wald's test *p* = .35 and .50 for self-administered mindfulness and biofeedback, respectively.

Second, we tested whether the effects of self-administered mindfulness interventions persisted over time after the end of the treatment (by looking at follow-up measurements when reported). There was no difference between measures immediately after the end of the intervention and follow-up measurements, with *p* = .52 and *p* = .54 for self-administered mindfulness (*k* = 35 post-intervention measurements, *k* = 28 follow-up measurements) and biofeedback (*k* = 31 post-intervention measurements, *k* = 4 follow-up measurements), respectively. Note however that only very few studies were eligible for the analysis for biofeedback[14].

Including lower-quality studies

[14] Detailed results for all the above-described analyses and for those that were not pre-registered can be found in the material section on the OSF.

As the primary analysis, we aimed to investigate whether self-administered mindfulness and biofeedback reduce stress. We conducted our analyses by excluding effects based on designs with a high risk of bias assessment and those based on inconsistent means or standard deviations). We did not pre-register this analysis, however, as the scope of a meta-analysis is a comprehensive assessment of the evidence on a given phenomenon, we also carried out a synthesis involving *all* theoretically relevant effects regardless of methodological quality or issues of reporting integrity. The full set of the coded effects included 119 effects (72 for self-administered mindfulness and 42 for biofeedback).

When analyzing all the coded effects (k = 72) related to the efficacy of self-administered mindfulness interventions, the overall meta-analytic estimate was g = .31, 95% CI [.10, .51] with the true effect in a new study being expected to fall between -0.66 and 1.27. In this full set of studies, the amount of heterogeneity was even more substantial than in our primary analysis, τ = .45, I^2 = 79.4% (18.3% due to between- and 61.1% due to within-cluster heterogeneity).

After adjusting for publication bias, the 4-parameter selection model still yielded a non-significant result, g = .21, 95% CI [-.22, .65], p = .34. The PET model again did not find an underlying effect for self-administered mindfulness g = -.03, 95% CI [-.72, .65], p = .92. When we analyzed the full set of effects for biofeedback, the overall meta-analytic estimate was at g = .65, 95% CI [.39, .92], The 95% prediction interval was wide, ranging from -0.28 to 1.59. The heterogeneity was, again, high, τ = .42, I^2 = 65.5% (35.5% due to between- and 30.0% due to within-cluster heterogeneity).

For what concerns the correction for publication bias, the 4-parameter selection model was significant g = .51, 95% CI [.09, .93], p = .02. However, the PEESE estimate provided a non-significant estimate g = .17, 95% CI [-.25, .58], p = .41. The fact that the two techniques did not converge on the same conclusion means that the overall effect for biofeedback regarding the regulation of stress should only be considered as suggestive concerning the presence of an effect. Thus, there was a lack of evidence for both strategies even when we included *all* theoretically relevant effects regardless of methodological quality or issues of reporting integrity. Table 3 summarizes the results for the overall effect of self-administered mindfulness and biofeedback.

	k	*g* *[95% CI]*	*SE*	τ	I^2	4PSM estimate	4PSM *p*-value	PET-PEESE estimate	PET-PEESE *p*-value
Self-administered mindfulness	76	.31 [.10, .51]	.07	.45	79.4%	.21 [-.22, .65]	.34	-.03 [-.72, .65]	.92
biofeedback	43	.65 [.39, .92]	.12	.42	65.5%	.51 [.09, .93]	.02	.17 [-.25, .58]	.41

Table 5: Meta-analysis for self-administered mindfulness and biofeedback. Values in brackets represent 95% *CI*.

Discussion

In this pre-registered meta-analysis, we evaluated whether two popular non-pharmacological strategies - self-administered mindfulness and biofeedback - have demonstrated efficacy in decreasing levels of stress. Random-effects meta-analysis assuming no selection on statistical significance suggested the presence of systematic, but highly heterogeneous true effects for both strategies. However, after correcting for publication bias, we found no evidence for the effects of self-administered mindfulness or biofeedback on stress regulation, regardless of whether we analyzed the results with more strict or less strict exclusion criteria. The lack of observed evidence is largely due to the high degree of heterogeneity, the low power, and publication bias, which all severely limited the interpretability of our results. The lack of evidence was mirrored for the components of the stress response and for the affective consequences of stress for both self-administered mindfulness and biofeedback.

Note that the current meta-analysis only focused on the literature in English. It is unclear how our results extend to literature in other languages. In what follows we refine our conclusions by assessing the quality of the literature and by discussing seemingly contradictory findings.

How interventions may differ across situations and across people

We were not able to draw any conclusions regarding the role of individual differences in moderating the effects of self-administered mindfulness on stress regulation, based on what we found in primary studies. Individual differences were mostly used as exclusion criteria and were not taken into account in any of the conducted analyses. It is theoretically plausible – perhaps even likely - that people differ in how much they benefit from different types of stress regulation strategies.

For traditional mindfulness protocols, some studies have investigated a potential connection between personality traits and effectiveness of mindfulness training (Tang & Braver, 2020). Trait mindfulness (i.e., the dispositional ability of being grounded in the present moment) has often been linked to mindfulness training outcomes and effectiveness (Tang & Braver, 2020), as such training seems more efficient for individuals scoring high in this trait. Moreover, neuroticism is the Big Five dimension with the strongest link with mindfulness ($r = -0.45$) according to a recent meta-analysis (Giluk et al., 2009), while De Vibe et al. (2015) found that participants higher in neuroticism showed greater improvement in psychological distress and well-being compared to those in a control group after a MBSR intervention. For biofeedback the literature appears to be less developed as to which individual differences may play a major role, compared to mindfulness. One possibility to find potentially relevant moderators for biofeedback would be to test the traits that have already been found as moderators in other stress regulation interventions (e.g., neuroticism; Schneider et al., 2012). However, because of the lack of information on the topic, our meta-analysis cannot speak to these ideas.

It is important that primary studies on stress regulation in general, and self-administered mindfulness and biofeedback specifically, start taking into account personality traits that might play a role in the effectiveness of stress regulation strategies. When researchers record these personality traits and other potentially relevant aspects, they can deposit them in a database so that meta-analysts can understand across different studies for whom, under which conditions, and with which kind of parameters the interventions function most optimally. We created a protocol in Appendix R, which includes traits that we think should be measured to better understand how stress is regulated by different people.

Quality of the literature

The most pernicious issue concerns the quality of the stress regulation literature. For self-administered mindfulness, we had to exclude 54% of the effects because they were based on inconsistent means or standard deviations or because the study designs they came from were judged to be at a high risk of bias. For biofeedback, the percentage of excluded studies was lower but still substantial, with approximately one third of the effects being excluded. We nevertheless repeated our analyses this time including these low-quality studies, but these did not change our conclusions[15].

Another problematic issue is that, for most of the studies (30/32 for self-administered mindfulness and 23/31 for biofeedback) included in our meta-analysis, the experimental group was matched with a passive instead of an active control condition. Future studies need to employ a well-matched active control that is structurally equivalent to the experimental

[15] For biofeedback, when we excluded effects from designs having high risk of studies ($k = 28$), both bias-correction techniques failed to find an effect. However, when we analyzed the full set of effects ($k = 41$) the 4-parameter selection model suggested the presence of an effect, while the PEESE was not able to detect any signal.

condition for a more accurate estimate of the effect of interest (MacCoon et al., 2013). Even when we exclude these low-quality studies, we conclude that both literatures suffer from publication bias and from substantial heterogeneity, severely limiting what we can infer from the literature in terms of evidential value.

Why do the results from different techniques differ?

When analyzing the data, the naive meta-analytic estimate and the *p*-curve seemed to provide evidence in favor of the efficacy of self-administered mindfulness and biofeedback strategies, while PET-PEESE and 4PSM did not provide evidence in favor of the efficacy of either strategy. *P*-curve, however, is known to overestimate the effect under conditions of substantial heterogeneity (Simonsohn, Nelson, & Simmons, 2014), while PET-PEESE has been found to perform poorly in scenarios when used on a small number of studies that are underpowered and are highly heterogeneous (Carter et al., 2019). This means that because of the relative imprecision of these publication bias techniques, we have an effect that has been corrected, but we cannot know how accurate it is.

When considering different interventions to downregulate stress based on the available evidence, it is vital to be accurate and clear about which evidence exists. Choosing the wrong strategy could have no effect on stress regulation or, worse, cause adverse effects (IJzerman et al., 2020). For instance, mindfulness interventions have at times (and anecdotally) resulted in psychotic episodes, panic attacks, and depersonalization (Van Gordon et al., 2017). Moreover, some studies have even found that mindfulness interventions do not reduce people's self-focus and desire to be better than others, but instead bolstered self-enhancement (Gebauer et al., 2018; Vaughan-Johnston et al., 2021).

We thus believe that our conservative approach was the most suitable: Only if all bias correction techniques point in the same direction did we accept that the stress regulation strategy was efficacious. At present, our best answer is that we simply do not know whether biofeedback or self-administered mindfulness work in downregulating stress. The only way to answer this question is by improving the quality of primary studies in these literatures.

Prevalence and use of self-administered mindfulness and biofeedback

Self-administered mindfulness has seen a rise of accessibility and popularity in recent years (Cavanagh et al., 2018). This increase was due to a combination of factors, such as the secularization of the Buddhist concept of mindfulness, introduced in North America by Jon Kabat-Zinn in 1970, and the emergence of technologies (e.g., smartphone applications) that enabled people's access to forms of self-administered mindfulness. In particular, in the last few years there has been a surge in downloads of mindfulness applications, especially after the COVID-19 outbreak. For example, 'Calm', the top English-language meditation app, saw almost 4 million downloads in April 2020, the month in which more people around the globe experienced the first lockdown (Perez, 2020).

A large and very heterogeneous group of people very likely tried to apply self-administer mindfulness through these apps. But to date, there is no evidence that self-administered mindfulness is efficacious in reducing stress, depression, or anxiety. Moreover, *if* self-administered mindfulness does work, we have no knowledge how it functions across different individuals. A number of studies have been conducted using Headspace (Headspace, n.d.), a widely used mindfulness application, to investigate the beneficial effects of self-administered mindfulness. Two studies using Headspace for its intervention met our inclusion criteria for our meta-analysis and they had similar problems as what we observed throughout the literature: They were underpowered (Economides et al., 2018; Rich et al., 2021) and/or only employed a passive, not an active, control condition (Rich et al., 2021). Given the lack of evidence behind self-administered mindfulness that we observe in our meta-analysis, we warn the reader to put too much stock into the evidence cited by such commercial initiatives until high-quality pre-registered studies are conducted about their efficacy.

Heart rate variability biofeedback saw a smaller increase in prevalence than self-administered mindfulness, perhaps due to a general higher cost of the practice. These higher costs may be the reason that biofeedback is often more employed in clinical settings, such as in the treatment of some disorders (asthma, COPD, fibromyalgia; Gevirtz, 2013) than in non-clinical settings. Consistent with this possibility, eight studies out of twenty included in our meta-analysis involved a clinical population. Although various manuals and protocols exist - in contrast to self-administered mindfulness - the need for high-quality, pre-registered studies here is similar as for self-administered mindfulness before clinicians apply them in their practice. We simply do not observe evidence behind the idea that biofeedback downregulates stress, depression, or anxiety. In what follows we tried to understand why we were unable to find evidence, followed by methodological recommendations for future studies.

The signal might be present, but we were unable to detect it

The fact that we did not detect a signal does not necessarily mean that self-administered mindfulness or biofeedback are not effective in reducing stress levels. Indeed, the primary studies we analyzed were either of insufficient quality or they were largely underpowered to detect a small-to-modest effect size. The median power for detecting a small effect size (d = .20) was at .21 for self-administered mindfulness and .11 for biofeedback. Studies on these two strategies require much larger sample sizes to detect a signal, if it is present.

Our conclusions seemingly contradict prior meta-analyses on self-administered mindfulness and biofeedback, who did detect a moderate size effect (Cavanagh et al., 2014; Goessl et al., 2017; Taylor et al., 2021). However, we think that the conclusions of our and their meta-analyses are largely consistent with each other and that the differences can be explained by two factors. First, our selection criteria were much stricter in checking the quality of the literature: a) In comparison to Taylor et al. (2021), we excluded studies based on mathematically inconsistent data (they did not). Further b) their active comparator conditions consisted of many heterogeneous treatments (e.g., progressive muscle relaxation, psychoeducation for pain and stress managements). With such a heterogeneous active control group their estimation of the effect sizes tends to be biased, because the mindfulness groups are compared with treatments with a different effectiveness in terms of stress reduction. For that reason, we instead opted for active control conditions in which participants engaged in some tasks *not* intended to diminish levels of stress. All in all, this means that we ended up with much fewer studies (26 versus 47 for the relevant variables, not taking into account the different additional foci by Taylor et al. on wellbeing and quality of life that we did not focus on, which consisted of a further 36 number of studies).

Second, our meta-analysis relies on the state-of-the-art in publication bias correction techniques and provides a much more accurate estimate of the effect size that can be expected when using similar sample sizes and study designs in comparison to all three prior meta-analyses on the topic (Cavanagh et al., 2015; Goessl et al., 2017; Taylor et al., 2021). All together, we believe that our stringent workflow provides a much more precise estimate of the target effect, as compared to the three prior meta-analyses. Before correction of the effect, we also find moderately sized effects, but these effects disappear when we provide our stricter - and more accurate - corrections.

The efficacy of these two strategies may depend on the type of protocol used

For both strategies we observed a substantial variance for what concerns the type of the treatment. For self-administered mindfulness, the number of intervention sessions in the included studies ranged from 1 to 64 (M = 29.3; Mdn = 27). The same was true for biofeedback, where the number of interventions ranged from 1 to 40 (M = 9.5; Mdn = 6). The total time spent by participants doing meditation or alternatively having a biofeedback intervention also varied substantially: Self-administered mindfulness included really brief (lasting 5 minutes) and longer interventions (lasting overall 16 hours), the same being true for biofeedback (5 minutes for the shortest to 22 hours for the longest). Furthermore, the total duration of the intervention could last a single day or up to two months for self-administered mindfulness and could range from a single day up to six weeks for biofeedback.

The heterogeneity in the self-administered mindfulness literature should not come as a surprise. Self-administered mindfulness does not currently and necessarily follow a well-established protocol like other traditional mindfulness programs such as mindfulness-based stress reduction (MBSR; Kabat-Zinn, 2008) or mindfulness-based cognitive therapy (MBCT; Segal et al., 2002). Similarly, attempts have been made to develop a manualized heart rate variability protocol for biofeedback. Lehrer et al. (2000) developed a ten-sessions HRV-biofeedback protocol with two kinds of instruments, for office and home use respectively. This protocol was further improved (Lehrer et al., 2013) and the number of sessions was halved. The reason for shortening this protocol is that HRV biofeedback could be learned in fewer visits, according to Vaschillo, Vaschillo, and Lehrer (2006).

However, as we have seen in our meta-analysis, each study seemed to rely on a different pattern regarding the length, the frequency, and duration of sessions. This suggests that there is no consensus on a standardized protocol for experimental and clinical practice. According to the results of our meta-analysis, the frequency of the intervention could potentially play a role in the efficacy in reducing stress (i.e., the more weekly interventions, the lower the experienced stress). The inferences from these analyses should, nevertheless, be regarded with considerable caution, as we could not apply publication bias correction techniques to these subgroup analyses.

Keys area for improvement

Given the current state of the literature on stress, rigorous experimental studies must be conducted to evaluate whether such regulation strategies are able to effectively help decrease the level of stress. Potential sources of bias (e.g., selective reporting) can be mitigated by the implementation of open sciences practices such as pre-registration and Registered Report (a peer-reviewed pre-registration; Chambers, 2013). To further improve the literature, we recommend publishing data of studies on stress regulation strategies (including, but not limited to, self-administered mindfulness and biofeedback) into an open repository (e.g., PsychOpen CAMA; Burgard et al., 2021). This would foster transparency and would allow other researchers to contribute with their data to accumulate evidence of different stress regulation strategies. In sharing data to an open repository, we recommend using our protocol in Appendix I so that meta-analysts can understand the differences across context and people across studies.

Constraints on Generality section (Simmons et al., 2017)

Results of our meta-analysis apply only to studies that satisfied the inclusion criteria of our meta-analytic approach. For instance, we do not know whether our findings are the same for people under 18 and we are not aware if studies of other languages would have led to different results. Nevertheless, given the state of the literature, we expect similar problems with publication bias to hold.

Conclusions

We conducted a meta-analysis of strategies contributing to stress regulation, including 26 articles and 35 effects for self-administered mindfulness and 20 articles and 31 effects for biofeedback. The results do not allow us to conclude that either strategy has a demonstrated efficacy in reducing stress. We therefore suggest Registered Reports as an intervention to better understand the (lack of) efficacy of either strategy, taking into account the limitations of the existing literature that we highlighted in the present meta-analytic synthesis.

Declaration of Competing Interest

Declarations of interest: None.

Data availability

Underlying data

Data from this article can be found on the OSF page (http://bit.ly/39MifmH).

Extended data

To ensure methodological rigor and transparency, we have made our data and the script available on the Open Science Framework (http://bit.ly/39MifmH)

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0). At least some of the data/evidence that will be used to answer the research question already exists AND is accessible in principle to the authors (e.g., residing in a public dataset or with a colleague). The authors used the data to create a coding scheme BUT the authors certify that they have not yet accessed any part of summary statistics.

Contributors

Conceptualization: Alessandro Sparacio, Ivan Ropovik, Gabriela Jiga-Boy, and Hans IJzerman.

Data curation: Alessandro Sparacio and Ivan Ropovik.

Formal analysis: Alessandro Sparacio, Ivan Ropovik, and Hans IJzerman.

Funding acquisition: Hans IJzerman and Gabriela Jiga-Boy.

Investigation: Alessandro Sparacio.

Methodology: Alessandro Sparacio, Ivan Ropovik, and Hans IJzerman.

Project administration: Alessandro Sparacio, Ivan Ropovik, Gabriela Jiga-Boy, and Hans IJzerman.

Resources: Alessandro Sparacio, Ivan Ropovik, Gabriela Jiga-Boy, and Hans IJzerman.

Software: Alessandro Sparacio and Ivan Ropovik.

Supervision: Ivan Ropovik, Gabriela Jiga-Boy, Patrick Forscher and Hans IJzerman.

Validation: Ivan Ropovik, Gabriela Jiga-Boy, and Hans IJzerman.

Visualization: Alessandro Sparacio, Bastien Paris and Ivan Ropovik.

Writing - original draft: Alessandro Sparacio, Ivan Ropovik, and Hans IJzerman.

Writing - review & editing: Alessandro Sparacio, Ivan Ropovik, Gabriela Jiga-Boy, and Hans IJzerman**.**

Preface of Study III

In the following chapter of the present Ph.D. project we wanted to evaluate more in depth the efficacy of one of the strategies we included in our previous meta-analyses with a multi-site project. Out of the four strategies, we choose self-administered mindfulness. The choice was partly arbitrary, and partly due to the fact that self-administered mindfulness was the most scalable intervention out of the four. The pre-registered meta-analysis showed that the quality of self-administered mindfulness studies available currently was insufficient; this multi-site study was therefore one of the first sufficiently powered and high-quality studies to be conducted on the topic. Additionally, it is unclear whether single brief stand-alone mindfulness exercises are sufficient to reduce stress of individuals. Thus, the scope of this multi-site project: Test four self-administered mindfulness exercises (i.e., Body Scan, Mindful Walking, Mindful Breathing, Loving-kindness) to see whether they reduce stress levels of individuals.

Chapter 4: A large, multi-site test of self-administered mindfulness among English speakers

This project was pre-registered on March 23th, 2022.

Author Note: Our pre-registration document can be found at https://osf.io/6w2zm/. Our OSF page can be found here (https://osf.io/6w2zm/). The authors that contributed to this manuscript are: Alessandro Sparacio, Hans IJzerman, Ivan Ropovik, Filippo Giorgini, Christoph Spiessens, Bert N. Uchino, Joshua Landvatter, Tracey Tacana, Sandra J. Diller, Jaye L. Derrick, Joahana Segundo, Jace D. Pierce, Robert M. Ross, Zoë Francis, Amanda LaBoucane, Christine Ma-Kellams, Maire B. Ford, Kathleen Schmidt, Celia C. Wong, Wendy C. Higgins, Bryant M. Stone, Samantha K. Stanley, Gianni Ribeiro, Paul T. Fuglestad, Valerie Jaklin, Andrea Kübler, Philipp Ziebell, Crystal L. Jewell, Yulia Kovas, Mahnoosh Allahghadri, Michael F. Baranski, Hannah Burgess, Annika B. E. Benz, Maysa DeSousa[27], Catherine E. Nylin, Janae C. Brooks, Caitlyn M. Goldsmith, Jessica M. Benson, Siobhán M. Griffin, Stephen Dunne, William E. Davis, Tam J. Watermeyer, William B. Meese, Jennifer L. Howell, Laurel Standiford Reyes, Megan G. Strickland, Sally S. Dickerson, Samantha Pescatore, Shayna Skakoon-Sparling, Zachary I. Wunder, Martin V. Day, Shawna Brenton, Audrey H. Linden, Christopher E. Hawk, Léan V. O'Brien, Tenzin Urgyen, Jennifer S. McDonald, Gabriela Jiga-Boy

Abstract

Over the past several years, self-administered mindfulness interventions have become increasingly popular. However, the effectiveness of these interventions for regulating stress is unclear. In the present multi-site study (N_{sites} = 37, $N_{participants}$ = 2,239; all fluent English speakers), we investigated the efficacy of four single, brief stand-alone mindfulness exercises (versus three active control conditions) on self-reported stress with Bayesian mixed-effects models. All mindfulness exercises proved to be more efficacious than the active control conditions in reducing participants' self-reported stress levels. Between the control condition (M = 1.95) and the condition with lowest stress levels (Body Scan; M = 1.68), there was a mean difference of 0.27 on a five-point scale, (Cohen's d = -.56) indicating a small decrease in stress. These results suggest that brief mindfulness exercises can reduce stress, even if we cannot fully disentangle the presence of demand effects. Future studies should use designs that minimize demand effects (e.g., using physiological measurements) together with self-reported measures to further test the efficacy of self-administered mindfulness on stress regulation.

