Journal of Economic Behavior and Organization 217 (2024) 207–226

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Mindfulness training, cognitive performance and stress reduction

Gary Charness [a], Yves Le Bihan [b], Marie Claire Villeval [c,d,*]

[a] Department of Economics, University of California at Santa Barbara, United States
[b] Institut Français du Leadership Positif. 4 place Amédée Bonnet 69002 Lyon, France
[c] Univ Lyon, CNRS, GATE UMR 5824, F-69130 Ecully, France
[d] IZA, Bonn, Germany

ARTICLE INFO

JEL codes:
C91
I12

Keywords:
Mindfulness
Attention
Cognition
Stress
Lab-in-the-field experiment

ABSTRACT

Improving cognitive function and reducing stress may yield important benefits to individuals' health and to society. We conduct an experiment involving a three-month within-firm training program based on the principles of mindfulness and positive psychology at three large companies. We find an improvement in the difference-in-differences across the training and control groups in all five non-incentivized measures and in seven of the eight incentivized tasks. but only the non-incentivized measures and one of the incentivized measures reached a standard level of significance (above 5 %), showing strong evidence of its impact on both reducing perceived stress and increasing self-reported cognitive flexibility and mindfulness. At the aggregate level, we identify an average treatment effect on the treated for the non-incentivized measures and some effect for the incentivized measures. Remarkably, the treatment effects persisted three months after the training sessions ended. Overall, mindfulness training seems to provide benefits for psychological and cognitive health in adults.

1. Introduction

Cognitive function and performance are critically important elements of health and vitality.[1] The Diagnostic and Statistical Manual of Mental Disorders published by the American Psychiatry Association (American Psychiatric Association, 2013) defines six areas of cognitive function which are important to general health. They include executive function, learning and memory, perceptual-motor function, language, complex attention, and social cognition. It is often felt that a healthy mind leads to better physical health ("Mens sana in corpore sano").[2] The practice of mindfulness is considered by some to contribute to cognitive health by helping one fine-tune one's focus, thereby improving sleep, and reducing stress and anxiety. Thus, we might expect positive outcomes for participating in mindfulness trainings. We tested whether a mindfulness and positive-psychology intervention, conducted with employees at their workplaces, improves outcomes for both psychological health conditions measured by questionnaires and cognitive performance on a selection of incentivized lab-in-the-field tasks.

* Corresponding author at: Univ Lyon, CNRS, GATE UMR 5824, F-69130 Ecully, France.
E-mail address: villeval@gate.cnrs.fr (M.C. Villeval).

[1] For example, the National Institute on Aging, 2023 states: "In general, staying active is known to lower the risk of high blood pressure, stroke, and symptoms of depression, all of which in turn can improve cognitive health." It defines cognitive health as the ability to clearly think, learn, and remember (https://www.nia.nih.gov/health/cognitive-health-and-older-adults; consulted on November 9, 2023).

[2] The pre-Socratic philosopher Thales wrote : "What man is happy? He who has a healthy body, a resourceful mind and a docile nature." Longitudinal studies have identified strong relationships between mental health and physical health (*e.g.*, Ohrnberger et al., 2017).

https://doi.org/10.1016/j.jebo.2023.10.027
Received 11 January 2023; Received in revised form 30 August 2023; Accepted 22 October 2023
Available online 21 November 2023

Mindfulness training programs have been recently explored by firms in response to the increasing stress induced by modern life and work requirements, with the aim of reinforcing employees' ability to use their positive energy to cope with stress, reduce mental strain, and to improve wellbeing and self-regulation.[3] Inspired by Buddhist philosophy, mindfulness focuses on the monitoring and awareness of immediate conscious experience through non-judgmental attention to internal and external stimuli and improved attentional control (Kabat-Zinn, 1990).[4] If effective, such interventions might not only increase individuals' wellbeing, but might also have beneficial effects on companies' performance and on public-health budgets.

Mindfulness training can complement the tools inspired by positive psychology, seen as a "science of positive subjective experience, positive individual traits, and positive institutions" (Seligman and Csikszentmihalyi, 2000). Positive-psychology interventions within organizations aim at reducing work stress, cultivating valued subjective experiences, developing character strength, and improving engagement and wellbeing (for meta-analyses, see, *e.g.*, Meyers et al., 2013, Donaldson et al., 2019). Thus, through its aim of developing the cognitive ability to focus more of one's thoughts on the present and cultivating a state of awareness, mindfulness training has a positive potential and can be used as a complement of positive-psychology training that focuses more on flourishing, enhancing human thriving, developing a positive attitude, and reducing negative emotions.