A large, multi-site test of self-administered mindfulness among English speakers

Engaging in mindfulness meditation, in theory, appears simple: One is asked to sit and focus one's attention on the breath and on the present moment, without needing complex settings or apparatus. Partly due to this apparent simplicity, mindfulness meditation protocols that can be self-administered have increased in accessibility and popularity in recent years (Cavanagh et al., 2018).

Self-administered mindfulness protocols are free or less expensive compared to traditional protocols (Spijkerman, Pots, & Bohlmeijer, 2016) and readily available (Cavanagh et al. 2014; Wahbeh et al. 2014). These mindfulness interventions can be accessed as self-help books, or via computer programs, smartphone apps, or audio and video recordings. For example, the last few years witnessed a surge in downloads of mindfulness applications such as 'Calm', with 40 million downloads (Gebel, 2019), and 'Headspace', reportedly with 70 million downloads (Headspace, 2021). Features such as the duration or frequency of these interventions are less certain, because while the applications might offer modular sessions of varying lengths, the onus is on the consumer to decide for how long and how often to engage with these apps.

Despite their easy accessibility, the evidence behind these mindfulness interventions is debated and some questions remain unanswered, such as: Are these types of intervention truly effective in regulating individuals' stress levels? Or which types of self-administered mindfulness exercises, from the plethora of those available, work to downregulate stress? We attempted to answer these questions first by conducting a survey aimed at mindfulness practitioners in order to identify which are the most used and the most popular mindfulness exercises. Based on the results of the survey, we then designed a multi-site, highly powered study to test the effects and the boundary conditions of four types of self-administered mindfulness meditation exercises on stress regulation.

Can self-administered mindfulness interventions reduce levels of stress?

Mindfulness can be defined as “paying attention in a particular way: On purpose, in the present moment, and nonjudgmentally” (Kabat-Zinn, 1994, p. 4). Mindfulness emphasizes attention to the present moment, with awareness of one’s bodily sensations or one’s mental content (thoughts, emotions, memories, etc.). Self-administered mindfulness interventions share these features, but compared to traditional established mindfulness protocols (e.g., mindfulness-based stress reduction - MBSR or mindfulness-based cognitive therapy - MBCT), do not require the presence of an instructor or and the commitment to be in a particular place and time regularly, for 8 weeks, allowing practitioners to meditate whenever they choose so (Spijkerman, Pots, & Bohlmeijer, 2016).

Because self-administered mindfulness interventions hold promises for reducing stress levels and are accessible to virtually anyone, it is important to understand whether they bring about the expected results. While some studies have shown positive results in terms of reducing self-reported stress (Cavanagh et al., 2013; Cavanagh et al., 2018), and a recent meta-analysis pointed in the same direction (Taylor et al., 2021), other studies reported null results. For instance, Glück and Maercker (2011) did not find evidence that a self-guided mindfulness training effectively decreased levels of perceived stress and distress for a group of meditators compared to a waitlist control group. Finally, a meta-analysis failed to find robust evidence regarding the efficacy of self-administered mindfulness interventions on stress reduction after applying techniques that account for publication bias (Sparacio et al., 2022).

Which types of self-administered mindfulness exercises work to downregulate stress?

Another reason why we do not know yet whether self-administered mindfulness interventions are efficacious at regulating stress is that self-administered mindfulness studies include a range of different exercises (e.g., Body Scan, Mindful Breathing, Mindful Walking, etc.); current tests evaluate protocols in their entirety, instead of differentiating among these exercises. According to Germer et al. (2016), there are a plethora of mindfulness exercises that can be divided into three different categories: "Awareness", "present experience", and "acceptance". These different categories are not truly separate, nonetheless we suggest that splitting them into three different groups allows us to better understand the potential applied value of certain mindfulness exercises that can be self-administered.

"Awareness" mindfulness exercises are typically characterized by a series of steps: a) Disengage from an automatic train of thoughts (e.g., Interrupting repetitive thoughts by taking a long breath), b) focus the attention on an object that is used as an "anchor" (e.g., bringing the attention to the breath or to parts of the body), and c) return the attention to the object of focus when one realizes they have been distracted, and then watch where the mind wanders next.

"Present experience" mindfulness exercises instruct participants to pay attention completely to the activity being carried out (e.g., bringing the attention to the sole of the foot while walking). If the mind wanders, the attention is redirected to the present moment and the instructions that are given in this set of exercises have the function of fulfilling this objective.

"Acceptance" mindfulness exercises are characterized by applying a non-judgmental attitude of kindness and curiosity to one's experience. Practitioners that do these types of exercises are invited to cultivate positive feelings towards themselves and to others (e.g., directing Loving-Kindness to themselves or to someone else).

We currently do not know which category of the exercises described above is most effective for reducing stress levels. Some studies have tested single brief mindfulness exercises (Feldman et al., 2010; Hutcherson et al., 2008), however, to our knowledge, none investigated the effectiveness of brief stand-alone mindfulness exercises on stress reduction. Thus, the aim of our multi-site project is to test which self-administered mindfulness exercises were more effective in reducing individuals' stress levels, compared to three non-mindful active control conditions randomly sampled.

Study 1: Survey on self-administered mindfulness

To identify the most representative mindfulness exercises to use in our multi-site project, we conducted a survey to ask mindfulness practitioners for suggestions regarding self-administered mindfulness exercises. We then came across a study of Matko et al., (2019) that provided a list of the 20 most used meditation techniques. To compile a final list of exercises, we retained the most popular exercises suggested by the practitioners taking part in the survey that were also included in the study of Matko et al., (2019).

Procedure

We disseminated a Qualtrics survey to mindfulness practitioners (i.e., instructors and/or researchers) that we retrieved from lists of subscribers of various mindfulness associations and via social media networks (Facebook, Twitter, & LinkedIn) between November 17th and December 31st, 2020. Ninety-five mindfulness practitioners started the survey, 40 completed half of it, and 20 respondents reached the end.

The survey was divided into two parts: The first part asked respondents to suggest low-intensity self-administered mindfulness interventions designed to decrease short-term stress (i.e., 'state-level' stress), while the second part asked them to suggest high-intensity self-administered mindfulness interventions that can be aimed at changing personality traits that can impact on stress levels (i.e., 'trait-level' stress). A low-intensity self-administered mindfulness intervention was defined as a "mindfulness intervention that is brief, but long enough to have an impact on stress levels of individuals and give participants flexibility (to ensure that they do not abandon the study), while being effective in reducing stress". We did not expect this low-intensity intervention to substantially impact on traits of participants that predisposes them to higher levels of stress (e.g., trait anxiety, tendency to ruminate). A high-intensity self-administered mindfulness intervention was defined as a "mindfulness intervention that is longer, less flexible, but that might impact not only on stress levels of participants, but also on personality traits of participants that are relevant to the buildup of stress (e.g., trait anxiety, tendency to ruminate)."

We asked the same set of questions for both low and high-intensity self-administered mindfulness interventions, namely: 1) How long the intervention should last, in order to observe a significant reduction of participants' stress level (*from 1 day to 8 weeks or more*), 2) How many times a week participants should do exercises of self-administered mindfulness? (*1 to 7/week*), 3) Which type of source they thought would be more effective for delivering the intervention (*smartphone app, video file, audio file, self-help book, another type of source - specify which, or 'the source is not relevant'*), and 4) What types of exercises should be included in the intervention (*e.g., Body Scan, focus on breathing, etc.*). For each exercise they suggested, we asked them 1) If the order of the exercises they listed in the previous question was important to achieve stress reduction (*yes, maybe, no*), and 2) How long they thought it should be practiced and how often (*length - or daily duration of the exercise in minutes; weekly frequency*; *number of weeks of the intervention*).

Results

Results of our survey showed that respondents: 1) estimated that smartphone apps were the most effective types of self-administered mindfulness exercise for reducing stress levels as compared to other cited sources (e.g., books), 2) considered that the order of the mindfulness exercises could potentially be relevant to impact levels of stress[16] and 3) assessed that practicing mindfulness daily had the greatest impact on stress reduction. A full

[16] 18 respondents answered "maybe" to this question, 14 answered "yes" and 7 answered "no".

description of the results is beyond the scope of this manuscript, however a R script with the answers of the respondents and a complete report of the survey in PDF can be found on the OSF page of the project (https://osf.io/aw2vz/).

For planning the multi-site study, we relied on practitioners' answers that included a list of exercises to be considered for a low-intensity self-administered mindfulness intervention[17]. The list compiled practitioners' suggestions and consisted, initially, of 174 mindfulness exercises (see the list of the exercises document at https://osf.io/br86s/). We then removed duplicates and we merged the techniques that we considered to be similar[18] (e.g., 'Mindful Walking' and 'walk in mindfulness') or that were spelled differently but referred to the same exercise (e.g., 'Body Scanning' and 'Body Scan'). We then organized each technique according to its dominant element of mindfulness based on Germer's (2016) system of categorization, which included three categories previously mentioned: 'awareness', 'present experience', and 'acceptance'. Some mindfulness exercises suggested by the practitioners were already categorized according to Germer's system (e.g., 'Body Scan' was placed as belonging to the category 'awareness'). For the exercises that were not already categorized, the first author made an arbitrary choice based on the similarities between the uncategorized exercise and the other exercises already present in that category (e.g., 'mindful movements' was placed in the category 'present experience').

Finally, for each dominant element, we retained the two exercises that were the most cited: Body Scan (39 votes) and Mindful Breathing (37 votes) for 'awareness', Mindful Walking (10 votes) and mindful movement (10 votes) for 'present experience', and self-

[17] When we designed the survey, we had not yet outlined the details of the multi-site project, so we asked for general information related to self-administered mindfulness protocols. For the purpose of the multi-site project we used only a portion of all the information provided by the answers of mindfulness practitioners, however we believe that the unused information may help other mindfulness researchers design more rigorous studies on the topic.

[18] This process was done by the first author, however it is possible to infer how the simplification was done by comparing the files with all exercises (i.e., List of the exercises) and the one where similar exercises were merged (i.e., Mindfulness clusters MULTI-SITE; see https://osf.io/br86s/).

compassion break (5 votes) and Loving-Kindness meditation (3 votes) for 'acceptance' (See 'Mindfulness clusters MULTI-SITE' at https://osf.io/br86s/).

We then cross-referenced our set of self-administered mindfulness exercises with the list of 20 popular meditation techniques that Matko et al., (2019) detailed by consulting 637 meditators with an online survey. By retaining the exercises that were present in both approaches we arrived at this final list of self-administered mindfulness exercises: Body Scan, Mindful Breathing, Mindful Walking and Loving-Kindness meditation. We decided to test these four mindfulness exercises in the multi-site project described below.

Study 2: The multi-site project

Hypothesis

We hypothesized that participants allocated to any one of the experimental (mindfulness) conditions would experience lower levels of self-reported stress compared to participants allocated to one of the active control conditions[19]. We also explored effects on the dimensions of pleasure, arousal, and dominance of any of the experimental conditions as compared to the active control conditions. Similarly, we explored the moderating role of neuroticism in reducing levels of stress when comparing the experimental and control conditions.

Methods

Materials

This research was conducted in line with the CO-RE Lab Lab Philosophy v5 (Silan, Bellemin, Dujols, Sparacio, Adetula, & IJzerman, 2021). All materials used in the study, including the pre-registered document (https://osf.io/us5ae), the ethics (IRB) approval

[19] Participants who are randomly assigned to the control group will be randomly assigned to listen to an excerpt of one of the three stories that we selected for the present study.

documents of all the sites involved in the project and the meditations scripts are available on our OSF page (https://osf.io/6w2zm/). The data analytic script instead can be found on the GitHub repository of the project (https://github.com/alessandro992/A-large-multi-site-test-of-self-administered-mindfulness).

Participants and sites

As the audio files were in English, we limited the study to English native speakers or participants who self-assessed their level of English proficiency at the C1/C2 levels from the Common European Framework of Reference for Languages (Council of Europe, 2001). Participants were excluded if they reported having or having had a history of mental illnesses assessed via a pre-screening question, if they declared having meditated in the previous 6 months, or if they did not match the English level required. Each participant was asked to take part in the survey using a smartphone with headphones or earphones attached, to allow participants perform one of the mindfulness activities (i.e., Mindful Walking) that required walking. We asked each site to collect a number of participants that ranged from 70 to 120[20]; however, if a site collected fewer or more participants than was the target, we still used the data from those participants in the analysis. The data collection started on March 23rd 2022 and continued until June 30th.

Exclusion of participants

After removing test answers, we had a total of 3,627 answers to our survey. Among all participants, 1,307 were meditators or declared that they had meditated in the previous 6 months of the experiment, 776 did not match the English level requirement, and 981 had a history of mental illnesses. After removing participants that did not meet our inclusion criteria the number dropped to 2,463. We then removed 224 participants that according to our

[20] Each site collected a different number of participants; the site with the fewest participants collected only one, while the site with the highest number of participants reached 179. Our Rpubs page displays the total number of participants per site (https://rpubs.com/ale-sparacio92/920457).

criteria were judged as being careless respondents and we had a final sample of 2,239 valid observations (of these, 611 self-identified as males, 1,576 as females, seven as transgender males, two as transgender females, 27 did not identify with any choice, 16 preferred not to say; M_{age} = 22.4, SD_{age}= 10.1; range 17-87; 94.2% students). Table 6 displays the number of valid observations for each condition.

Condition	Mindful Walking	Mindful Breathing	Loving - Kindness	Body Scan	Book chapter (control)
N	416	469	427	449	478

Table 6: Distribution of participants across conditions

Distribution of participants across sites

Of the original 42 sites that expressed interest in participating in the multi-site study, 37 collected data. A list of the sites that participated in the data collection can be found on the OSF page of the project (https://osf.io/uh3pk). Participants could be recruited through the SONA system of the respective institution or via crowdsourcing platforms such as mTurk or Prolific academic. Participants could come from any geographic area if they met our inclusion criteria and could be given either credits or financial compensation in exchange for participating in the study.

Simulation of the sequential Bayesian design.

Prior to the data collection, we simulated data based on a Bayes Factor Design Analysis (BFDA; Schönbrodt & Wagenmakers, 2018) to assess the expected efficiency and informativeness of the present design. The aim of the simulation was to establish (1) the expected likelihood of the study to provide compelling relative evidence either in favor of H_0 (BF_{01} = 1/10) or H_1 (BF_{10} = 10); (2) the likelihood of obtaining convincing but misleading evidence, and (3) the likelihood that the study points into the correct direction even if stopped earlier due pragmatic constraints on sample size (Schönbrodt & Wagenmakers, 2018).

Given these aims, we modeled a sequential design with maximum N where the data collection continues until either the threshold for compelling evidence is met or the maximum N is reached. Despite forty-one labs indicated interest in the project, we took the conservative estimate of 30 data-collecting labs. Each lab was expected to collect data of at least $N = 70$ participants, with a maximum N at 120 (translating to $Min = 420$ and $Max = 720$ participants per condition). Our goal was to be able to detect an effect size of $d = 0.20$; we modeled the true value to vary between labs by repeatedly (for each simulation) drawing from a normal distribution, $\delta \sim N(0.20, 0.05)$, with a 95% probability that the effect size falls between $d =$ 0.10 and 0.30.

We tested the effectiveness of four stand-alone interventions using a between-participants adaptive group design, where upon hitting a threshold of compelling evidence in one condition, we planned to allocate the rest of the participants into other conditions where the threshold had not been met yet. The simulation, however, assumed a conservative scenario with equal N across all conditions, therefore, simplifying the computations to a single between-participants t-test scenario.

The results (see Figure 8) show that, given the assumed design, the probability of the test to arrive at the boundary of compelling evidence (BF = 10 or 1/10) was .79 (.72 at H_1 and .07 erroneously at H_0). The probability of terminating at maximum N of 720 per condition was .21; .05 of showing some evidence for H_1 ($BF > 3$), .13 of being inconclusive ($3 > BF >$ 1/3), and .03 of showing evidence for H_0 ($BF < 1/3$). For the test of a single condition against controls, the sequential design is expected to be 27% more efficient than collecting a fixed maximum N per lab, with the average N at stopping point (BF boundary and maximum N) at 526. Even conservatively assuming a balanced-N situation, the informativeness of the design thus appears to be adequate and the use of the adaptive design would likely enhance informativeness and/or resource efficiency. Figure 8 shows the results of the simulation of the sequential design with 10,000 iterations.

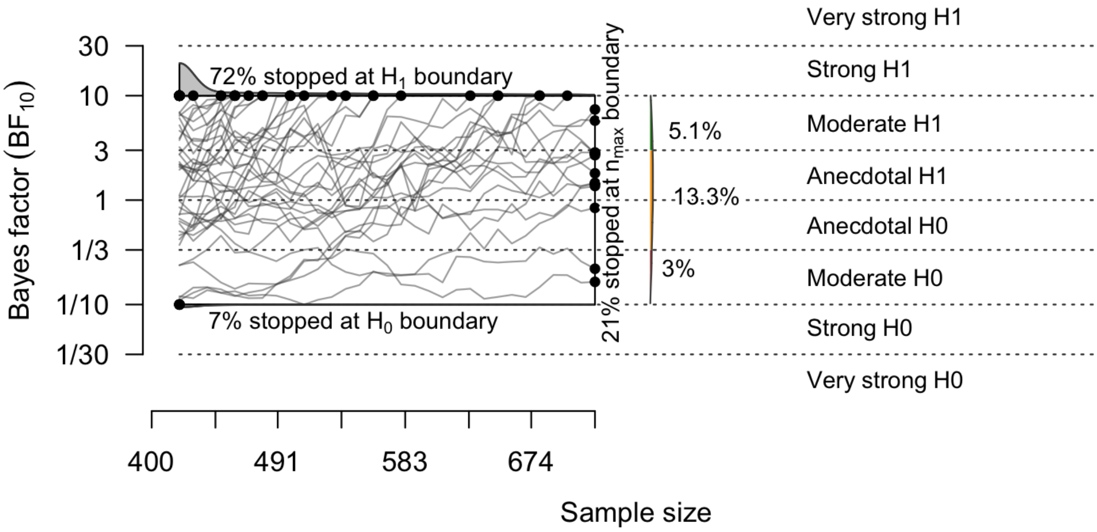


Figure 8. Design analysis for the sequential Bayesian design with maximum *N*

Procedure

Participants accessed the experiment via a Qualtrics link. We provided participants with detailed information about the study (see "Participants Information Sheet" included in the IRB package, https://osf.io/6w2zm/) and we asked them to provide consent prior to participating. We asked each participant to take part in the survey using a smartphone with headphones or earphones attached and asked them not to use a computer or laptop. We included a question in which we asked participants whether they started the survey from a device other than a smartphone; if they answered negatively, we asked them to exit the survey and to restart it, this time using a smartphone with headphones or earphones attached.

We then asked participants to sit in a quiet place such as a room where they would not be disturbed for 20 minutes. After providing informed consent, we invited participants to complete the Neuroticism subscale of the IPIP - 5 NEO domains (IPIP, 2001) to assess their self-reported levels of neuroticism. Examples of items were “I often feel blue” or “I am filled with doubts about things” (answered on a five-point scale ranging from 1 = *Very inaccurate*, to 5 = *Very Accurate*; $\omega_u = 0.90$[21]).

After the completion of the neuroticism measure, we randomly allocated participants to one of five conditions[22]. In each condition, participants were presented with a different 15-minutes recording. The four audio files of the mindfulness exercises and the three audio files of the stories of the non-mindful active control condition were recorded by the same certified meditation trainer. Below we provide a description of the mindfulness conditions and of the three different stories of the non-mindful active control condition. The exact text of the seven meditations and of the three stories used in the active control conditions can be found at our OSF project page (https://osf.io/6w2zm/). The seven recordings instead can be found at the Soundcloud page of the project (https://soundcloud.com/listening-385769822).

Body Scan. In Body Scan, the meditation trainer invited participants to “scan” their parts of the body. Every time the mind wandered, the meditation trainer invited participants to bring back the awareness and the attention to the part of their body they were “scanning”.

Mindful Breathing. During Mindful Breathing, the meditation trainer invited participants to “stay with their breath”, without changing the way they were breathing. When their mind wandered, the meditation trainer told participants to bring their attention back to their breath with kindness and patience.

Loving-Kindness meditation. During this practice, the meditation trainer encouraged participants to direct Loving-Kindness toward themselves and then to extend these feelings of Loving-Kindness towards somebody else.

[21] The subscript *u* indicates that we used the variant of omega based on a unidimensional model as the IPIP - 5 NEO scale measured an unidimensional construct.
[22] Participants in the non-mindful active control condition could listen to the excerpt of one of three stories that we selected with an equal chance of getting one story over another.

Mindful Walking. During Mindful Walking, the meditation trainer asked participants to walk in a quiet place (preferably indoors or in a place as isolated as possible from distractions), while listening to the instructions. During this practice, the meditation trainer invited participants to bring their awareness to the experience of walking and subsequently the meditator invited them to "feel" the physical sensations of contact of their feet with the ground.

Non-mindful active control condition. Participants in the active control condition had an equal probability of listening to an excerpt coming from one of these three books: "Silverview" (word count: 1838) by John le Carré (Le Carré's, 2021), "the old man and the sea" (word count: 2039) by Ernest Hemingway (Hemingway, 1952) and "Smith of Wootton Major" (word count: 2309) by J. R. R. Tolkien (Tolkien, 1967). We added more than one story to increase variance of the control conditions, allowing for somewhat greater generalizability across stimuli (Judd et al., 2012). These three stories were chosen because they had a similar word count, were written in easy-to-understand English, and did not feature major plot changes. Altogether, this meant that they were unlikely to elicit strong emotions.

State-Trait Anxiety scale. After the 15 minutes of mindfulness exercises or active control listening, participants first answered the 20-item State-Trait Anxiety Inventory, Form Y (STAI; Spielberger, Gorshu & Lushene, 1970). Participants were asked to indicate how they felt in that exact moment on a 4-points scale (1 = *Not at all*, 2 = *Somewhat*, 3 = *Moderately so,* 4 = *Very much so*) on 20 items (e.g., "I am tense"; "I feel frightened"; ω_u = 0.92[23]). In line with our pre-registration, we only analyzed the scale items assessing state

[23] The subscript *u* indicates that we used the variant of omega based on a unidimensional model as the State-Trait Anxiety scale measured an unidimensional construct.

anxiety as we were interested in self-reported stress of individuals with a short-term efficacy span.

Self-Assessment Manikin scale. Participants then filled in the Self-Assessment Manikin scale, a non-verbal pictorial assessment technique that measures three different dimensions: Pleasure, arousal, and dominance in relation to an object or a stimulus (Bradley & Lang, 1994). We invited participants to observe a series of images and we asked them to select the image that best represented their emotional state in that exact moment. The Self-Assessment Manikin scale is the picture-oriented version of the widely used Semantic Differential Scale devised by Mehrabian and Russell (1974). This instrument measures the three-dimensional structure of stimuli, objects, and situations with 18 bipolar adjective pairs that can be rated along a 9-point scale. Factor analyses of the resulting 18 ratings generate scores on the dimensions of pleasure, arousal, and dominance. The correlations between scores obtained using the Self-Assessment Manikin scale and those obtained from Mehrabian and Russell (1977) emotion adjectives are high for pleasure (.93) and arousal (.93) and lower for dominance (.66; Morris, 1995; Lang. 1985). As compared to the tool of Mehrabian and Russell (1974), the Self-Assessment Manikin scale requires a limited investment of effort and time and does not rely on a verbal rating system making this tool available also for non-English speakers.

Demographics. Finally, participants answered questions related to their demographic information: Age, gender, country of birth, country of residence, and whether they were students or not. We asked student participants to indicate in which university they were studying and if participants were not currently students, we inquired what their current job is. At the end, we thanked and debriefed participants. If sites awarded credit or payment, then participants were redirected to another, locally hosted page to receive credits/payments for completing the study.

Analyses

To assess the effectiveness of the chosen mindfulness exercises against the control conditions at reducing stress in participants in an efficient manner, we conducted an adaptive sequential design. We sampled participants until a Bayes Factor threshold for compelling evidence was met or until a maximally feasible number of participants was collected (Schönbrodt et al., 2017). Unlike in the frequentist approach, the aim of this procedure was not the control of error rates but to gauge relative support in favor of one of two competing hypotheses, H_1 and H_0. In the frequentist approach, the practice of adjusting the sampling plan after peaking at the data would have increased the Type I error rate. In the Bayesian approach, the data collection could be stopped anytime without having any adverse consequences on the interpretation of the Bayes Factor (Rouder, 2014). We divided our analysis workflow into confirmatory and exploratory. The sequential testing was tied to the confirmatory analysis phase; the exploratory analyses had no role in the decisions related to the stopping rule.

Confirmatory analyses

As the aim of the present study was to assess the effectiveness of four brief stand-alone mindfulness interventions, we planned to carry out four independent comparisons with the active controls. We therefore carried out four independent-samples Bayesian *t*-tests to determine whether there was a difference between each mindfulness exercise and the active control condition. We employed a two-tailed test using a non-informative JZS Cauchy prior for the alternative hypothesis with a default *r*-scale of $\sqrt{2}/2$ (Rouder, Speckman, Sun, Morey, & Iverson 2009). To account for the hierarchical nature of the data, we compared the condition means using a Bayesian mixed-effects model that involved a random intercept for site and for the different stories used in the non-mindful active control condition.

We implemented the sequential design as adaptive, stopping collecting the data for a given test of a condition after obtaining a Bayes Factor (BF_{10}) of 10 in favor of H_1 or a Bayes Factor of 1/10 favoring H_0., and re-assigning the remaining participants into other conditions, where such a threshold had not been met. We chose a Bayes Factor of 10, because according to the classification of Lee and Wagenmakers (2013), it demarcates the threshold between moderate and strong evidence. We started monitoring the Bayes Factor when there were at least 100 participants in each group. We chose 100 somewhat arbitrarily, because with a smaller sample size, the probability of misleading evidence (i.e., a Bayes Factor trajectory that arrives at the "wrong" boundary, that is, at the H_0 boundary when H_1 is correct and the other way around) would have been higher. Here, using a sample size of 100 and a Bayes Factor of 10, we aimed to substantially decrease the probability of misleading evidence (Stefan et al., 2017).

We monitored the Bayes Factor continuously and recorded the interim analysis at least every three days to check whether the threshold would have been reached. If the threshold would have not been reached (BF_{10} of 10 or of 1/10) we would have stopped the data collection after that the maximum number of participants would have been reached. If for any of the four tests, one of these three conditions was not met (i.e., BF_{10} of 10 in favor of H_1; BF_{10} of 1/10 in favor of H_0 ; the maximum number of participants was reached), we would have terminated the data collection on June 30th, 2022.

When one of these three conditions was met, we stopped allocating participants to that particular experimental condition, and we allocated them to the remaining conditions for which this has not been the case yet. If the data collection ended before all the sites were able to collect at least 70 participants (because we would have reached a BF_{10} of 10 in favor of H_1 or a BF_{10} of 1/10 favoring H_0. for each condition test), we would have continued with the data collection by reopening all groups and randomly allocating participants across groups until each site would have collected at least 70 participants until June 30, 2022 (see also our pre-registration at https://osf.io/us5ae). Finally, we decided not to screen for and exclude outliers and we did not perform any (non-linear) transformations contingent on the observed data.

Exploratory analyses

After the data collection was concluded, we also carried out analyses exploring the effect of the experimental conditions on pleasure, arousal, and dominance and for the moderating effect of neuroticism. The sequential testing was tied to the confirmatory analysis phase while the exploratory analyses had no role in our sampling decisions. We performed three different Bayesian *t*-tests for each dimension of the Self-Assessment Manikin scale (pleasure, arousal, and dominance) comparing our experimental conditions with the control condition. We then looked at the Bayes Factor to establish whether the data favored H_1 or H_0. As we did for the confirmatory analyses, we compared the condition means using a Bayesian mixed-effects model with a random intercept for lab and for the different stories used in the non-mindful active control condition to account for the hierarchical nature of the data.

To check whether neuroticism moderated the effects of the four experimental conditions on stress, we compared the model with the interaction to the model with only the main effects (using the lmBF function) and we reported the corresponding BF. If the model with the interaction was preferred to the model with only the main effects of a BF of 10 or more, we regarded it as solid evidence of the moderation of neuroticism on stress.

Dealing with careless responders

We applied a set of rules to deal with careless or insufficient effort responders (C/IE; Currant, 2016), to reduce the random variance component in the data. First, as one of the proactive measures, we forced the answers for the questions connected to our exclusion criteria (meditation experience, English level, and mental illnesses). For the questionnaires related to our dependent variables/moderator instead we alerted respondents about unanswered questions, but they had the possibility to continue with the survey without providing a response. Second, the programmed survey prevented participants from skipping listening to the 15 minutes audio file (for both mindfulness exercises and control conditions) by blocking the screen with the audio of the meditation/control condition for 14 minutes, so as not to allow participants to proceed to the following survey page until the meditation was finished. Third and finally, we identified and excluded participants who provided identical responses to a long series of items (i.e., always selecting the answer "strongly agree") by performing a long-string analysis. Using long-string analyses, we excluded participants with a string of consistent responses equal or greater than 10 (i.e., half of the scale length; see Huang et al., 2012; $N = 0$).

Ethics information, administrative organization, and data protection

The study first received ethical approval from Swansea University while the sites that participated in the data collection either received ethical approval from their local IRBs or stated that they were exempt. Each site's IRB protocols with ethics details and acceptance of each protocol can be found on the OSF project page at https://osf.io/6w2zm/.

Swansea University and Université Grenoble Alpes carried the administrative organization for the study. Swansea University was also the data controller for this project. Personal data of participants were processed for the purposes outlined in the information sheet (see the document Information Sheet at https://osf.io/xuznc/). Standard ethical procedures involved participants providing their consent to participate in this study by completing the consent form that was administered at the beginning of the online survey used for the experiment.

Results

Confirmatory Analyses

Development of the sequential design and closing of conditions.

On our RPubs page (https://rpubs.com/ale-sparacio92) we documented the development of the Bayes Factor comparing each condition against the active control conditions on the main dependent variable, state-stress assessed by the State-Trait Anxiety Inventory, Form Y. After the beginning of the data collection, we checked the Bayes Factor continuously as soon as the minimum number of participants per condition ($N = 100$) was reached, which happened on April 4, 2022[24]. Already upon first inspection (total sample size $N = 571$), Mindful Walking favored H_0 with a $BF_{10} = .01$, after which we closed the allocations of participants to this group. The following day (on April 5, 2022) one group (Body Scan) reached the threshold of compelling evidence with a $BF_{10} = 14.55$, while another group (Loving-Kindness) reached the threshold of non-compelling evidence with a $BF_{10} =.07$. As a result, we then closed allocation of participants to these two groups. On April 7, 2022, the last group (i.e., Mindful Breathing) reached a $BF_{10} =.02$, leading us to close the allocation of participants to that group. At that stage, Body Scan was the only condition that surpassed a $BF_{10} = 10$ in favor of H_1 , and Mindful Breathing, Loving-Kindness, and Mindful Walking surpassed the threshold of $BF_{10} = 0.1$ in favor of H_0.

However, since most of the sites had not yet collected the minimum number of participants required (set to $N = 70$), following the pre-registration we re-opened data collection for all groups and we randomly assigned the remaining participants to each of the experimental and active control conditions. At this point we monitored the Bayes Factor every three days, until June 30th, when we closed the data collection.

[24] When we checked the Bayes Factor each group had at least 110 participants across each condition, which was slightly higher than the pre-registration threshold of 100. This occurred because data had to be entered into the analysis from a Qualtrics export and this made it more difficult to monitor the Bayes factor continuously.

Results for each self-administered mindfulness condition as compared to the active control conditions for self-reported stress

We recoded the reverse items and then averaged the scores for the 20 state-focused items of the State-Trait Anxiety Inventory, Form Y. We used four Bayesian mixed-effects models to determine whether there was a difference between each mindfulness exercise taken individually and the active control conditions. For these Bayesian mixed-effect models, we included a random intercept for site and for different stories used in the non-mindful active control condition to account for the hierarchical nature of the data.

At the end of the data collection the group with the highest Bayes Factor was Body Scan with a Bayes Factor of 3.69×10^{11}, indicating that the observed data are 3.69×10^{11} times more likely to occur under H_1 than under H_0 (denoting "extreme evidence"; Lee & Wagenmakers, 2013), meaning that Body Scan reduces self-reported stress as compared to the active control condition (see Table 2 for means and standard deviations). The observed effect thus went in the expected direction.

For what concerns the other conditions, they all surpassed the threshold of compelling evidence of 10 in favor of H_1 . Table 7 summarizes the results for all the different mindfulness conditions in comparison to the active control conditions.

Condition	*N*	*M*	*SD*	BF_{10}
Active Control Conditions	478	1.95	0.50	/
Body Scan	449	1.68	0.46	3.7×10^{11}
Mindful Breathing	469	1.73	0.50	2.3×10^{5}
Loving-Kindness	427	1.70	0.49	1.1×10^{7}
Mindful Walking	416	1.73	0.46	4.8×10^{2}

Table 7: Results of the four independent comparisons with the active controls for the State-Trait Anxiety Inventory, Form Y

In summary, the Bayesian mixed-effects models provided strong evidence that all four mindfulness conditions were effective in slightly reducing self-reported stress levels of participants as compared to the active control conditions. From the control condition (M = 1.95) to the lowest-scoring condition, Body Scan (M = 1.68), there was a mean difference of 0.27 on a five-point scale, thus slightly decreasing people's stress levels.

Why did the Bayesian sequential result provide a different conclusion compared to the overall analysis?

According to our Bayesian sequential analysis plan we pre-registered that we would have stopped the allocation to participants to a group when, upon reaching a minimum number of participants across conditions, we would have reached a $BF_{10} = 10$ in favor of H_1 or a $BF_{10} = 0.1$ in favor or H_0. Body Scan reached a $BF_{10} = 10$, while the other three conditions reached a $BF_{10} = 0.1$. When these conditions were met for each condition, following our pre-registration we re-opened all the groups, and we randomly assigned the remaining participants until all the sites would have collected at least 70 participants. Some of the sites still failed to collect at least 70 valid observations as data collection stopped on June 30th, 2022. When we used the Bayesian mixed effect models with the total sample (n = 2,239) each condition test surpassed the threshold of a $BF_{10} = 10$, contradicting the results of the sequential analysis plan. How can this difference be explained?

According to Stefan et al. (2018), the probability of misleading evidence (i.e., when a Bayes Factor's trajectory exceeds the wrong boundary) can be influenced by several factors: The sample size, the effect size, the choice of the boundaries, and the choice of the priors. Although we tried to lower the chance to obtain misleading evidence by setting high Bayes Factor boundaries (we collected data until the Bayes factor was larger than 10 or smaller than 1/10) a small effect might have increased the chances of misleading evidence. For instance, when the true effect is $\delta = 0.35$, and when a non-informative prior is used, to obtain a Bayes

Factor larger than 10, with a probability of p = 0.8, it is necessary to allocate 220 participants to each condition (Stefan et al., 2018). The fluctuations of the mindfulness conditions from a $BF_{10} < 1/10$ to a $BF_{10} > 10$ are probably due to randomness that is determined by the fragility of the effect. For each of these conditions the effect began to stabilize when at least 200 participants were assigned to each condition[25]. In Figure 9 we plotted how the Bayes Factor, reported on a natural logarithmic scale, varied for each mindfulness condition as a function of the increasing sample size.

[25] This condition was not met for body scan, where the effect started to be stable before having 200 participants for conditions. However, it is likely that this difference was caused by randomness.

Figure 9: Evolution of the Bayes Factor, reported on a natural logarithmic scale, for all mindfulness conditions. Red horizontal lines represent the threshold of $BF_{10} = 10$ and of $BF_{10} = 1/10$ respectively

Exploratory analyses

Does neuroticism moderate the effectiveness of self-administered mindfulness exercises?

Some studies found that neuroticism moderates the psychological effects of mindfulness training (de Vibe et al., 2015, Tang & Braver, 2020). For this reason, we investigated whether neuroticism moderated the relationship between mindfulness exercises and stress using the 20-item Neuroticism subscale of the IPIP - 5 NEO domains. We merged the different mindfulness conditions and compared the merged conditions to the active control conditions to achieve higher power. We failed to find any evidence for the moderation of neuroticism as the ratio between the two models (the full model and the one with only the interaction) yielded an inconclusive Bayes Factor ($BF_{10} = 0.11$).

Does the English-level moderate the effectiveness of self-administered mindfulness exercises?

Language plays a role in the acquisition of knowledge to make meaning of emotional experiences and perceptions (Lindquist et al., 2015). If certain levels of knowledge of a particular language are not reached, the processes of making meaning out of emotional experiences could be compromised. Thus, we investigated whether the participants' English level moderated the effect of the mindfulness exercises on stress. Of the total 2,239 participants included in the analyses, 647 were non-native English speakers at least C1/C2 level, while 1,592 were native English speakers. We again merged the mindfulness conditions into a single group to increase statistical power. We did not find evidence for an interaction effect ($BF_{10} = 0.05$).

Cohen's ds for each condition compared to the active control condition.

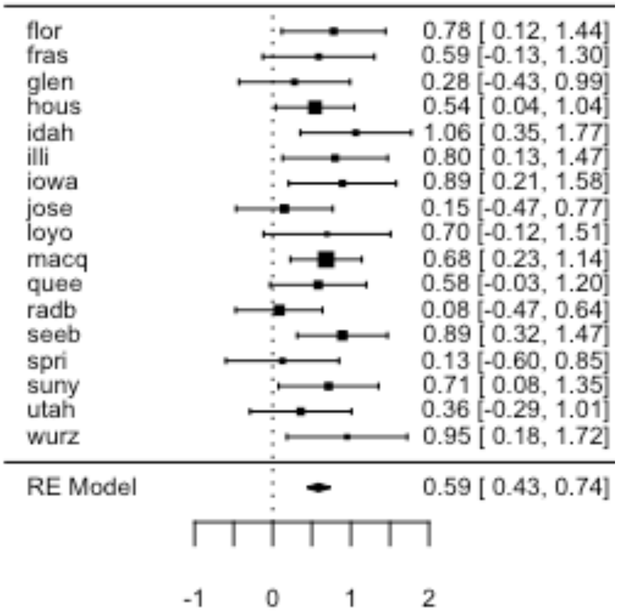
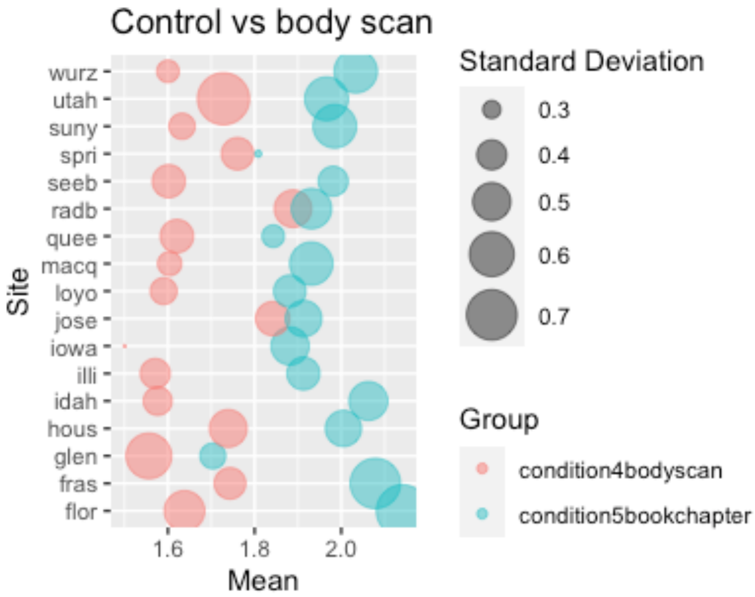
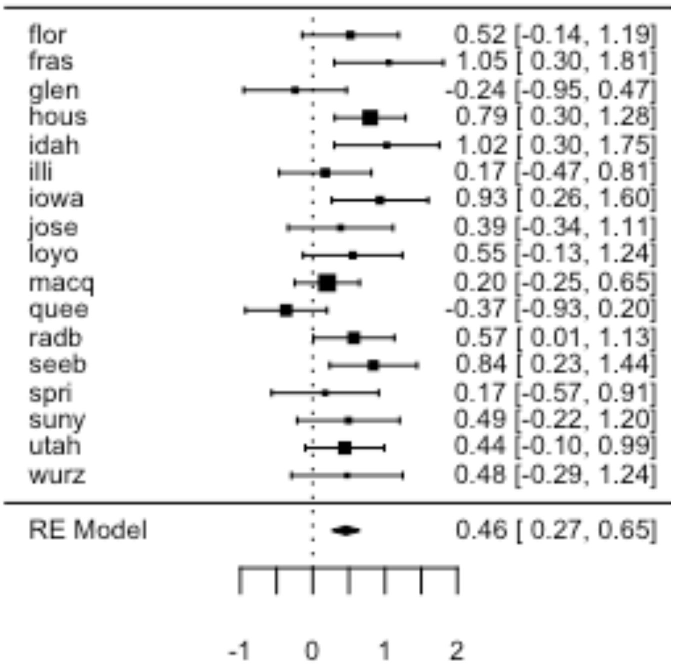
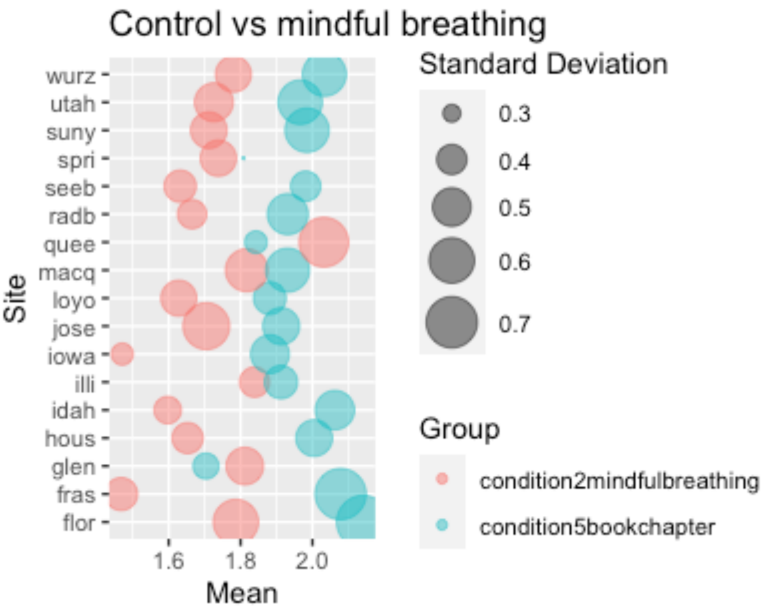
We calculated Cohen's *d* for each condition test using the escalc function of the *metafor* package using sample means and sample *SD*s. Even if we relied on a Bayesian framework, we used Cohen’s *d* as an estimate of the magnitude of the effect because Cohen’s *d* can be interpreted as the standard mean difference between two independent samples. Table 8 summarizes the effect sizes for all the conditions when compared to the control condition.