Our aim was to go beyond the existing literature and further examine the impact of mindfulness on psychological health and cognitive performance under the relatively strict practices of experimental economics. Previous studies on the effects of mindfulness practice and positive-psychology interventions on cognitive performance have typically been conducted without monetary incentives and relatively frequently with no control group (for example, only 61 % of the studies included in the meta-analysis of Eby et al. (2019) include a control group). In the case of eliciting attitudes and feelings, one can only use self-reported responses. However, in the case of decision-making using incentives, associating the rewards given to the participants with their outcomes is considered by experimental economists as a gold standard to increase attention and engagement in the tasks, to reveal the individuals' preferences, and to avoid hypothetical biases (e.g., Smith, 1976; Camerer and Hogarth, 1999; Camerer and Mobbs, 2017). Therefore, we provide incentivized tasks as measures of cognitive performance. We also included a control group because having a control group is needed to measure differences-in-differences that can identify the effects of an intervention on the evolution of behavior over time separately from other sources of evolution, which guarantees the internal validity of a study.

Given that our motivation for this study was the idea that mindfulness shows promise in terms of improvements in psychological states and in cognitive performance, we also wished to take this outside the usual student setting used in laboratory experiments and have working adults at their workplace be participants in the experiment. Our view is that this is the best way to identify effects in a setting with relatively high external validity. We collaborated with a training institute for in-company mindfulness and positive-psychology programs, collecting data from both incentivized and non-incentivized tasks at three large global companies in France (two in the pharmaceutical industry and one in the energy sector) that implemented a eight-session training program in mindfulness and positive psychology before the COVID-19 pandemic. Our team includes a professional trainer with a great deal of experience in conducting such trainings.[5]

We assessed the causal impact of this program on various cognitive and non-cognitive abilities of individuals by means of difference-in-differences (DiD) comparisons. Our primary *ex-ante* hypothesis was that, by helping to focus the mind and cultivate a state of awaraness, mindfulness training will improve cognitive performance in a variety of both incentivized and non-incentivized tasks. We measured the evolution of the participants' incentivized performance and self-reported assessment across three moments in time (before the training started, at the end of the three-month training program, and three months after the end of the training program), and we compared these changes across the Training and the Control groups. Eight incentivized tasks, described and motivated in Section 3.3.1, were selected to assess the impact of the training program on decisions in settings involving strategic reasoning, attention, and/or concentration. Five non-incentivized self-assessments were included, described in Section 3.3.2.

It is important to evaluate the effects of such interventions since they may affect policy decisions. As an example, although 23,000 children had already received mindfulness meditation training at school, the French government rejected a large-scale mindfulness meditation program in primary and secondary schools in February of 2022, in part because it considered the scientific proof of the effectiveness of such programs to be very weak. We do present new evidence regarding the effectiveness of such interventions on adults in a study in the field. Of course, ours is far from the last word on this issue, and we hope that other studies will continue to explore the benefits of mindfulness practice for performance and stress reduction. Note also that we did not intend to disentangle the effects of the mindfulness intervention from those of the positive-psychology tools, since these were always used together during the training.

We estimated the difference-in-differences for performance in our Training and Control groups on our incentivized tasks. Seven of these eight measures moved in the predicted direction of improvement. We found statistically-significant improvement only for one

[3] In the U.S., large companies like Ford, Intel, Google, or Target have introduced such programs (see Eby et al., 2019). In France where we conducted our study, a survey of Initiative Mindfulness France, 2020 identified more than 60 mindfulness programs including more than three sessions, implemented in large companies including Air France, BNP Paribas, Danone, EDF, L'Oréal, MAIF, SAP, and Schneider Electric.

[4] Mindfulness programs were developed first as a cognitive therapy for patients, notably to reduce psychological disorders and stress, with validated benefits in the treatment of depressions or post-traumatic stress disorders, reduction in the risk of relapse in several illnesses like breast cancer, and improved wellbeing of patients with chronic diseases (for a recent meta-analysis on health effects, see Gotink et al. (2015) and Zhang, Lee et al. (2021).