Condition test (against control)	Cohens’ *d* [95% *CI*]	*SE*
Body Scan	- .56 [.43, .69]	.07
Mindful Breathing	-.46 [.30, .61]	.08
Loving-Kindness	-.48 [.35, .62]	.07
Mindful Walking	-.45 [.32, .59]	.07

Table 8: Effect sizes for all the condition test (against control).

Heterogeneity per site.

To investigate the differences in the results between the sequential design and the result of the overall analysis we computed a series of analyses to try to explain effect size variance per site. However, for each of the mindfulness exercises we did not detect significant heterogeneity. Forest plots and means and *SD*s of each condition across sites are displayed in figure 10. Additionally, in table 9 we reported the heterogeneity values for each mindfulness condition across sites.


flor 0.78 [0.12, 1.44]
fras 0.59 [-0.13, 1.30]
glen 0.28 [-0.43, 0.99]
hous 0.54 [0.04, 1.04]
idah 1.06 [0.35, 1.77]
illi 0.80 [0.13, 1.47]
iowa 0.89 [0.21, 1.58]
jose 0.15 [-0.47, 0.77]
loyo 0.70 [-0.12, 1.51]
macq 0.68 [0.23, 1.14]
quee 0.58 [-0.03, 1.20]
radb 0.08 [-0.47, 0.64]
seeb 0.89 [0.32, 1.47]
spri 0.13 [-0.60, 0.85]
suny 0.71 [0.08, 1.35]
utah 0.36 [-0.29, 1.01]
wurz 0.95 [0.18, 1.72]
RE Model 0.59 [0.43, 0.74]
-1 0 1 2

Control vs body scan
wurz
utah
suny
spri
seeb
radb
quee
macq
loyo
jose
iowa
illi
idah
hous
glen
fras
flor
Site
1.6 1.8 2.0
Mean
Standard Deviation
0.3
0.4
0.5
0.6
0.7
Group
condition4bodyscan
condition5bookchapter

flor 0.52 [-0.14, 1.19]
fras 1.05 [0.30, 1.81]
glen -0.24 [-0.95, 0.47]
hous 0.79 [0.30, 1.28]
idah 1.02 [0.30, 1.75]
illi 0.17 [-0.47, 0.81]
iowa 0.93 [0.26, 1.60]
jose 0.39 [-0.34, 1.11]
loyo 0.55 [-0.13, 1.24]
macq 0.20 [-0.25, 0.65]
quee -0.37 [-0.93, 0.20]
radb 0.57 [0.01, 1.13]
seeb 0.84 [0.23, 1.44]
spri 0.17 [-0.57, 0.91]
suny 0.49 [-0.22, 1.20]
utah 0.44 [-0.10, 0.99]
wurz 0.48 [-0.29, 1.24]
RE Model 0.46 [0.27, 0.65]
-1 0 1 2

Control vs mindful breathing
wurz
utah
suny
spri
seeb
radb
quee
macq
loyo
jose
iowa
illi
idah
hous
glen
fras
flor
Site
1.6 1.8 2.0
Mean
Standard Deviation
0.3
0.4
0.5
0.6
0.7
Group
condition2mindfulbreathing
condition5bookchapter

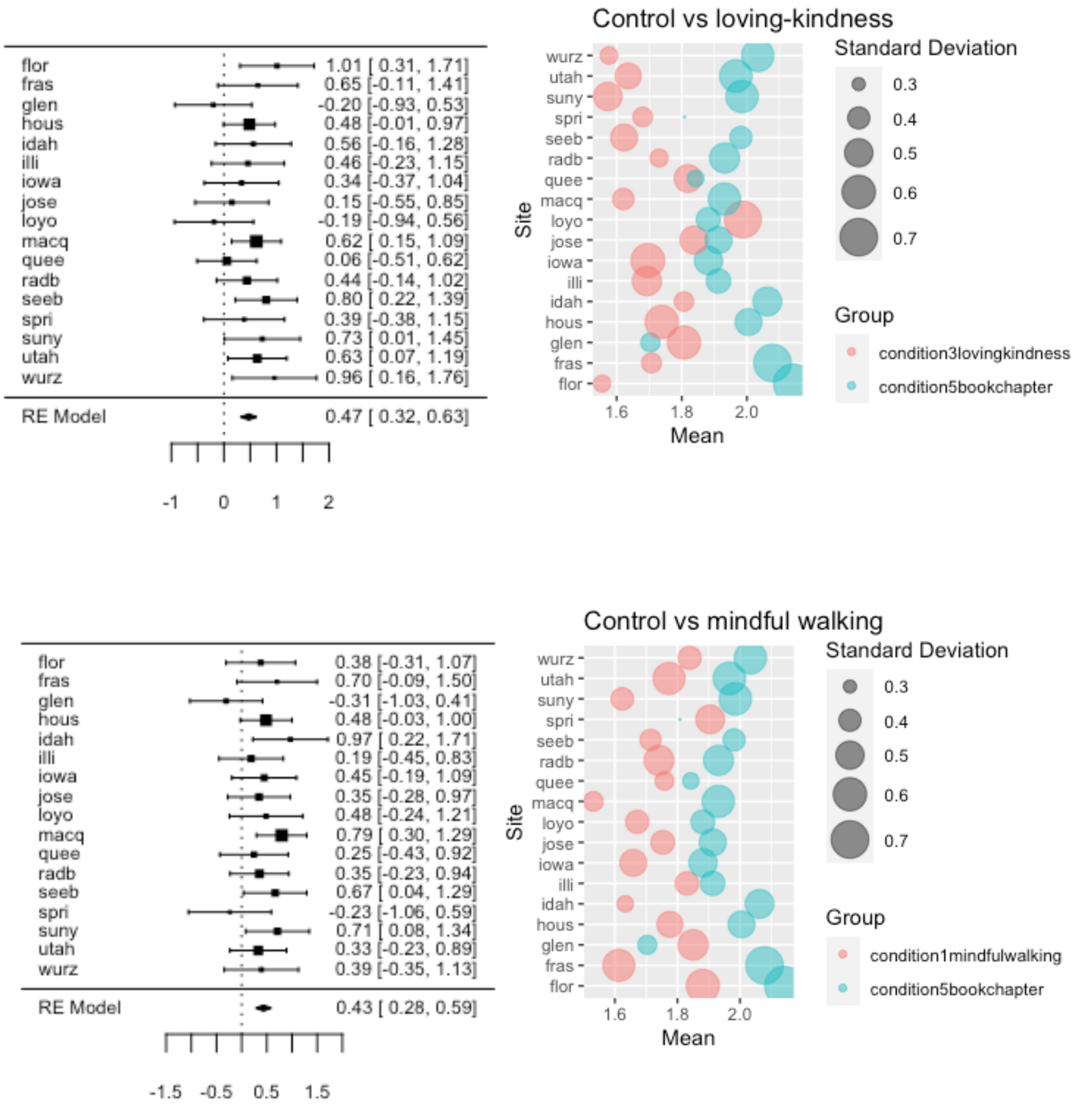


Figure 10: Forest plot, means and *SD*s for stress for each mindfulness exercise compared to the active control conditions.

Condition	Cochran's Q test (p-value)	τ	I^2
Body Scan	.83	0	0%
Mindful Breathing	.17	.21	24.24%
Loving-Kindness	.50	0	0%
Mindful Walking	.67	0	0%

Table 9: Heterogeneity values for each mindfulness condition across sites.

Self-assessment Manikin scale

We explored the effects of the mindfulness exercises on the dimensions of pleasure, arousal, and dominance as compared to the active control conditions again using four Bayesian mixed-effects models. We found that only for the dimension of pleasure and only for Mindful Breathing the Bayes Factor favored H_1, surpassing the threshold we set (BF_{10} = 16.1) meaning that participants felt more pleasant during the Mindful Breathing as compared to participants that listened to the story in the active control condition. Table 10 summarizes the Bayes Factor we found for the three dimensions of the Self-assessment Manikin scale for each condition test.

Condition test (against control)	pleasure	arousal	dominance
Body Scan	2.95	5.0×10^{-8}	0.43
Mindful Breathing	16.1	1.6×10^{-7}	9.7×10^{-5}
Loving-Kindness	0.01	7.1×10^{-8}	2.3×10^{-5}
Mindful Walking	0.30	1.2×10^{-7}	3.9×10^{-4}

Table 10: Bayes Factors for the four independent comparisons with the active controls for the Self-assessment Manikin scale

Discussion

In the current multi-site project, we investigated whether four different mindfulness exercises taken singularly, were effective in reducing participants' stress levels as compared to three active control conditions. We found that all the four mindfulness exercises (i.e., Body Scan, Mindful Breathing, Mindful Walking, Loving-Kindness) slightly decreased self-reported stress of participants as compared to the listening of three different stories randomly sampled.

We describe the advantages and improvements of our multi-site study in relation to the existing research on self-administered mindfulness in the following section. We then highlight potential shortcomings of the present project, and we provide suggestions on how to further improve it. Finally, we provide constraints of generality for the current project to define under which circumstances our conclusions are valid.

Improvements to the literature

A previous meta-analysis failed to find evidence for the reduction of participants' stress levels for self-administered mindfulness interventions mostly due to the low-quality of the included studies (Sparacio et al., 2022). For instance, some of the included studies lacked sufficient power (Economides et al., 2018; Rich et al., 2021), others were based solely on a passive, rather than an active, control condition (Rich et al., 2021), still others did not adhere to open science practices as they were not based on protocols or pre-registered (e.g., Lee & Jung, 2018; Lundqvist et al., 2018).

In particular, the absence of a standardized protocol for self-administered mindfulness is particularly concerning. Each study on self-administered mindfulness relies on a different protocol that varies in length, duration, and types of self-administered mindfulness exercises (Sparacio et al., 2022). Currently, for these interventions there is not a well-established protocol like in other conventional mindfulness programs like mindfulness-based stress reduction (MBSR; Kabat-Zinn, 2008) or mindfulness-based cognitive therapy (MBCT; Segal et al., 2002).

Our multi-site design attempts to provide solutions for these shortcomings. Indeed, compared with previous mindfulness studies, our multi-site design is 1) adequately powered to find the effect of each mindfulness exercise as compared to the control conditions, 2) compares each mindfulness condition with an active control group (and not a passive control group or a waiting list), and was 3) pre-registered. In addition, the results of this multi-site study can serve as a basis for building protocols on self-administered mindfulness. In fact, we determined that the four mindfulness exercises included in our study are effective in reducing stress levels and thus are good candidates for creating protocols based on them. In the next section we highlight potential shortcomings of the present project that may cast doubts on the effectiveness of our multi-site study, and we further provide suggestions on how to improve it.

Shortcomings of the current multi-site project

A limitation of the current multi-site project is that the effect we found for each mindfulness exercise was small, furthermore we relied only on self-report assessments to measure participants' stress levels. Self-reported measures have some limitations that indirect measurements do not have. Indeed participants that answer to a questionnaire may be imprecise in self-reported assessment (Devaux & Sassi, 2016), may lack introspective ability leading to biased estimates about their levels of stress (Levin-Aspenson, & Watson, 2018) and may be subjected by the demand effect (i.e., participants are more likely to confirm the hypothesis of the positive effect of mindfulness, than reinterpreting information that would counter the hypothesis they hold; Nichols, & Maner 2008; Weber, & Cook, 1972).

However, given the lack of evidence behind the efficacy of prior studies on self-administered mindfulness, this project is an important step towards the utility of the intervention.

A possible way to solve these shortcomings is to measure stress not only with self-report measures, but also with physiological assessments of the autonomic nervous system (e.g., assessment of catecholamines, assessment of the ANS via skin conductance, cortisol, heart rate, systolic and diastolic blood pressure; Bally et al., 2003; Berntson et al., 1993). Thus, future studies investigating the efficacy of single brief self-administered mindfulness exercises should include both psychological and physiological measures to have more reliable estimates of stress levels and to rule out the possibility of a demand effect.

Constraints on generality (Simons et al., 2017)

Based on the results of this multi-site study, brief stand-alone mindfulness exercises have an efficacy in reducing stress levels of individuals as compared to an active control condition. However, this applies only for particular contextualized situations and for some types of populations. In relation to the first constraint the results we obtained relate to short-term stress, and the effect sizes we found were small. We are not aware of the effect of these exercises when the individual is exposed to high stressful situations (e.g., living in an area exposed to a war). In relation to the second constraint the findings of our study apply to: 1) Participants that are older than 18, 2) participants that are fluent/native English speakers coming from Australia, Europe, UK, and US, 3) participants that are non-meditators, and finally 4) participants that did not have a history of mental illnesses. For instance, we do not know whether our findings are the same for other types of populations that have different characteristics. To improve generalizability of the current findings future studies should try to replicate our results by including meditators in the sample and a population of non-English speakers which also comes from different geographical areas of the ones included in the current project.

Conclusions

In the current multi-site study 37 sites were involved in the data collection and gathered, all together, 2239 valid observations. This is the first large scale project investigating the efficacy of single-brief mindfulness interventions. We found that each mindfulness condition that we included in the study (i.e., Body Scan, Mindful Breathing, Mindful Walking, Loving-Kindness) was slightly more effective in reducing self-reported levels of stress of participants as compared to an active control condition. These interventions should be intended as being effective in the short-term and are unlikely to affect dispositional traits (such as chronic stress). Although we found an effect for single brief mindfulness exercises our multi-site study carries the limitations of using only self-report measures. Well-powered studies and with physiological assessment of the autonomic nervous system are warranted to corroborate the results of the current multi-site project.

Data availability

To ensure methodological rigor and transparency, we have made our data and the script available on the Open Science Framework (https://osf.io/6w2zm/) and via the Github repository (https://github.com/alessandro992/A-large-multi-site-test-of-self-administered-mindfulness)

Declaration of Competing Interest

The certified mindfulness instructor that collaborated on the multi-site project only recorded the meditations, but was not involved in the writing of any parts of the manuscript (e.g., choice of the analyses). We do not have further conflict of interests to declare.

Conceptualization: Alessandro Sparacio, Hans IJzerman, and Gabriela Jiga-Boy.

Data curation: Alessandro Sparacio, Ivan Ropovik, and Filippo Giorgini.

Formal analysis: Alessandro Sparacio and Ivan Ropovik.

Funding acquisition: Hans IJzerman, Ivan Ropovik, and Gabriela Jiga-Boy.

Investigation: Bert N. Uchino, Joshua Landvatter, Tracey Tacana, Sandra J. Diller, Jaye L. Derrick, Joahana Segundo, Jace D. Pierce, Robert M. Ross, Zoë Francis, Amanda LaBoucane, Christine Ma-Kellams, Maire B. Ford, Kathleen Schmidt, Celia C. Wong, Wendy C. Higgins, Bryant M. Stone, Samantha K. Stanley, Gianni Ribeiro, Paul T. Fuglestad, Valerie Jaklin, Andrea Kübler, Philipp Ziebell, Jewell L. Crystal, Yulia Kovas, Mahnoosh Allahghadri, Michael F. Baranski, Hannah Burgess, Annika B. Benz, Maysa DeSousa, Catherine E. Nylin, Janae C. Brooks, Caitlyn M. Goldsmith, Jessica M. Benson, Siobhán M. Griffin, Stephen Dunne, William E. Davis, Tam J. Watermeyer, William B. Meese, Jennifer L. Howell, Laurel Standiford Reyes, Megan G. Strickland, Sally S. Dickerson, Samantha Pescatore, Shayna Skakoon-Sparling, Zachary I. Wunder, Martin V. Day, Shawna Brenton, Audrey H. Linden, Christopher E. Hawk, Léan V. O'Brien, Tenzin Urgyen, and Jennifer S. McDonald.

Methodology: Alessandro Sparacio, Hans IJzerman, and Ivan Ropovik.

Project administration: Alessandro Sparacio, Hans IJzerman, and Gabriela Jiga-Boy.

Resources: Alessandro Sparacio, Hans IJzerman, Christoph Spiessens, and Gabriela Jiga-Boy.

Software: Alessandro Sparacio.

Supervision: Alessandro Sparacio, Hans IJzerman, Ivan Ropovik, and Gabriela Jiga-Boy.

Validation: Alessandro Sparacio, Hans IJzerman, and Ivan Ropovik.

Visualization: Alessandro Sparacio.

Writing - original draft: Alessandro Sparacio, Hans IJzerman, Ivan Ropovik, and Gabriela Jiga-Boy.

Writing - review & editing: Alessandro Sparacio, Hans IJzerman, Ivan Ropovik, and Gabriela Jiga-Boy.

Chapter 5: Discussion

The goal of the present Ph.D. project was to shed light on stress regulation processes, by examining the evidence behind various stress regulation practices to better explain when and why these practices reduced stress. To investigate which strategy effectively reduce stress we conducted a Registered Report meta-analysis on two external regulation strategies (i.e., being in nature and emotional social support) a pre-registered meta-analysis on two internal regulation strategies (i.e., self-administered mindfulness and biofeedback), and a multi-site study experimentally examining the effect of one of the strategies (self-administered mindfulness) on stress regulation.

The aim of the two meta-analyses was to evaluate the current state of the art of stress regulation research and to offer suggestions for how to improve the field.
The two meta-analyses had common goals: 1) For the four strategies, to evaluate which components of the stress response (i.e., affective, physiological, cognitive), as well as affective consequences of stress, had sufficient empirical support. 2) To apply publication bias correction techniques to have more accurate estimates of the *ES* linked to each strategy. 3) To investigate whether personality traits were used in moderator analyses in studies on stress regulation.

The aim of the multi-site study was to provide high-quality evidence on the foundational claims of self-administered mindfulness, testing four different self-administered mindfulness exercises in a population of non-meditators to examine which of these exercises were effective in reducing individuals' stress levels, compared to three, randomly sampled, non-mindful active control conditions (we created multiple control conditions to be less dependent on a single stimulus).

In what follows we summarize the results of our empirical studies trying to provide an answer to the goals mentioned above for each strategy taken separately emphasizing strengths and weaknesses of our findings. Furthermore, we provide guidelines for building protocols for each strategy, and we provide suggestions for building a stress management intervention based on the results of our empirical studies. Finally, we define the limits of stress regulation strategies in situations in which structural changes are necessary or that require the intervention of a specialist (e.g., psychiatrist or therapist).

External regulations strategies (included in the Registered Report meta-analysis)

In the following paragraphs we summarize what we found for the two external regulation strategies mentioning the goals we set for this meta-analysis.

Being in nature

Being in nature was the only strategy (out of the four) for which, according to our inferential criteria, we could find empirical evidence. Indeed, we found a naive meta-analytic effect of Hedges' $g = -.42$, that held after we applied our main publication bias technique (4PSM's Hedges' $g = -.60$; $p = .006$). For the components of the stress response, for which we had a sufficient number of effects, we found small naïve meta-analytic estimates (for the physiological and the affective component, Hedges' $g = -.31$ and Hedges' $g = -.49$ respectively), while for the affective consequences of stress we had too few effects to calculate an estimate. We found no studies in which personality traits were used as moderators.

Although we found that interventions that involved being in nature reduces stress, the literature on the named strategy has some limitations: When we carried out an assessment of the the risk of bias with the RoB 2 (Sterne et al., 2019), 25% of the studies had a high risk of bias, while the 50% of them had a moderate risk. When we examined the tests to assess whether means or *SD*s were mathematically impossible we found that for being in nature the 37% of the total effects had inconsistent means or *SD*s. Finally, when we screened the papers for inconsistencies in the reported statistics, we found that for being in nature the 96% of the effects were reported correctly, but only a minority of the studies was reported correctly in the *APA* format (47%). In conclusion, although there were statistical issues in the literature on being in nature, these issues did not significantly affect the outcome of our meta-analysis. We thus rely on our inferential criteria according to which, being in nature does, in fact, lessen stress.

Emotional social support

Emotional social support was the strategy that according to the literature has the most empirical evidence (e.g., Cohen, 2004; Lakey & Cronin, 2008). Although we found a signal (Hedges' $g = -.14$), the signal was not present anymore when we applied publication bias techniques (4PSM's Hedges' $g = -.15$; $p = .15$). We found a naïve meta-analytic estimate for the physiological (Hedges'$g = -.26$), and affective (Hedges'$g = -.11$) components. No studies were found for the cognitive component, and we had insufficient data to calculate an estimate of the affective consequences of stress. For emotional social support, only in one study a personality trait was used as a moderator in the focal analyses.

For what concerns the quality of the literature the situation does not differ much from that of being in nature: All studies except one scored high on the risk of bias assessment. No studies had inconsistent means or *SD*s, however the 15% of the effects had some inconsistencies in the reported statistics with only the 23% of the studies reported statistics correctly in the *APA* format. We conclude that emotional social support, based on our findings, cannot be considered an effective strategy for the downregulation of stress levels.

Internal regulations strategies (included in the pre-registered meta-analysis)

In the following paragraphs we present the results of the two internal regulation strategies that were part of the pre-registered meta-analysis.

Biofeedback

For heart-rate variability biofeedback we found an effect of Hedges' g = .59, however the publication bias correction techniques failed to converge on the presence of an effect. The majority of the effects was categorized as failing into the physiological component (k = 11; Hedges'g = -.51), followed by the affective component (k = 10; Hedges'g = .61). We did not analyze results for what concerns the cognitive component and the affective consequences of stress because we had too few effects (respectively k = 1, and k = 7 for the cognitive component and for the affective consequences of stress).

In 17 studies out of 20 individual differences were assessed, however they were used solely as exclusion criteria or as baseline measurements but did not enter in the main analyses. For what concerns the quality assessment of the literature the situation is concerning: Thirty-nine percent of the included studies had some risk of bias, while the 24% scored high in the risk of bias assessment, additionally, the 26% of the coded effects were based on inconsistent *SD*s or means. Overall, based on our inferential criteria, we cannot conclude that heart rate variability biofeedback demonstrated sufficient empirical evidence for the reduction of stress.

Self-administered mindfulness

We found a signal for self-administered mindfulness (Hedges' g = .37) that was not detectable anymore after applying publication bias correction techniques. We found a naive meta-analytic effect for the affective component (k = 26, Hedges' g = .39), for the physiological, the cognitive component and the affective consequences of stress too few effects were collected, thus we did not analyze the results. Finally, as regards the quality of the included studies a sizable portion of the studies had some risk of bias (76%) with a small portion had high risk (9%), however no studies were excluded because of inconsistencies in the means or *SD*s of the reported studies.

While we could not find an effect for self-administered mindfulness in our meta-analysis, in our multi-site study we found evidence for what concerns the reduction of stress levels of single-brief self-administered mindfulness exercises. In what follows, we outline the results of our multi-site study.

Multi-site study on self-administered mindfulness

In our multi-site study, we found that four self-administered mindfulness exercises (Body Scan, Mindful Breathing, Mindful Walking, and Loving-Kindness) reduced participants' self-reported stress when compared to three non-mindful active control conditions (i.e., the listening of three different stories chosen at random). The present project improved some shortcomings of the literature as studies on self-administered mindfulness tended to be underpowered (e.g., Economides et al., 2018; Rich et al., 2021), were based on passive or waitlist control groups (e.g., Rich et al., 2021) lacked pre-registrations or were not based on a protocol (e.g., Lee & Jung, 2018; Lundqvist et al., 2018).

Our project instead had sufficient power to find the effect of interest, it was based on three active control conditions, and it was pre-registered on the OSF (https://osf.io/6w2zm/). Although we could not disentangle the possibility of a demand effect, and we did not use physiological measures for the assessment of stress levels, our multi-site study is the first sufficiently powered study that tests the effectiveness of single brief stand-alone mindfulness exercises on stress reduction.

Four strategies, but a plethora of protocols

In addition to showing that the quality of the literature was poor across the different strategies, the two meta-analyses have shown that there was no convergence regarding the choice of a shared protocol either. Indeed, for emotional social support, being almost all the included studies observational, we had no indications on how a protocol should be built. On the other hand, for being in nature, each study was based on a different type of intervention, as regards the type of natural environment in which individuals do the experiment (e.g., park, forest), the type of activity that individuals performed during the trial (e.g., nature walking or nature seeing) and how individuals' stress levels were measured (e.g., with self-reported measures or with physiological assessments).

Similarly, for self-administered mindfulness and biofeedback, we have seen substantial heterogeneity across studies. The number of intervention sessions for self-administered mindfulness in the included studies ranged from 1 to 64 (M = 29.3; Mdn = 27). For biofeedback the numbers of trials varied from 1 to 40 (M = 9.5; Mdn = 6). The amount of time participants spent meditating or receiving a biofeedback intervention also changed significantly: Self-administered mindfulness included both short interventions (lasting 5 minutes) and lengthier ones (lasting a total of 16 hours). Similarly for biofeedback the shortest intervention lasted 5 minutes, with the longest arrived at a total of 22 hours.

Because of the high heterogeneity which characterizes the literature of the four stress regulation strategies, we have tried to work on a protocol that could provide guidelines for future studies on the named strategies. In each protocol we specified the information that should be recorded and the scales that should be administered to participants. We hope these protocols will provide a red thread for researchers who want to do research on these stress regulation strategies. (See Appendices G, H & R).

Towards the building of a stress management intervention

The two meta-analyses we conducted, showed that being in nature is the only one of the four strategies we have analyzed, which, based on our inference criteria, has been shown to be effective in reducing stress levels. The multi-site study instead demonstrated how single self-administered mindfulness exercises could be effective in reducing individuals' stress levels. It is therefore reasonable to start from these two strategies to build stress management interventions. We found that the effect of being in nature held regardless of the type of exposure, meaning that individuals benefit from seeing a natural environment even in video or with the help of a virtual reality simulation. The use of a virtual medium to exert the effects of being in nature on individuals allows flexibility as it makes this strategy more affordable for those who do not live near a forest or a park. In addition, the effect of being in nature can easily be combined with mindfulness to have cumulative and synergistic effects.

For instance, in a meta-analysis Schutte & Malou (2018) found a relationship between the sense of being in tune with nature and the ability to be mindful (i.e., trait mindfulness). Furthermore, according to Kaplan (2001), practicing meditation and spending time in nature together had beneficial effects that go beyond their individual effects alone. This result appears to have been evidenced by a meta-analysis according to which nature-based mindfulness is moderately superior to mindfulness conducted in non-natural settings for what concerns psychological and physiological outcomes (Djernis et al., 2019).

The exercises that we included in our multi-site study can be easily practiced in a natural environment as well. For instance, individuals can practice Body Scan while sitting in a chair in a park or can practice Mindful Walking while walking inside a forest, paying attention not only to the walk, but also to the trees, bushes, and other natural elements that are around. The ease with which mindfulness can be combined with being in nature lays the groundwork for future studies that can combine these two interventions to reduce people's stress levels. In the next paragraph we underline the limitations of these stress regulation strategies in particular cases in which structural changes should be implemented to reduce people's stress levels.

Limits of stress regulation strategies

Although strategies for reducing stress, like those included in the current Ph.D. project, may be useful for certain people, they might not be enough to decrease stress levels in some particular situations. For instance, mindfulness can reduce stress levels, but it won't suffice if a person is always stressed out at work because his coworkers don't appreciate them or because they are not comfortable at their job position. In a similar way, those who cannot pay their bills with their salary and make it to the end of the month will hardly be able to say that they have solved the problem after receiving encouragement and support from a friend. People in this condition are exposed to chronic stress and may develop allostatic load and all the negative consequences that come with it (e.g., depression, hippocampus impairment; Goldstein, & McEwen, 2002; McEwen 1988; Simmons et al., 1984). For those people, stress regulation interventions may be helpful, but structural changes are more urgent to reduce chronic stress (e.g., change jobs, find other sources of income that can allow the individual to pay all the expenses).

Furthermore, these strategies should not be considered a panacea for people suffering from severe psychopathological disorders (e.g., depression). For a clinical population, it is advisable to carefully select the interventions to be administered: Choosing the incorrect approach can have little impact or, worse, have negative impacts on stress regulation (IJzerman et al., 2020). For instance, anecdotally, mindfulness practices occasionally cause psychotic episodes, panic attacks, and depersonalization (Van Gordon et al., 2017). For these people it is recommended that they consult a mental health specialist.

There are therefore situations in which stress regulation strategies have limited effectiveness; in cases where the individual has to make structural changes to improve their living conditions and in cases where the individual suffers from certain psychopathologies that require the intervention of a clinical staff.

Conclusions

Regulating stress is vital: For this reason, the current Ph.D. project aimed to identify non-pharmacological and cheap stress regulation strategies that could effectively reduce stress of individuals. To reach this goal we conducted two meta-analyses on four stress regulation strategies (i.e., being in nature, emotional social support, self-administered mindfulness, biofeedback) and a multi-site study focused on one of the strategies introduced in the meta-analysis (i.e., self-administered mindfulness). The two meta-analyses pointed to being in nature as the sole reliable strategy for the downregulation of stress. However, the presence of publication bias, the high heterogeneity, and the low quality of the included studies decreased the strength of evidence of the results.

Although we could not detect a signal for self-administered mindfulness, the results of the multi-site study showed that four single self-administered mindfulness exercises effectively reduced stress compared with the three active control conditions. The inconsistency in the results between the multi-site study and the meta-analysis on self-administered mindfulness suggests that the lack of an effect could be due to the low quality of the studies on stress regulation. For this reason, we recommend that future studies be pre-registered or preferably conducted in the form of Registered Reports to improve the state of the literature on stress regulation.

References

Adam E. K. (2006). Transactions among adolescent trait and state emotion and diurnal and momentary cortisol activity in naturalistic settings. *Psychoneuroendocrinology*, *31*(5), 664–679.

Adam, E. K., & Kumari, M. (2009). Assessing salivary cortisol in large-scale, epidemiological research. *Psychoneuroendocrinology*, *34*(10), 1423–1436.

Aldwin, C. M. (2007). *Stress, coping, and development: An integrative perspective* (2nd ed.). Guilford Press.

Allen, L., Scott, J., Brand, A., Hlava, M., & Altman, M. (2014). Credit where credit is due. *Nature*, *508*, 312–313. https://doi.org/10.1038/508312a.

Alloy, L. B., Abramson, L. Y., Hogan, M. E., Whitehouse, W. G., Rose, D. T., Robinson, M. S., Kim, R. S., & Lapkin, J. B. (2000). The temple-wisconsin cognitive vulnerability to depression project: Lifetime history of axis I psychopathology in individuals at high and low cognitive risk for depression. *Journal of Abnormal Psychology*, *109*(3), 403–418.

American Psychological Association (2017). Stress in America: The state of our nation. Stress in America™ Survey. https://www.apa.org/news/press/releases/stress/2017/state-nation.pdf.

Anaya, J. (2016). The GRIMMER test: A method for testing the validity of reported measures of variability. *PeerJ Preprints,* https://doi.org/10.7287/peerj.preprints.2400v1.

Antonelli, M., Barbieri, G., & Donelli, D. (2019). Effects of forest bathing (shinrin-yoku) on levels of cortisol as a stress biomarker: A systematic review and meta-analysis. *International Journal of Biometeorology*, *63*(8), 1117–1134.

Aurelius, M., & Gardini, N. (2005) *Colloqui con se stesso*. Medusa.

Bally, K., Campbell, D., Chesnick, K., & Tranmer, J. E. (2003). Effects of patient-controlled music therapy during coronary angiography on procedural pain and anxiety distress syndrome. *Critical Care Nurse, 23*(2), 50–58.

Bartoš, F., Maier, M., Wagenmakers, E., Doucouliagos, H., & Stanley, T. D. (2021). No need to choose: Robust bayesian meta-analysis with competing publication bias adjustment methods. *PsyArxiv,* https://doi.org/10.31234/osf.io/kvsp7.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1 - 48.

Baujat, B., Mahé, C., Pignon, J.-P., & Hill, C. (2002). A graphical method for exploring heterogeneity in meta-analyses: Application to a meta-analysis of 65 trials. *Statistics in Medicine, 21*(18), 2641–2652.

Beckes, L., & Coan, J. A. (2011). Social baseline theory: The role of social proximity in emotion and economy of action. *Social and Personality Psychology Compass, 5*(12), 976–988.

Beil, K., & Hanes, D. (2013). The influence of urban natural and built environments on physiological and psychological measures of stress: A pilot study. *International Journal of Environmental Research and Public Health*, *10*(4), 1250–1267.

Berntson, G. G., Cacioppo, J. T., & Quigley, K. S. (1993). Cardiac psycho physiology and autonomic space in humans: Empirical perspectives and conceptual implications. *Psychological Bulletin*, *114*(2), 296–322.

Blandini, F., Martignoni, E., Sances, E., Bono, G., & Nappi, G. (1995). Combined response of plasma and platelet catecholamines to different types of short-term stress. *Life Sciences*, *56*(13), 1113–1120.

Bolger, N., DeLongis, A., Kessler, R. C., & Schilling, E. A. (1989). Effects of daily stress on negative mood. *Journal of Personality and Social Psychology*, *57*(5), 808–818.

Bowers, C. A., & Gesten, E. L. (1986). Social support as a buffer of anxiety: An experimental analogue. *American Journal of Community Psychology, 14*(4), 447–451.

Bratman, G. N., Hamilton, J. P., Hahn, K. S., Daily, G. C., & Gross, J. J. (2015). Nature experience reduces rumination and subgenual prefrontal cortex activation. *Proceedings of the National Academy of Sciences, 112*(28), 8567–8572.

Brown, N. J. L., & Heathers, J. A. J. (2016). The GRIM test. *Social Psychological and Personality Science, 8*(4), 363–369.

Brugha, T., Bebbington, P., MacCarthy, B., Potter, J., Sturt, E. & Wykes, T. (1987) Social networks, social support and the type of depressive illness. *Acta Psychiatrica Scandinavica, 76*(6), 664–673.

Burgard, T., Bosnjak, M., & Studtrucker, R. (2021). Towards cumulative evidence and reproducible meta-analyses. Introduction and demonstration of PsychOpen CAMA. ZPID (Leibniz Institute for Psychology). *Available via PsychArchives:* https://doi.org/10.23668/PSYCHARCHIVES.4809.

Cannon, W. B. (1915). *Bodily changes in pain, hunger, fear and rage: An account of recent researches into the function of emotional excitement.* D Appleton & Company.

Cannon, W.B. (1929). Organization for physiological homeostasis. *Physiological Reviews*, *9*(3), 399-431.

Cannon, W. B. (1932) *The wisdom of the body*. W W Norton & Co.

Cannon, W.B. (1935). Stresses and strains of homeostasis. *The American Journal of the Medical Sciences*, *189*, 13–14.

Carter, E. C., Schönbrodt, F. D., Gervais, W. M., & Hilgard, J. (2019). Correcting for bias in psychology: A comparison of meta-analytic methods. *Advances in Methods and Practices in Psychological Science, 2*(2), 115–144.

Cavanagh, K., Strauss, C., Cicconi, F., Griffiths, N., Wyper, A., & Jones, F. (2013). A randomized controlled trial of a brief online mindfulness-based intervention. *Behaviour Research and Therapy, 51*(9), 573–578.

Cavanagh, K., Strauss, C., Forder, L., & Jones, F. (2014). Can mindfulness and acceptance be learnt by self-help?: A systematic review and meta-analysis of mindfulness and acceptance-based self-help interventions. *Clinical Psychology Review*, *34*(2), 118–129.

Cavanagh, K., Churchard, A., O'Hanlon, P., Mundy, E., Votolato, P., Jones, F., Gu, J. & Strauss, C. (2018). A randomised controlled trial of a brief online mindfulness-based intervention in a non-clinical population: Replication and extension. *Mindfulness*, *9*(4), 1191–1205.

Chambers, C. D. (2013). Registered reports: A new publishing initiative at Cortex. *Cortex, 49,* 609-610.

Christley, R., M., (2010). Power and error: Increased risk of false positive results in underpowered studies. *The Open Epidemiology Journal*, *3*, 16-19.

Coan, J. A., Schaefer, H. S., & Davidson, R. J. (2006). Lending a hand: Social regulation of the neural response to threat. *Psychological Science*, *17*(12), 1032–1039.

Coan, J. A., & Sbarra, D. A. (2015). Social Baseline Theory: The social regulation of risk and effort. *Current Opinion in Psychology*, *1*, 87–91.

Coan, J. A., Beckes, L., Gonzalez, M. Z., Maresh, E. L., Brown, C. L., & Hasselmo, K. (2017). Relationship status and perceived support in the social regulation of neural responses to threat. *Social Cognitive and Affective Neuroscience*, *12*(10), 1574–1583.

Cohen, S., Doyle, W. J., Skoner, D. P., Rabin, B. S., & Gwaltney, J. M. Jr. (1997). Social ties and susceptibility to the common cold. *Journal of the American Medical Association*, *277*(24), 1940- 1944.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress.

Journal of Health and Social Behavior, *24*(4), 385–396.

Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin*, *98*(2), 310-357.

Cohen, S. (2004). Social relationships and health. *American Psychologist, 59*(8), 676–684.

Congedo, V. (2021). *Le guide di headspace: Una Serie Netflix alla Scoperta della Meditazione*. State of Mind. Retrieved August 15, 2022, from https://www.stateofmind.it/2021/04/headspace-meditazione-recensione/

Council of Europe (2001). Common European framework of reference for languages: Learning, teaching, assessment. Retrieved from https://www.coe.int/en/web/common-european-framework-reference-languages

Chambers, R., Gullone, E., & Allen, N. B. (2009). Mindful emotion regulation: An integrative review. *Clinical Psychology Review*, *29*(6), 560–572.

Chang, E. C. (1998). Dispositional optimism and primary and secondary appraisal of a stressor: Controlling for confounding influences and relations to coping and psychological and physical adjustment. *Journal of Personality and Social Psychology*, *74*(4), 1109–1120.

Chu, P. S., Saucier, D. A., & Hafner, E. (2010). Meta-analysis of the relationships between social support and well-being in children and adolescents. *Journal of Social and Clinical Psychology, 29*(6), 624–645.

Cicero, M. T., Reid, J. S., & Kelsey, F. W. (1949). *Laelius de Amicitia*. Allyn and Bacon.

Cooper, C. L., & Payne, R. (1991). *Personality and stress: individual differences in the stress process*. John Wiley & Sons.

Coles, N. A., March, D. S., Marmolejo-Ramos, F., Larsen, J. T., Arinze, N. C., Ndukaihe, I., Willis, M. L., Foroni, F., Reggev, N., Mokady, A., Forscher, P. S., Hunter, J. F., Kaminski, G., Yüvrük, E., Kapucu, A., Nagy, T., Hajdu, N., Tejada, J., Freitag, R., Zambrano, D., … Liuzza, M. T. (2022). A multi-lab test of the facial feedback hypothesis by the Many Smiles Collaboration. *Nature Human Behaviour*, 10.1038/s41562-022-01458-9. Advance online publication.

Corbin, T. J., Purtle, J., Rich, L. J., Rich, J. A., Adams, E. J., Yee, G., & Bloom, S. L. (2013). The prevalence of trauma and childhood adversity in an urban, hospital-based violence intervention program. *Journal of Health Care for the Poor and Underserved*, *24*(3), 1021–1030.

COVID-19 Mental Disorders Collaborators (2021). Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. *Lancet*, *398*(10312), 1700–1712.

Crailsheim, E., & Dolores, E. P.-G. M. *The representation of external threats: From de middle ages to the modern world*. Brill, 2019.

De Bruin, E. I., Topper, M., Muskens, J. G. A. M., Bögels, S. M., & Kamphuis, J. H. (2012). Psychometric properties of the five facets mindfulness questionnaire (FFMQ) in a meditating and a non-meditating sample. *Assessment, 19*(2), 187–197.

Devaux, M., Sassi, F. (2016). Social disparities in hazardous alcohol use: self-report bias may lead to incorrect estimates. *The European Journal of Public Health*, *26*(1), 129–134.

De Vibe, M., Solhaug, I., Tyssen, R., Friborg, O., Rosenvinge, J. H., Sørlie, T., … Bjørndal, A. (2015). Does personality moderate the effects of mindfulness training for medical and psychology students? *Mindfulness, 6*(2), 281–289.

Dewe, P. (1991), Primary appraisal, secondary appraisal and coping: Their role in stressful work encounters. *Journal of Occupational Psychology*, *64*(4), 331-351.

De Witte, M., Spruit, A., van Hooren, S., Moonen, X., & Stams, G.-J. (2019). Effects of music interventions on stress-related outcomes: A systematic review and two meta-analyses. *Health Psychology Review*, *14*(2), 294-324.

Desteno, D., Lim, D., Duong, F., & Condon, P. (2018). Meditation inhibits aggressive responses to provocations. *Mindfulness, 9*(4), 1117-1122.

Dettori, J., R. (2011). Loss to follow-up. *Evidence-based Spine-care Journal*, *2*(1), 7–10.

Ditzen, B., Neumann, I. D., Bodenmann, G., von Dawans, B., Turner, R. A., Ehlert, U., & Heinrichs, M. (2007). Effects of different kinds of couple interaction on cortisol and heart rate responses to stress in women. *Psychoneuroendocrinology*, *32*(5), 565–574.

Donnelly, D. A., & Murray, E. J. (1991). Cognitive and emotional changes in written essays and therapy interviews. *Journal of Social and Clinical Psychology*, *10*(3), 334–350.

Du, J., Huang, J., An, Y., & Xu, W. (2018). The relationship between stress and negative emotion: The mediating role of rumination. *Clinical Research and Trials*, *4*(1), 1–5.

Jones, E., Fear, N. T., & Wessely, S. (2007). Shell shock and mild traumatic brain injury: A historical review. *The American Journal of Psychiatry*, *164*(11), 1641–1645.

Economides, M., Martman, J., Bell, M.J., & Sanderson, B. (2018). Improvements in stress, affect, and irritability following brief use of a mindfulness-based smartphone app: A randomized controlled trial. *Mindfulness*, *9*(5), 1584-1593.

Eddie, D., Kim, C., Lehrer, P., Deneke, E., & Bates, M. E. (2014). A pilot study of brief heart rate variability biofeedback to reduce craving in young adult men receiving inpatient treatment for substance use disorders. *Applied Psychophysiology and Biofeedback*, *39*(3-4), 181–192.

Ein-Dor, T., Coan, J. A., Reizer, A., Gross, E. B., Dahan, D., Wegener, M. A., Carel, R., Cloninger, C. R., & Zohar, A. H. (2015). Sugarcoated isolation: Evidence that social avoidance is linked to higher basal glucose levels and higher consumption of glucose.

Frontiers in Psychology, *6*, 492.

Epictetus, et al. (1990). *Manuale*. Garzanti.

Erickson, K., & Schulkin, J. (2003). Facial expressions of emotion: A cognitive neuroscience perspective. *Brain and Cognition*, *52*(1), 52–60.

Feldman, L. A. (1995). Valence focus and arousal focus: Individual differences in the structure of affective experience. *Journal of Personality and Social Psychology*, *69*(1), 153–166.

Feldman, G., Greeson, J., & Senville, J. (2010). Differential effects of mindful breathing, progressive muscle relaxation, and loving-kindness meditation on decentering and negative reactions to repetitive thoughts. *Behaviour Research and Therapy*, *48*(10), 1002–1011.

Ferguson, C., & Heene, M. (2012). A vast graveyard of undead theories. *Perspectives on Psychological Science, 7*, 555 - 561.

Frank, D. L., Khorshid, L., Kiffer, J. F., Moravec, C. S., & McKee, M. G. (2010). Biofeedback in medicine: who, when, why and how? *Mental Health in Family Medicine*, *7*(2), 85–91.

Fraley, R. C., Heffernan, M. E., Vicary, A. M., & Brumbaugh, C. C. (2011). The experiences in close relationships-relationship Structures questionnaire: A method for assessing attachment orientations across relationships. *Psychological Assessment*, *23*(3), 615–625.

Fueshko, T.M. (2016). The Intricacies of shell shock: A chronological history of the Lancet's publications by Dr. Charles S. Myers and his Contemporaries. *Peace Change*, *41*: 38-51.

Garland, E. L., Hanley, A., Farb, N. A., & Froeliger, B. E. (2015). State mindfulness during meditation predicts enhanced cognitive reappraisal. *Mindfulness*, *6*(2), 234–242

Gebauer, J. E., Nehrlich, A. D., Stahlberg, D., Sedikides, C., Hackenschmidt, A., Schick, D., Stegmaier, C. A., Windfelder, C. C., Bruk, A., & Mander, J. (2018). Mind-Body practices and the self: Yoga and meditation do not quiet the ego but instead boost self-enhancement. *Psychological Science*, *29*(8), 1299–1308.

Gebel, M. (2019). *Calm, the 7-year-old meditation app, says it's now valued at $1 billion*. Business Insider. Retrieved February 4, 2022, from https://www.businessinsider.com/relaxation-app-calm-first-meditation-startup-1-billion-valuation-2019-2?r=US&IR=T..

Germer, C. K., Siegel, R. D., & Fulton, P. R. (2016). Mindfulness and psychotherapy. *Guilford Press.*

Gevirtz, R. (2013). The promise of heart rate variability biofeedback: Evidence-based application. *Biofeedback, 41*(3), 110–120.

Giluk, T. L. (2009). Mindfulness, big five personality, and affect: A meta-analysis. *Personality and Individual Differences*, *47*(8), 805–811.

Glück, T. M., & Maercker, A. (2011). A randomized controlled pilot study of a brief web-based mindfulness training. *BMC Psychiatry*, *11*(1).

Goessl, V. C., Curtiss, J. E., & Hofmann, S. G. (2017). The effect of heart rate variability biofeedback training on stress and anxiety: A meta-analysis. *Psychological Medicine*, *47*(15), 2578–2586.

Goldstein, D. S., & McEwen, B. (2002). Allostasis, homeostats, and the nature of stress. *Stress*, *5*(1), 55–58.

Gorgias, MacDowell, D. M., Gorgias, & Gorgias. (2005). *Encomium of helen*. Bristol Classical Press.

Grazuleviciene, R., Vencloviene, J., Kubilius, R., Grizas, V., Dedele, A., Grazulevicius, T., Ceponiene, I., Tamuleviciute-Prasciene, E., Nieuwenhuijsen, M. J., Jones, M., &

Gidlow, C. (2015). The effect of park and urban environments on coronary artery disease patients: A randomized trial. *BioMed Research International*, *2015*, 403012.

Green, Z. A., Faizi, F., Jalal, R., & Zadran, Z. (2021). Emotional support received moderates academic stress and mental well-being in a sample of Afghan university students amid COVID-19. *The International Journal of Social Psychiatry* 207640211057729. Advance online publication.

Grinker, R. R., & Spiegel, J. P.. *Men under Stress*. Blakiston Company. 1945.

Gross, C. H. (1998). Claude Bernard and the constancy of the internal environment. *Neuroscientist*, *4*, 380 –385.

Gross, J. J. (1998a). Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression, and Physiology. *Journal of PersonalIty and Social Psychology*, *74*, 224-237.

Gross, J. J. (1998b). The emerging field of emotion regulation: An integrative review. *Review of General Psychology*, *2*(3), 271–299.

Gross, J. J., & Thompson, R. A. (2007). *Emotion Regulation: Conceptual Foundations. In J. J. Gross (Ed.), Handbook of emotion regulation* (pp. 3–24). The Guilford Press.

Gummidela, V., Silva, D., & Gutierrez-Osuna, R. (2021). Evaluating the role of breathing guidance on game-based interventions for relaxation training. *Frontiers in Digital Health*, *3*, 760268.

Halgren, E. (1992). *Emotional neurophysiology of the amygdala within the context of human cognition. The amygdala: Neurobiological aspects of emotion, memory, and mental dysfunction* (pp. 191–228) In J. P. Aggleton (Ed.).

Harandi, T. F., Taghinasab, M. M., & Nayeri, T. D. (2017). The correlation of social support with mental health: A meta-analysis. *Electronic Physician*, *9*(9), 5212–5222.

Harrer, M., Cuijpers, P., Furukawa, T.A., & Ebert, D.D. (2021). *Doing meta-analysis with R: A hands-on guide.* Boca Raton, FL and London: Chapman & Hall/CRC Press. ISBN 978-0-367-61007-4.

Headspace. (n.d.). *Researching meditation and mindfulness*. https://www.headspace.com/science/meditation-research.

Headspace (2021). *Researching meditation and mindfulness*. Retrieved from https://www.headspace.com/science/meditation-research

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, *1*(1)*,* 39–65.

Hedges, L.V. & Olkin, I. (1985). *Statistical methods for meta-analysis.* Academic Press.

Heim, C., & Nemeroff, C. B. (2001). The role of childhood trauma in the neurobiology of mood and anxiety disorders: preclinical and clinical studies. Biological psychiatry, *49*(12), 1023–1039.

Hellhammer, D. H., Wüst, S., & Kudielka, B. M. (2009). Salivary cortisol as a biomarker in stress research. *Psychoneuroendocrinology*, *34*(2), 163–171.

Hemingway, E. (1952). *The old man and the sea*. Life.

Heraclitus, & Marcovich, M. (1978). *Eraclito: Frammenti*. La nuova Italia.

Hinkle, L. E. (1974). The concept of "stress" in the biological and social sciences. *International Journal of Psychiatry in Medicine*, *5*(4), 335–357.

Holt-Lunstad, J., Birmingham, W. A., & Light, K. C. (2008). Influence of a "warm touch" support enhancement intervention among married couples on ambulatory blood pressure, oxytocin, alpha amylase, and cortisol. *Psychosomatic Medicine*, *70*(9), 976–985.

Holmes (1986). Claude Bernard, the "Milieu Intérieur", and regulatory physiology. *History and Philosophy of the Life Sciences*, *8*(1), 3–25.

Hong, S., & Reed, W. R. (2020). Using Monte Carlo experiments to select meta-analytic estimators. *Research Synthesis Methods 12*(2), 192–215.

Hutcherson, C. A., Seppala, E. M., & Gross, J. J. (2008). Loving-kindness meditation increases social connectedness. *Emotion*, *8*(5), 720–724.

IJzerman, H., Lewis, N. A., Przybylski, A. K., Weinstein, N., DeBruine, L., Ritchie, S. J., ... & Anvari, F. (2020). Use caution when applying behavioural science to policy. *Nature Human Behaviour, 4*(11), 1092-1094.

IJzerman, H., Hadi, R., Coles, N. A., Paris, B., Sarda, E., Fritz, W., Klein, R., Ropovik, I. (2022). Social Thermoregulation: A Meta-Analysis. *Available via PsyArxiv,* https://doi.org/10.31234/osf.io/fc6yq.

International Personality Item Pool. (2001). A scientific collaborator for the development of advanced measures of personality traits and other individual differences. Retrieved from http://ipip.ori.org/.

Ioannidis, J. P. A. (2008). Why most discovered true associations are inflated. *Epidemiology*, *19*(5), 640–648.

Iyengar, S., & Greenhouse, J., B. (1988). Selection models and the file drawer problem. *Statistical Science*, *3*(1), 109–117.

Irvine, W. B. *Guide to the good life: The ancient art of stoic joy*. Oxford University Press. 2009.

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, *23*(5), 524–532.

Kabat-Zinn, J. (1994). *Wherever you go, there you are: Mindfulness meditation in everyday* *life*. Hyperion.

Kabat-Zinn, J. (2008). *Full catastrophe living: using the wisdom of your body and mind to face stress, pain, and illness*. Random House.

Khantzian E. J. (1985). The self-medication hypothesis of addictive disorders: Focus on heroin and cocaine dependence. *The American Journal of Psychiatry*, *142*(11), 1259–1264.

Khoury, B., Lecomte, T., Fortin, G., Masse, M., Therien, P., Bouchard, V., … Hofmann, S. G. (2013). Mindfulness-based therapy: A comprehensive meta-analysis. *Clinical Psychology Review, 33(*6), 763–771.

Kirschbaum, C., & Hellhammer, D. H. (1994). Salivary cortisol in psychoneuroendocrine research: Recent developments and applications. *Psychoneuroendocrinology*, *19*(4), 313–333.

Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Adams, R. B., Alper, S., Aveyard, M., Axt, J. R., Babalola, M. T., Bahník, Š., Batra, R., Berkics, M., Bernstein, M. J., Berry, D. R., Bialobrzeska, O., Binan, E. D., Bocian, K., Brandt, M. J., Busching, R., ... Nosek, B. A. (2018). Many labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science*, *1*(4), 443–490.

Kolek, L., Ropovik, I., Sisler, V., van Oostendorp, H., & Brom, C. (2022). Video games and attitude change: A Meta-analysis. *PsyArxiv,* https://doi.org/10.31234/osf.io/8y7jn.

Kotera, Y., Richardson, M., & Sheffield, D. (2020). Effects of Shinrin-Yoku (forest bathing) and nature therapy on mental health: a systematic review and meta-analysis. *International Journal of Mental Health and Addiction, 20*(1), 337-361.

Lambert, E. A., & Lambert, G. W. (2011). Stress and its role in sympathetic nervous system activation in hypertension and the metabolic syndrome. Current Hypertension Reports, *13*(*3*), 244–248.

Lane, R. D., & Nadel, L. (Eds.). (2000). *Cognitive neuroscience of emotion*. Oxford University Press.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Lakey, B., & Cronin, A. (2008). Low social support and major depression: Research, theory and methodological issues. In K. S. Dobson & D. J. A. Dozois (Eds.), *Risk factors in depression* (pp. 385–408). Elsevier Academic Press.

Lazarus, R. S. *Psychological stress and the coping process*. McGraw-Hill. 1966.

Lazarus, R. S., & Folkman, S. *Stress, appraisal, and coping*. Springer, 1984.

Le Carré's, J. *'They told me I was grown up enough to keep a secret': Exclusive extract from Silverview, John le Carré's final novel*. The Guardian. Retrieved February 4, 2022, from https://www.theguardian.com/books/ng-interactive/2021/oct/09/they-told-me-i-was-grown-up-enough-to-keep-a-secret-exclusive-extract-from-silverview-john-le-carres-final-novel.

Levin-Aspenson, H. F., & Watson, D. (2018). Mode of administration effects in psychopathology assessment: Analyses of gender, age, and education differences in self-rated versus interview-based depression. *Psychological Assessment, 30*(3), 287–295.

Lee, J., Bum-Jin Park, Yuko Tsunetsugu, Takahide Kagawa & Yoshifumi Miyazaki (2009). Restorative effects of viewing real forest landscapes, based on a comparison with urban landscapes, *Scandinavian Journal of Forest Research*, *24*(3), 227-234.

Lee, M. D., & Wagenmakers, E.-J. (2013). Bayesian cognitive modeling: A practical course. *Cambridge University Press.*

Lee, R. A., & Jung, M. E. (2018). Evaluation of an mHealth App (DeStressify) on university students' mental health: Pilot trial. *JMIR mental health*, *5*(1), e2.

Lehrer, P. M., Vaschillo, E., & Vaschillo, B. (2000). Resonant frequency biofeedback training to increase cardiac variability: Rationale and manual for training. *Applied Psychophysiology and Biofeedback*, *25*(3), 177–191.

Lehrer, P., Vaschillo, B., Zucker, T., Graves J., Katsamanis, M., Aviles, M., & Wamboldt F., (2013). Protocol for heart rate variability biofeedback training. *Biofeedback, 41(3)*, 98–109.

Lehrer, P. M., & Gevirtz, R. (2014). Heart rate variability biofeedback: How and why does it work? *Frontiers in Psychology*, *5,* 756.

Le Moal M. (2007). Historical approach and evolution of the stress concept: A personal account. *Psychoneuroendocrinology*, *32*(1), S3–S9.

Lepore S. J. (1997). Expressive writing moderates the relation between intrusive thoughts and depressive symptoms. Journal of Personality and Social Psychology, *73*(5), 1030–1037.

Levens, S. M., Elrahal, F., & Sagui, S. J. (2016). The role of family support and perceived stress reactivity in predicting depression in college freshman. *Journal of Social and Clinical Psychology*, *35*(4), 342–355

Lin, I. M., Fan, S. Y., Yen, C. F., Yeh, Y. C., Tang, T. C., Huang, M. F., Liu, T. L., Wang, P. W., Lin, H. C., Tsai, H. Y., & Tsai, Y. C. (2019). Heart rate variability biofeedback increased autonomic activation and improved symptoms of depression and insomnia among patients with major depression disorder. *Clinical Psychopharmacology and Neuroscience, 17*(2), 222–232.

Lippman, A., E. (1963). Hellenic conceptions of harmony. *Journal of the American*

Musicological Society, *16*(1), 3–35.

Lüdecke, D. (2019). esc: Effect size computation for meta analysis (Version 0.5.1). https://CRAN.R-project.org/package=esc.

Lundqvist, C., Ståhl, L., Климт Kenttä, G., & Thulin, U. (2018). Evaluation of a mindfulness intervention for paralympic leaders prior to the paralympic games. *International Journal of Sports Science & Coaching*, *13*(1), 62–71.

Lindquist, K. A., MacCormack, J. K., & Shablack, H. (2015). The role of language in emotion: predictions from psychological constructionism. *Frontiers in Psychology*, *6*, 444.

MacLehose W. (2020). Historicising stress: Anguish and insomnia in the middle ages. *Interface Focus*, *10*(3), 20190094.

MacCoon, D. G., Imel, Z. E., Rosenkranz, M. A., Sheftel, J. G., Weng, H. Y., Sullivan, J. C., Bonus, K. A., Stoney, C. M., Salomons, T. V., Davidson, R. J., & Lutz, A. (2012). The validation of an active control intervention for Mindfulness Based Stress Reduction (MBSR). *Behaviour Research and Therapy*, *50*(1), 3–12.

Maggio, L. A., Tannery, N. H., & Kanter, S. L. (2011). Reproducibility of literature search reporting in medical education reviews. *Academic Medicine*, *86*(8), 1049–1054.

Marselle, M. R., Irvine, K. N., & Warber, S. L. (2013). Walking for well-being: Are group walks in certain types of natural environments better for well-being than group walks in urban environments? *International Journal of Environmental Research and Public Health*, *10*(11), 5603–5628.

Marselle, M., Irvine, K., & Warber, S. (2014). Examining group walks in nature and multiple aspects of well-being: A large-scale study. *Ecopsychology, 6(3), 134-147.*

Matko, K., & Sedlmeier, P. (2019). What Is Meditation? Proposing an empirically derived classification system. *Frontiers In Psychology*, *10*, 2276.

Maxwell, S. E. (2004). The Persistence of underpowered studies in psychological research: causes, consequences, and remedies. *Psychological Methods*, *9*(2), 147–163.

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean? *American Psychologist*, *70*(6), 487–498.

May Barry, K., Woods, M., Martin, A., Stirling, C., & Warnecke, E. (2018). A randomized controlled trial of the effects of mindfulness practice on doctoral candidate psychological status. *Journal of American College Health*, *67*(4), 299-307.

McEwen, B. S., & Stellar, E. (1993). Stress and the individual. *Mechanisms Leading to Disease*. *Archives of Internal Medicine*, *153*(18), 2093–2101.

McEwen, B. S., & Magarinos, A. M. (1997). Stress effects on morphology and function of the hippocampus. *Annals of the New York Academy of Sciences*, *821*, 271–284.

McEwen B. S. (1998). Stress, adaptation, and disease. Allostasis and allostatic load. *Annals of the New York Academy of Sciences*, *840*, 33–44.

McEwen B. S. (2000). The neurobiology of stress: From serendipity to clinical relevance. *Brain Research*, *886*(1-2), 172–189.

McEwen B. S. (2006). Protective and damaging effects of stress mediators: Central role of the brain. *Dialogues in clinical neuroscience*, *8*(4), 367–381.

McEwen B. S. (2007). Physiology and neurobiology of stress and adaptation: Central role of the brain. *Physiological Reviews*, *87*(3), 873–904.

McFarland, C., Buehler, R., von Rüti, R., Nguyen, L., & Alvaro, C. (2007). The impact of negative moods on self-enhancing cognitions: The role of reflective versus ruminative mood orientations. *Journal of Personality and Social Psychology*, *93*(5), 728–750.

McNamee, R. (2003). Confounding and confounders. *Occupational and Environmental*

Medicine, *60*(3), 227–234.

McShane, B. B., Böckenholt, U., & Hansen, K. T. (2016). Adjusting for publication bias in meta-analysis. *Perspectives on Psychological Science*, *11*(5), 730–749.

Moreau, D., & Gamble, B. (2022). Conducting a meta-analysis in the age of open science: Tools, tips, and practical recommendations. *Psychological Methods, 27*(3), 426 - 432.

Moore, A., & Malinowski, P. (2009). Meditation, mindfulness and cognitive flexibility. *Consciousness and Cognition*, *18*(1), 176–186.

Morris, J. D. (1995). Observations: SAM: The self-assessment manikin: An efficient cross-cultural measurement of emotional response. *Journal of Advertising Research, 35*(6), 63–68.

Myers, C. S. (1916). Contributions to the study of shell shock: Being and account of certain disorders of cutaneous sensibility. *Lancet*, *187*, 608–613.

Nichols, A. L., & Maner, J. K. (2008). The good-subject effect: Investigating participant demand characteristics. *The Journal of General Psychology*, *135*(2), 151–165.

Nyklíček I., & Irrmischer M. (2017). For whom does mindfulness-based stress reduction work? Moderating effects of personality. *Mindfulness, 8*, 1106–1116.

Olfson, M., King, M., & Schoenbaum, M. (2015). Benzodiazepine use in the United States. *JAMA Psychiatry*, *72*(2), 136–142.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), 943–952.

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews. *Systematic Reviews*, *5*(1), 210–220.

Pelletier C. L. (2004). The effect of music on decreasing arousal due to stress: A meta-analysis. *Journal of Music Therapy*, *41*(3), 192–214.

Perez, S. (2020). *Meditation and mindfulness apps continue their surge amid pandemic.* TechCrunch. https://techcrunch.com/2020/05/28/meditation-and-mindfulness-apps-continue-their-surge-amid-pandemic/.

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience*, *9*(2), 148–158.

Phelps, E. A. (2006). Emotion and cognition: Insights from studies of the human amygdala. *Annual Review of Psychology*, *57*(1), 27–53.

Philippon, J., & Poirier, J. *Joseph Babinski: A Biography*. Oxford University Press, 2009.

Prinsloo, G. E., Rauch, H. G. L., Lambert, M. I., Muench, F., Noakes, T. D., & Derman, W. E. (2010). The effect of short duration heart rate variability (HRV) biofeedback on cognitive performance during laboratory induced cognitive stress. *Applied Cognitive Psychology*, *25*(5), 792–801.

Pustejovsky, J. (2017). *You wanna PEESE of d's? Blogpost.* Retrieved from https://www.jepusto.com/pet-peese-performance/

Pustejovsky, J., Tipton, E. (2022). Meta-analysis with robust variance estimation: Expanding the range of working models. *Prevention Science, 23*(3), 425-438.

Puterman, E., DeLongis, A., & Pomaki, G. (2010). Protecting us from ourselves: Social support as a buffer of trait and state rumination. *Journal of Social and Clinical Psychology, 29*(7), 797–820.

Raffone, A., & Srinivasan, N. (2017). Mindfulness and cognitive functions: Toward a unifying neurocognitive framework. *Mindfulness*, *8*(1), 1–9.

Raman, T.L., Abdul Aziz, N.A., & Yaakob, S.S.N. (2021). The effects of different natural environment influences on health and psychological well-being of people: A Case study in Selangor. *Sustainability*, *13*, 8597.

Ramsay, D. S., & Woods, S. C. (2014). Clarifying the roles of homeostasis and allostasis in physiological regulation. *Psychological Review*, *121*(*2*), 225–247.

Ratanasiripong, P., Ratanasiripong, N., & Kathalae, D. (2012). Biofeedback intervention for stress and anxiety among nursing students: A randomized controlled trial. *ISRN Nursing*, *2012*, 827972.

Reblin, M., & Uchino, B. N. (2008). Social and emotional support and its implication for health. *Current Opinion in Psychiatry*, *21*(2), 201–205.

Revelle, W., & Condon, D., M. (2018). Reliability. In Irwing, P., Booth, T., & Hughes, D., editors. *Wiley-Blackwell Handbook of Psychometric Testing*. Wiley-Blackwell.

Revenson, T. A., Schiaffino, K. M., Majerovitz, S. D., & Gibofsky, A. (1991). Social support as a double-edged sword: The relation of positive and problematic support to depression among rheumatoid arthritis patients. *Social Science & Medicine*, 33*(7),* 807–813.

Rich, R.M., Ogden, J. & Morison, L. (2021). A randomized controlled trial of an app-delivered mindfulness program among university employees: effects on stress and work-related outcomes. *International Journal of Workplace Health Management*. *14* (2), 201-216.

Robinson, A., M. (2018). Let's talk about stress: History of stress research. *Review of General Psychology*, *22*(3), 334–342.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*(3), 638–641.

Rouder, J. N. (2014). Optional stopping: No problem for Bayesians. *Psychonomic Bulletin & Review, 21*(2), 301–308.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review,* *16*(2), 225–237.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178.

Sanada, K., Montero-Marin, J., Alda Díez, M., Salas-Valero, M., Pérez-Yus, M. C., Morillo, H., Demarzo, M. M., García-Toro, M., & García-Campayo, J. (2016). Effects of mindfulness-based interventions on salivary cortisol in healthy adults: A meta-analytical review. *Frontiers in Physiology*, *7,* 471.

Sapolsky, R. M., (1996). Why stress is bad for your brain. *Science 273*, 749–750.

Schäfer, T., & Schwarz, M. A. (2019). The meaningfulness of effect sizes in psychological research: Differences between sub-disciplines and the impact of potential biases. *Frontiers in Psychology*, *10*, 813.

Schardt, C., Adams, M. B., Owens, T., Keitz, S., & Fontelo, P. (2007). Utilization of the PICO framework to improve searching PubMed for clinical questions. *BMC Medical Informatics and Decision Making*, *7*, 16.

Schneiderman, N., Ironson, G., & Siegel, S. D. (2005). Stress and health: Psychological, behavioral, and biological determinants. *Annual Review of Clinical Psychology*, *1*, 607–628.

Schneider, T. R., Rench, T. A., Lyons, J. B., & Riffle, R. R. (2012). The influence of neuroticism, extraversion and openness on stress responses. *Stress and Health: Journal of the International Society for the Investigation of Stress*, *28*(2), 102–110.

Schneiderman, N., Ironson, G., & Siegel, S. D. (2005). Stress and health: Psychological, behavioral, and biological determinants. *Annual Review of Clinical Psychology*, *1*, 607–628.

Schönbrodt, F. D., Wagenmakers, E. J., Zehetleitner, M., & Perugini, M. (2017). Sequential hypothesis testing with Bayes factors: Efficiently testing mean differences.

Psychological Methods, *22*(2), 322–339.

Schutte, N. S. & John M. Malouff, J. M. (2018) Mindfulness and connectedness to nature: A meta-analytic investigation. *Personality and Individual Differences*, *27*(10-14).

Schwarzer, R., & Leppin, A. (1989). Social support and health: A meta-analysis. *Psychology & Health, 3*(1), 1–15.

Segal, Z. V., Teasdale, J. D., Williams, J. M., & Gemar, M. C. (2002). The mindfulness-based

cognitive therapy adherence scale: Inter-rater reliability, adherence to protocol and treatment distinctiveness. *Clinical Psychology & Psychotherapy*, *9*(2), 131–138.

Selye H. (1946). The general adaptation syndrome and the diseases of adaptation. *The Journal of Clinical Endocrinology and Metabolism*, *6*, 117–230.

Selye, H. (1956). *The stress of life*. McGraw-Hill.

Selye H. (1998). A syndrome produced by diverse nocuous agents. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *10*(2), 230–231.

Shakespeare, W., Melchiori, G., & Weil, H. (2007). King Henry Iv. Cambridge Univ. Pr. Stress. (n.d.) In Oxford Dictionary Online. Retrieved from http://www .oxforddictionaries.com/us/definition/american_english/stress.

Silan, M., Bellemin, R., Dujols, O., Sparacio, A., Adetula, A., & IJzerman, H. (2021). CO-RE Lab lab philosophy v5. Retrieved from https://psyarxiv.com/6jmhe/.

Simmons, J., Nelson, L., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366.

Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science*, *12*(6), 1123-1128.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General*, *143*(2), 534–547.

Sparacio, A., Ropovik, I., Jiga-Boy, G. M., Forscher, P. S., Paris, B., & IJzerman, H. (2022). Stress regulation via self-administered mindfulness and biofeedback interventions in adults: A pre-registered meta-analysis. *Available via PsyArXiv*, https://psyarxiv.com/zpw28/.

Sparacio, A., IJzerman, H., Ropovik, I., Giorgini, F., Spiessens C., … Jiga-Boy, G., M., 2022. (article in preparation) A large, multi-site test of self-administered mindfulness in English speakers. *Available via OSF,* https://osf.io/6w2zm/.

Spielberger, C. D., Gorshu, R. L. and Lushene, R. D. *Test manual for the state trait anxiety inventory*. Consulting Psychologist Press, 1970.

Spijkerman, M. P. J., Pots, W. T. M., & Bohlmeijer, E. T. (2016). Effectiveness of online mindfulness based interventions in improving mental health: A review and meta-analysis of randomized controlled trials. *Clinical Psychology Review*, *45*, 102–114.

Stanley T.D. & Doucouliagos C.H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods*, *5*(1), 60–78.

Stefan, A. M., Gronau, Q. F., Schönbrodt, F. D., & Wagenmakers, E. J. (2019). A tutorial on bayes factor design analysis using an informed prior. *Behavior Research Methods*, *51*(3), 1042–1058.

Sterling, P., & Eyer, J. (1988). *Allostasis: A new paradigm to explain arousal pathology*. In S.

Fisher & J. Reason (Eds.), Handbook of life stress, cognition and health (pp. 629–649). John Wiley & Sons.

Sterling, P. *What is health?: Allostasis and the evolution of human design*. The MIT Press, 2020.

Stephens, M. A. C., & Wand, G. (2012). Stress and the HPA axis: Role of glucocorticoids in alcohol dependence. *Alcohol Research: Current Reviews, 34*(4), 468-483.

Sterne, J., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., A., … Higgins, J. (2019). RoB 2: A revised tool for assessing risk of bias in randomised trials. *The BMJ*, *366*, l4898.

Storbeck, J., & Clore, G. L. (2008). *The affective regulation of cognitive priming. Emotion*, *8*(2), 208–215.

Shalev, A. Y., Yehuda, R., & McFarlane, A. C. (Eds.). *International handbook of human response to trauma*. Kluwer Academic Publishers, 2000.

Strictus. (n.d.) In latin-dictionary online. Retrieved from http://www.latindictionary.org/strictus.

Sutton, A. J., Duval, S. J., Tweedie, R. L., Abrams, K. R., & Jones, D. R. (2000). Empirical assessment of effect of publication bias on meta-analyses. *The BMJ*, *320*(7249), 1574–1577.

Szabelska, A., Pollet, T. V., Dujols, O., Klein, R. A., & IJzerman, H. (2022). A tutorial for exploratory research: An eight-step approach. *Available on Psyarxiv* https://psyarxiv.com/cy9mz/download/?format=pdf.

Tang, R., & Braver, T. S. (2020). Towards an individual differences perspective in mindfulness training research: Theoretical and empirical considerations. *Frontiers in Psychology*, *11*, 818.

Taylor H., Strauss C., Cavanagh K., (2021). Can a little bit of mindfulness do you good? A systematic review and meta-analyses of unguided mindfulness-based self-help interventions. *Clinical Psychology Review*. *89,* 102078.

Teasdale, J. D., Segal, Z. V., & Williams, J. M. G. (2006). Mindfulness Training and Problem Formulation. *Clinical Psychology: Science and Practice*, *10*(2), 157–160.

Teper, R., Segal, Z. V., & Inzlicht, M. (2013). Inside the mindful mind: How mindfulness enhances emotion regulation through improvements in executive control. *Current Directions in Psychological Science*, *22*(6), 449–454.

Thayer, J. F., & Friedman, B. H. (2002). Stop that! Inhibition, sensitization, and their neurovisceral concomitants. *Scandinavian Journal of Psychology*, *43*(2), 123–130.

Tolkien, J., R., R. *Smith of Wootton Major*. Redbook, 1967.

Tsuji, S., Bergmann, C., & Cristia, A. (2014). Community-augmented meta-analyses. *Perspectives on Psychological Science*, *9*(6), 661–665.

Twain, Mark, et al. *The Adventures of Tom Sawyer*. World Publications Group, Inc., 2013.

Ulrich, R.S. (1979). Visual landscapes and psychological well-being. *Landscape Research, 4*, 17-23.

Ulrich, R. S. (1983). Aesthetic and affective response to natural environment. In I. Altman & J. F. Wohlwill (Eds.), *Behavior and the Natural Environment* (pp. 85–125). New York: Springer.

Ustinova, Y., & Cardeña, E. (2014). Combat stress disorders and their treatment in ancient Greece. *Psychological Trauma: Theory, Research, Practice, and Policy*, *6*(6), 739–748.

Van Gordon, W., Shonin, E., & Garcia-Campayo, J. (2017). Are there adverse effects associated with mindfulness? *Australian & New Zealand Journal of Psychiatry*, *51*(10), 977–979.

Vaughan-Johnston, T. I., Jacobson, J. A., Prosserman, A., & Sanders, E. (2021). Mind-body practices and self-enhancement: direct replications of gebauer et al.'s (2018) experiments 1 and 2. *Psychological Science*, 956797621997366. Advance online publication.

Vaschillo, E. G., Vaschillo, B., & Lehrer, P. M. (2006). Characteristics of resonance in heart rate variability stimulated by biofeedback. *Applied Psychophysiology & Biofeedback, 31*, 129–142.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48.

Viechtbauer, W., & Cheung, M. W. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, *1*(2), 112–125.

Virgile, Penna, A. L., Canali, L., & Pennacchietti, S. *Bucoliche*. Biblioteca Universale Rizzoli, 1990.

Vogt, K. *Seneca*. Stanford Encyclopedia of Philosophy. Retrieved August 17, 2022, from https://plato.stanford.edu/entries/seneca/.

Wahbeh, H., Svalina, M. N., & Oken, B. S. (2014). Group, one-on-one, or Internet? Preferences for mindfulness meditation delivery format and their predictors. *Open Medicine Journal*, *1*, 66–74.

Weber, S. J., & Cook, T. D. (1972). Subject effects in laboratory research: An examination of subject roles, demand characteristics, and valid inference. *Psychological Bulletin, 77*(4), 273–295.

Watson, D., Clark, L. A., & Carey, G. (1988). Positive and negative affectivity and their relation to anxiety and depressive disorders. *Journal of Abnormal Psychology*, *97*(3), 346–353.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., D., François, R., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686.

Wittmann, A., Braud, M., Dujols, O., Forscher, P., & IJzerman, H. (2022). Individual differences in adapting to temperature in French students are only related to attachment avoidance and loneliness. *Royal Society open science*, *9*(5), 201068.

Wurtman R. J. (2005). Genes, stress, and depression. Metabolism: Clinical and Experimental, *54*(5 Suppl 1), 16–19.

Yang, L., Zhao, Y., Wang, Y., Liu, L., Zhang, X., Li, B., & Cui, R. (2015). The effects of psychological stress on depression. *Current Neuropharmacology*, *13*(4), 494–504.

Yeung, S., K., Fillon, A., Protzko, J., Elsherif, M., Moreau D., & Feldman G. (2021). Experimental meta-analysis Registered Report template. Retrieved from https://osf.io/ytgrp/.

Zellars, K. L., & Perrewé, P. L. (2001). Affective personality and the content of emotional social support: Coping in organizations. *Journal of Applied Psychology, 86*(3), 459–467.

Appendices

Appendix A: Protocols and deviations

Any changes with respect to the choices established in this pre-registration will be fully disclosed on our OSF page and will be incorporated into this form: https://osf.io/6wpav/.

Appendix B: Call for unpublished data

Subject: Call for unpublished data for a meta-analysis: "**Stress regulation via being in nature and emotional social support for adults: A pre-registered meta-analysis"**

Dear Prof/Dr/Ms/Mr XXXX,

I am Alessandro Sparacio, PhD student in social psychology, at the University of Grenoble-Alpes and I'm conducting a meta-analysis on stress regulation, along with my co-authors Hans IJzerman, Ivan Ropovik, Gabriela Jiga-Boy & Patrick Forscher.

The pre-registered protocol for this meta-analysis is publicly available on the Open Science Framework (OSF) at [https://osf.io/6wpav/]

Our meta-analysis aims to address whether being in nature and emotional social support have any demonstrated efficacy in reducing stress levels.
As you have published studies relevant to this topic, we are getting in touch to see if you have any unpublished/file-drawer data, or papers in-press, which we may have missed through database searching, and which you would like to have included in the meta-analysis.

Feel free to email either the raw data (from which we will calculate summary scores) or the summary scores themselves. While any raw data emailed to us will of course remain confidential, please know that summary scores included in the meta-analysis will be made publicly available in a dataset on the OSF.

We are hoping to include as many relevant studies as possible, so any additional data is greatly appreciated.

Sincerely (also on behalf of my co-authors),

Alessandro Sparacio

This template was provided by Moreau and Gamble (2020)

Appendix C: Requesting for specific data

Subject: Requesting data for a meta-analysis, from your paper: ‘XXXX’

Dear Prof/Dr/Ms/Mr XXXX,

I am Alessandro Sparacio, PhD student in social psychology, at the University of Grenoble-Alpes and I’m conducting a meta-analysis on stress regulation, along with my co-authors Hans IJzerman, Ivan Ropovik, Gabriela Jiga-Boy & Patrick Forscher.

The pre-registered protocol for this meta-analysis is publicly available on the Open Science Framework (OSF) at [https://osf.io/6wpav/].

We think your study ‘XXXX’ meets inclusion criteria for our meta-analysis. However, the effect size we’re interested in (i.e., the correlation/difference between XXX and XXX) does not seem to be reported in the published paper.

We would be grateful if you could send either the summary scores or the raw data themselves (from which we can calculate the effect size). While any raw data emailed to us will of course remain confidential, please know that summary scores included in the meta-analysis will be made publicly available in a dataset on the OSF.

The latest we will be able to accept your data for inclusion is XXth of XXX, XXXX.

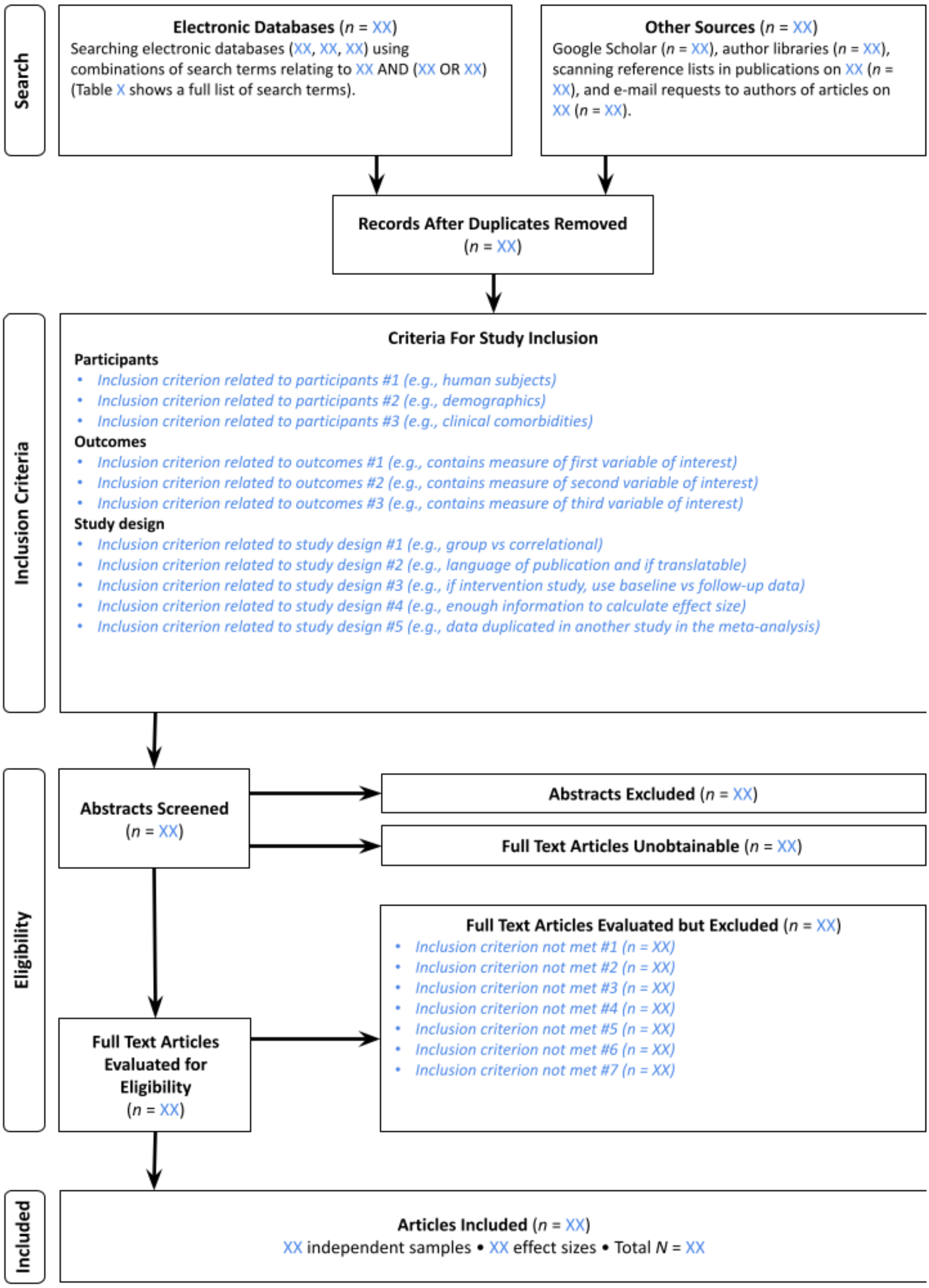
We are hoping to include as many relevant studies as possible, so any additional data is greatly appreciated.

Sincerely (also on behalf of my co-authors),

Alessandro Sparacio

This template was provided by Moreau and Gamble (2020)

Appendix D: Search criteria



This template was provided by Moreau and Gamble (2020)

Appendix E: Search strategy

BEING IN NATURE

PUBMED

(“natural space” OR “natural environment*” OR “natural landscape” OR “urban nature” OR “nearby nature” OR “nature view*” OR “outdoor nature” OR "natural space” OR “green area” OR “green environment” OR “nature contact” OR “contact with natur*” OR park OR “urban forest” OR “forest walking” OR “forest” OR “forest environment*” OR "shinrin" OR “forest bathing”) AND (walk* OR sitt* OR watch* OR view* OR stay* OR contact*) AND stress AND (“negative affect” OR “positive affect” OR emotion* OR cogniti* OR ruminati* OR physiological* OR biomarker* OR depression OR anxiety)

Search date: 12/4/2022
Results: 151 results
Notes:

PROQUEST (APA PsycArticles, APA Psycinfo, ProQuest Dissertations & Theses Global)

(“natural space” OR “natural environment*” OR “natural landscape” OR “urban nature” OR “nearby nature” OR “nature view*” OR “outdoor nature” OR "natural space” OR “nature contact” OR “contact with natur*” OR park OR “urban forest” OR “forest walking” OR “forest environment*” OR "shinrin"
OR “forest bathing”) AND (walk* OR sitt* OR watch* OR view* OR stay*) AND stress AND (“negative affect” OR “positive affect” OR emotion* OR cogniti* OR ruminati* OR physiological* OR biomarker* OR depression OR anxiety)

Search date: 12/4/2022
Results: 119
Notes:

SCOPUS

TITLE-ABS ((“greenspace*” OR “green space” OR “green landscape*” OR “natural space” OR “natural environment*” OR “natural landscape” OR “urban nature” OR “nearby nature”

OR “nature view*” OR “nature viewing” OR “viewing nature” OR “outdoor nature” OR "natural space” OR “nature contact” OR “contact with natur*” OR park OR “urban forest” OR “forest walking” OR “forest environment*” OR “nature therapy” OR “nature experience” OR “forest therapy” OR "shinrin" OR “forest bathing”) AND (walk* OR sitt* OR watch* OR view* OR stay*) AND stress AND (“negative affect” OR “positive affect” OR emotion* OR cogniti* OR ruminati* OR physiological* OR biomarker* OR depression OR anxiety)

Search date: 12/4/2022
Results: 196
Notes:

EMOTIONAL SOCIAL SUPPORT

PUBMED

("emotional support" OR "emotional social support") AND (encourage* OR help OR assist* OR love OR trust* OR contact or touch) AND stress AND ("negative affect" OR "positive affect" OR emotion* OR cogniti* OR ruminati* OR physiological* OR biomarker* OR depression OR anxiety)

Search date: 12/4/2022
Results: 288
Notes:

PROQUEST (APA PsycArticles, APA Psycinfo, ProQuest Dissertations & Theses Global)

("emotional support" OR "emotional social support") AND (encourage* OR help OR assist* OR love OR trust* OR contact or touch) AND stress AND ("negative affect" OR "positive affect" OR emotion* OR cogniti* OR ruminati* OR physiological* OR biomarker* OR depression OR anxiety)

Search date: 12/4/2022
Results: 497
Notes:

SCOPUS

TITLE-ABS (("emotional support" OR "emotional social support") AND (encourage* OR help OR assist* OR love OR trust* OR contact OR touch) AND stress AND ("negative affect" OR "positive affect" OR emotion* OR cogniti* OR ruminati* OR physiological* OR biomarker* OR depression OR anxiety))

Search date: 12/4/2022
Results: 413
Notes:

Appendix F: Correction for publication bias

1. *Primary confirmatory analysis: 4-parameter selection model (Carter et al., 2019; McShane, Böckenholt, & Hansen, 2016).*
   If there were less than four focal *p*-values per interval, the procedure fell back to the 3-parameter selection model. The selection models were implemented using a permutation-based procedure, iteratively selecting only a single focal effect size from each independent study, estimating the model in 5000 iterations, and averaging over the iterations by picking the model with the median ES estimate.
2. *Exploratory analyses*
   2.1. Vevea and Woods (2005) step function models with a priori defined selection weights, varying the assumed severity of bias, modeling moderate, severe, and extreme selection.
   2.2. Multi-level RVE-based implementation of the PET-PEESE model (Stanley & Doucouliagos, 2014), employing $\sqrt{(2/N)}$ and a $2/N$ terms instead of standard error and variance for PET and PEESE, respectively, as a measure of precision (see Pustejovsky, 2017). Additionally, the R code also allows the interested reader to use the 4PSM as a conditional estimator for PET-PEESE instead of traditional PET and explore the effect of such decision on the resulting inference (for more details, see IJzerman et al., 2022).
   2.3. Robust Bayesian model-averaging approach integrating the selection modeling and regression-based approaches (Bartoš et al., 2021), letting the data determine the contribution of each model by its relative predictive accuracy to fit the observed data.

Inferential criteria: Substantive inferences regarding the presence of an effect were guided by the estimates and inferential results of the 4-parameter selection model (4PSM) solely.

Appendix G: Study Protocol for being in nature

Based on meta-analytic synthesis of the literature, we have developed the following protocol for being in nature and stress.

Minimum Requirements:

1. Record the length of the intervention.
2. Record the type of green environment where participants are doing the experiment (e.g., whether is a park, farmland, near the sea, a forest, or an or an environment that has both natural elements and buildings).
3. Record any interaction that participants have with the natural environment (e.g., whether participants are only viewing the natural environment or whether they are engaging in any other activities).
4. If participants are in a virtual natural viewing condition, record the medium via which the intervention is administered (e.g., virtual reality).
5. In case of a randomized controlled trial, include both a passive and active control group.
6. Record the time between the intervention and stress measurement.
7. Let participants fill in the Trait Anxiety scale (Spielberger, 1970), the Experiences in Close Relationships Scale (ECR-R; Fraley et al., 2011), the Social Network Index (Cohen et al., 2000), and the Big Five inventory (John, 1991).
8. Record participants' native language, sex, gender, geographical origin, height, weight, and smoking status (if smoker, how many cigarettes).
9. Record whether the population is from students or not.
10. Record whether the population is a clinical sample or not.

Ideal Requirements:

Besides self-reported measures of stress record stress reactivity of participants at least through one physiological measure (e.g., assessment of catecholamines, assessment of the ANS via skin conductance, cortisol, heart rate, systolic and diastolic blood pressure; Bally, Campbell, Chesnick, & Tranmer, 2003; Berntson et al., 1993)

Appendix H: Study Protocol for Emotional social support

Based on meta-analytic synthesis of the literature, we have developed the following protocol for emotional social support and stress.

Minimum Requirements:

1. Record the type of the emotional social support (e.g., whether it was achieved through verbal expressions or via physical contact)
2. Record the source that provided the emotional social support (e.g., whether it was the partner, a friend, or a stranger)
3. In case of a randomized controlled trial, include an active and passive control group.
4. Record the time between the intervention and stress measurement.
5. Let participants fill in the Trait Anxiety scale (Spielberger, 1970), the Experiences in Close Relationships Scale (ECR-R; Fraley et al., 2011), the Social Network Index (Cohen et al., 2000), and the Big Five inventory (John, 1991).
6. Record participants' native language, sex, gender, geographical origin, height, weight, and smoking status (if smoker, how many cigarettes).
7. Record whether the population is from students or not.
8. Record whether the population is a clinical sample or not.

Ideal Requirements:

Besides self-reported measures of stress record stress reactivity of participants at least through one physiological measure (e.g., assessment of catecholamines, assessment of the ANS via skin conductance, cortisol, heart rate, systolic and diastolic blood pressure; Bally, Campbell, Chesnick, & Tranmer, 2003; Berntson et al., 1993).

Appendix I: Protocols and deviations

Any changes with respect to the choices established in this pre-registration will be fully disclosed on our OSF page, and will be incorporated into this form: https://osf.io/tcs64

Appendix L: Call for unpublished data

Subject: Call for unpublished data for a meta-analysis: "**Stress regulation via self-administered mindfulness and biofeedback intervention for adults: A pre-registered meta-analysis"**

Dear colleague,

I am Alessandro Sparacio, PhD student in social psychology, at the University of Grenoble-Alpes and I'm conducting a meta-analysis on stress regulation, along with my co-authors Hans IJzerman, Ivan Ropovik, Gabriela Jiga-Boy & Patrick Forscher.

The pre-registered protocol for this meta-analysis is pre-registered on the Open Science Framework (OSF) and on PROSPERO with protocol number CRD42020179810.

Our meta-analysis aims to address whether self-administered mindfulness and Heart-Rate Variability biofeedback have any demonstrated efficacy in reducing stress levels.
As you have published studies relevant to this topic, we are getting in touch to see if you have any unpublished/file-drawer data, or papers in-press, which we may have missed through database searching, and which you would like to have included in the meta-analysis.

Feel free to email either the raw data (from which we will calculate summary scores) or the summary scores themselves. While any raw data emailed to us will of course remain confidential, please know that summary scores included in the meta-analysis will be made publicly available in a dataset on the OSF.

We are hoping to include as many relevant studies as possible, so any additional data is greatly appreciated.

Sincerely (also on behalf of my co-authors),

Alessandro Sparacio
[redacted]

this template was provided by Moreau and Gamble (2020)

Dear colleagues,

I am Alessandro Sparacio, PhD student in social psychology, at the University of Grenoble-Alpes and I'm conducting a meta-analysis on stress regulation, along with my co-authors Hans IJzerman, Ivan Ropovik, Gabriela Jiga-Boy & Patrick Forscher.

The pre-registered protocol for this meta-analysis is pre-registered on the Open Science Framework (OSF) and on PROSPERO with protocol number CRD42020179810. We're trying to hunt down the grey literature of unpublished studies that have investigated whether self-administered mindfulness and Heart-Rate Variability biofeedback have any demonstrated efficacy in reducing stress levels.

If you have published studies relevant to this topic, we would love to hear from you about unpublished/file-drawer data or papers in-press that we may have missed through database searching.

Feel free to email either the raw data (from which we will calculate summary scores) or the summary scores themselves. While any raw data emailed to us will of course remain confidential, please know that summary scores included in the meta-analysis will be made publicly available in a dataset on the OSF.

We are hoping to include as many relevant studies as possible, so any additional data is greatly appreciated.

Sincerely (also on behalf of my co-authors),

Alessandro Sparacio
[redacted]

this template was provided by Moreau and Gamble (2020)

Appendix M: Requesting for specific data

Subject: Requesting data for a meta-analysis, from your paper: ‘XXXX’

Dear Prof/Dr/Ms/Mr XXXX,

I am Alessandro Sparacio, PhD student in social psychology, at the University of Grenoble-Alpes and I’m conducting a meta-analysis on stress regulation, along with my co-authors Hans IJzerman, Ivan Ropovik, Gabriela Jiga-Boy & Patrick Forscher.

The pre-registered protocol for this meta-analysis is publicly available on the Open Science Framework (OSF) at [link] and on PROSPERO [link] with protocol number [NUMBER]

We think your study ‘XXXX’ meets inclusion criteria for our meta-analysis. However, the effect size we’re interested in (i.e., the correlation/difference between XXX and XXX) does not seem to be reported in the published paper.

We would be grateful if you could send either the summary scores or the raw data themselves (from which we can calculate the effect size). While any raw data emailed to us will of course remain confidential, please know that summary scores included in the meta-analysis will be made publicly available in a dataset on the OSF.

The latest we will be able to accept your data for inclusion is XXth of XXX, XXXX.

We are hoping to include as many relevant studies as possible, so any additional data is greatly appreciated.

Sincerely (also on behalf of my co-authors),

Alessandro Sparacio
[redacted]

Subject: Requesting data for a meta-analysis, from your paper: ‘XXXX’

Dear Prof/Dr/Ms/Mr XXXX,

I am Alessandro Sparacio, PhD student in social psychology, at the University of Grenoble-Alpes and I’m conducting a meta-analysis on stress regulation, along with my co-authors Hans IJzerman, Ivan Ropovik, Gabriela Jiga-Boy & Patrick Forscher.

The pre-registered protocol for this meta-analysis is publicly available on the Open Science Framework (OSF) at [link] and on PROSPERO [link] with protocol number [NUMBER]

We think your study ‘XXXX’ meets inclusion criteria for our meta-analysis. I wanted to know if you had also conducted a study using this protocol and if you had any data to send me to include in my meta analysis

We would be grateful if you could send either the summary scores or the raw data themselves (from which we can calculate the effect size). While any raw data emailed to us will of course remain confidential, please know that summary scores included in the meta-analysis will be made publicly available in a dataset on the OSF.

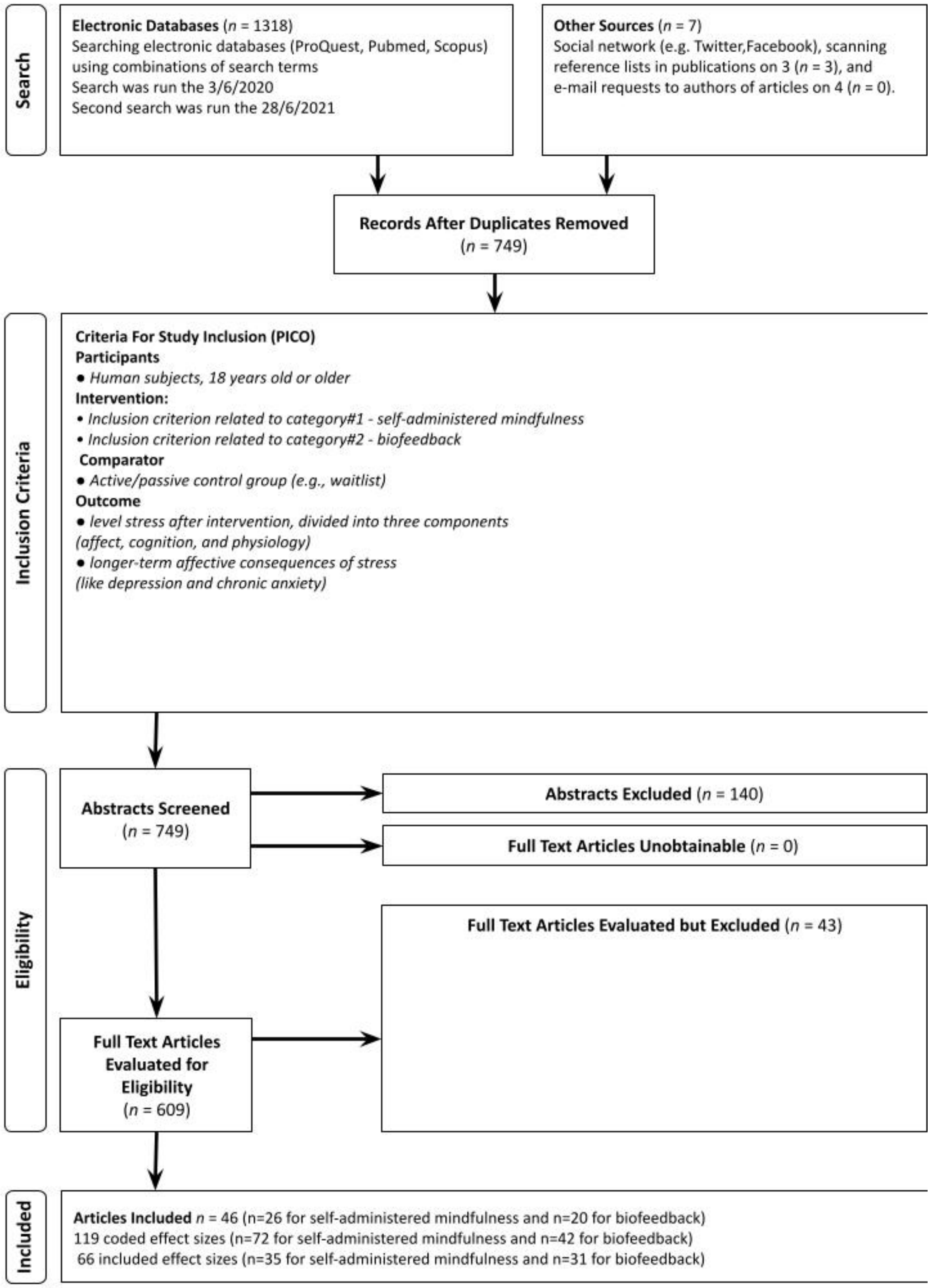
The latest we will be able to accept your data for inclusion is XXth of XXX, XXXX.

We are hoping to include as many relevant studies as possible, so any additional data is greatly appreciated.

Sincerely (also on behalf of my co-authors),

Alessandro Sparacio
[redacted]

Appendix N: Search criteria



Appendix O: Search strategy

SELF-ADMINISTERED MINDFULNESS (SAM) MEDITATION

PUBMED

mindful* AND (online OR web* OR self-help OR self-administered OR e-health OR video* OR audio OR computer* OR application* OR app) AND stress AND (affect* OR emotion* OR cogniti* OR ruminati* OR physiological* OR cortisol OR heart_rate)

Search date: XX
Results: XX
Notes:

PROQUEST (APA PsycArticles, APA Psycinfo, ProQuest Dissertations & Theses Global)

mindful* AND (online OR web* OR self-help OR self-administered OR e-health OR video* OR audio OR computer* OR application* OR app) AND stress AND (affect* OR emotion* OR cogniti* OR ruminati* OR physiological* OR cortisol OR heart_rate)

Search date: XX
Results: XX
Notes:

SCOPUS

mindful* AND (online OR web* OR self-help OR self-administered OR e-health OR video* OR audio OR computer* OR application* OR app) AND stress AND (affect* OR emotion* OR cogniti* OR ruminati* OR physiological* OR cortisol OR heart_rate)

Search date: XX
Results: XX
Notes:

HEART-RATE VARIABILITY (HRV) BIOFEEDBACK

PUBMED

(heart rate variability biofeedback OR HRVB OR respiratory sinus arrhythmia biofeedback OR RSA biofeedback OR resonance frequency feedback OR RFF) AND stress AND

(affect* OR emotion* OR cogniti* OR ruminati* OR physiological* OR cortisol OR heart_rate)

Search date: XX
Results: XX
Notes:

PROQUEST (APA PsycArticles, APA Psycinfo, ProQuest Dissertations & Theses Global)

(heart rate variability biofeedback OR HRVB OR respiratory sinus arrhythmia biofeedback OR RSA biofeedback OR resonance frequency feedback OR RFF) AND stress AND (affect* OR emotion* OR cogniti* OR ruminati* OR physiological* OR cortisol OR heart_rate)

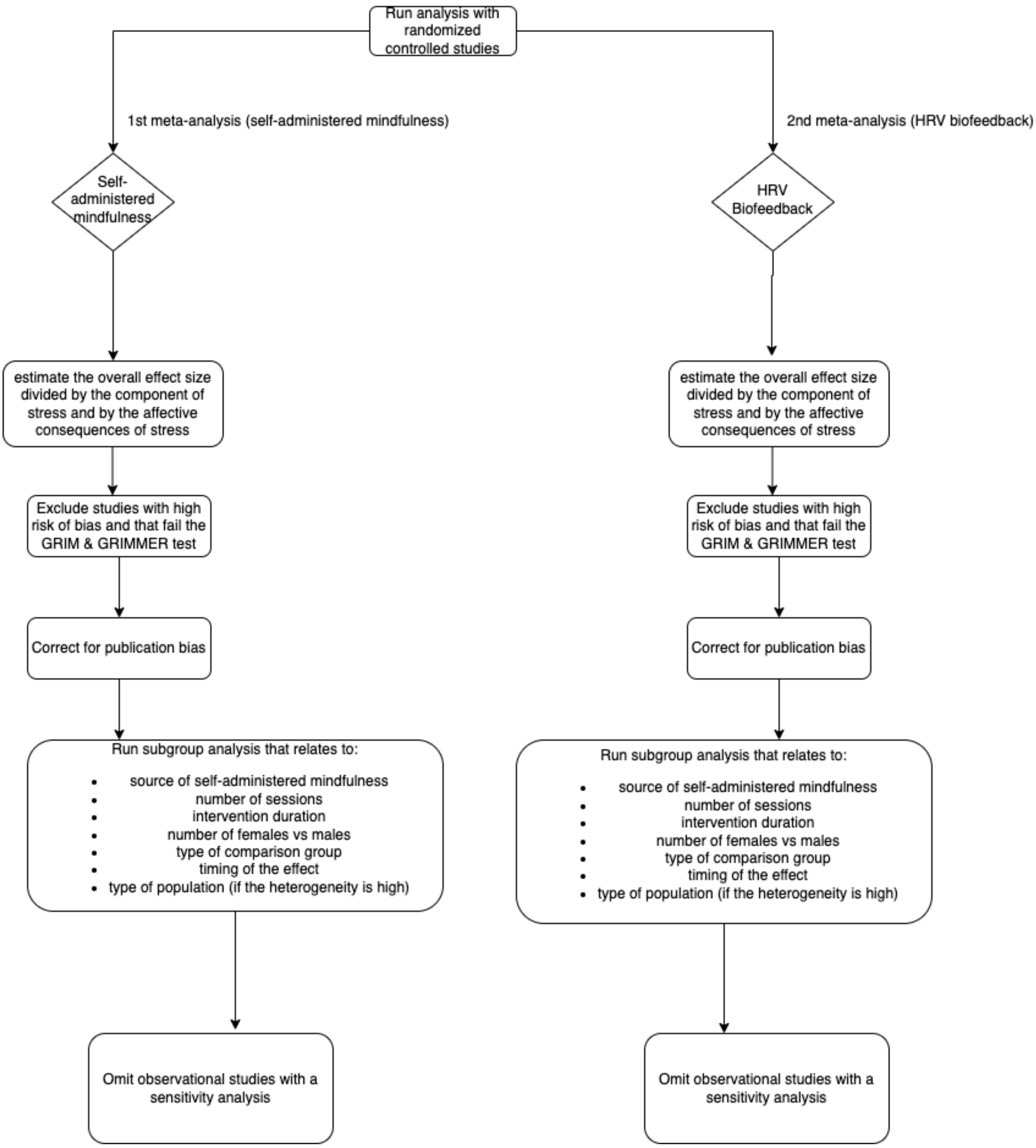
Search date: XX
Results: XX
Notes:

SCOPUS

(heart_rate_variability_biofeedback OR HRVB OR respiratory_sinus_arrhythmia_biofeedback OR RSA_biofeedback OR resonance_frequency_feedback OR RFF) AND stress AND (affect* OR emotion* OR cogniti* OR ruminati* OR physiological* OR cortisol OR heart_rate)

Search date: XX
Results: XX
Notes:

Appendix P: Analysis workflow



Appendix Q: Correction for publication bias

Procedure:

1. Conversion of effect sizes, calculation of variances and p-values
2. Data screening and outlier diagnostics
3. Test of evidential value → p-curve analysis using a permutation procedure
4. Estimation of a naive meta-analysis effect and effect heterogeneity: Hierarchical random-effect meta-analytic model using RVE
5. Correction for publication bias
6. Bias-adjusted test for the presence of an effect: 4-parameter selection model (4PSM) acting as a conditional estimator for PET-PEESE
    - 4PSM significant: Biased-adjusted estimation on the effect → Estimation by hierarchical implementation of PEESE using the RVE
    - 4PSM non significant: Biased-adjusted estimation on the effect → Estimation by hierarchical implementation of PET using the RVE

Appendix R: Study Protocol for Self-Administered Mindfulness and Biofeedback

The protocol below is based on our review of the literature as well as our meta-analytic synthesis. Despite that we don't find gender to moderate the effects for self-administered mindfulness, to be cautious (and because the literature contains many low-quality studies) we do recommend recording it in future studies.

Minimum Requirements:

1. Record the length of the session, the number of sessions asked for, and the type of intervention engaged in (e.g., for self-administered mindfulness, the body scan).
2. Record the medium via which the intervention is administered.
3. Include both a passive and active control group.
4. Record the time between the intervention and stress measurement.
5. Let participants fill in the Five Facet Mindfulness Inventory (De Bruin et al., 2012), the Trait Anxiety scale (Spielberger, 1983), the Experiences in Close Relationships Scale (Fraley et al., 2011) and the Big Five inventory (John, 1991).
6. Record participants' native language, sex, age, geographical origin.
7. Record whether participants practised mindfulness in the previous 6 months.
8. If participants are meditators, ask them how often and how long they typically meditate.
9. Record whether the population is from students or not.

Ideal Requirements:

1. Let participants fill in the Cardiff Anomalous Perceptions Scale (CAPS). On the basis of these measures, exclude participants that had a history of psychotic episodes and that suffered from depersonalisation and/or panic attacks. The measure can then also be used as a moderator to investigate the construct at subclinical levels.

2. Keep a diary during the intervention how often the participant (self)-administers the intervention.
3. If participants are meditators, ask them to keep a diary for a few weeks to record how often and how long they typically mediate.