[5] Crucially, to avoid conflict of interests we agreed from the beginning of the study that this team member would not have access to the dataset during the study and would not contribute to the data analysis. Moreover, the funding of the experiment to pay the participants was provided exclusively by the University of Lyon, independently from the French Institute of Positive Leadership and from the companies in which the training was implemented. There has been no (and there will be no) financial flows between the researchers and with this institute or the companies.

incentivized task (the Race-to-60, a game played against the computer that captures strategic sophistication) when we analyzed each game separately. Nevertheless, a one-tailed (directional) binomial test gives $p = 0.035$ for seven or more of these measures going in this direction. We do feel that the consistent directional effect is meaningful. In addition, the average treatment effect on the treated (ATET) estimates showed a causal positive effect of the training on two composite indices of performance at the aggregate level. Conducting our experiment on a population of mature adults at their workplace limited the number of participants that we could enroll in the study.[6] We acknowledge that the number of participants in our study (149), while large compared to most previous studies,[7] might be insufficient for yielding statistical significance for our relatively short intervention; in this respect our study is exploratory rather than conclusive.

We observed significant impacts on self-reported psychological dimensions, with all five measures moving in the predicted direction (one-tailed binomial test, $p = 0.031$). Differences-in-differences and ATET in the aggregate indices for these measures are significant at the 5 % or 1 % level. Moreover, in contrast with the incentivized measures, we found a significant causal effect of the training for most individual measures. In particular, the participants from the Training group reported a major reduction in perceived stress levels, a desirable change.[8] Other strong improvements were seen in the scores for cognitive flexibility and mindfulness when compared with participants from the Control group. These effects survive corrections for multiple-hypothesis testing and for attrition. We saw no benefits for patience. We did find improvement in resistance to change but this effect was not always significant and did not survive a correction for multiple-hypothesis testing. An important finding is that when an effect was significant, the effect typically persisted three months after the end of the intervention.

Overall, the picture is thus nuanced. We did find highly-significant beneficial effects of training for non-incentivized tasks. Yet, while we indeed found beneficial effects for incentivized tasks in the aggregate, these were typically not significant when considering each task in isolation. Our sense is that a longer (or more intensive) training would be needed for significance largely across the board. In any case, we see several possible reasons why we found stronger effects for the non-incentivized than for the incentivized measures. One is that by introducing monetary incentives in some tasks, the level of attention of all the participants (trained or not) increased; this may have hidden a training-induced difference in attention capability across the two groups. Alternatively, training may improve performance in all tasks, but the effect may be counterbalanced by a training-induced reduction in the value assigned to the monetary rewards (a reduced motivation). Another possible reason is a declarative bias. Individuals who spent energy to attend the training program (which increased time pressure on the regular workload) and to exercise between the training sessions may tend to exaggerate the benefits, especially if they wanted to reciprocate the company's effort in providing the training. Another possibility is that it is more difficult to improve cognitive abilities than psychological states. Further investigations would be needed to explore these factors.

Our exploratory study is the first in economics to analyze the impact of in-company mindfulness training on decision making. More generally, economists conducting health-focused interventions in the field have focused their attention on studying the effect of introducing monetary incentives to either encourage exercise and diet or to discourage smoking. In terms of our contribution to the literature, we feel that our contribution is fourfold: First, our strong questionnaire-based results contribute to the literature regarding the impact of mindfulness and positive-psychology training on mental health and stress. Second, we bring novel behavioral findings on the impact of mindfulness training at the workplace, based on incentivized tasks, to the literature in psychology and organizational behavior. Third, from a methodological standpoint, we demonstrate the effectiveness of conducting studies at firms. This setting offers a more diverse sample of participants than studies conducted on student subject pools and a modest attrition rate, which is a major improvement over that in most studies of which we are aware, and provides more confidence in our results. Finally, our study provides support for mindfulness and positive-psychology training in a controlled experimental setting having strong benefits in terms of psychological states and less-conclusive benefits in terms of incentivized cognitive effects. This seems sufficiently promising to offer a foundation for future studies.

In the remainder of this paper, Section 2 presents a brief literature review, while Section 3 describes our experimental design and the procedures. Section 4 introduces our conjectures, while Section 5 develops our results. Section 6 discusses these results and concludes.

2. Literature review

Recent studies in psychology and management have investigated the effectiveness of such mindfulness and positive-psychology

[6] One obstacle is that adding participation to the study to the training program is costly for companies since it increases the participants' unproductive working time. Furthermore, the COVID-19 pandemic did not allow us to involve more companies in the study.

[7] None of the studies listed in the meta-studies by Eby et al. (2019) or by Vonderlin (2020) show more participants than 149 (0 of 14 in the former and 0 of 14 in the latter), although some of the experiments in Hafenbrack and Vohs (2018) have similar numbers of participants.

[8] Stress can at times be a motivator to excel. Yet, individuals have different capabilities for managing stress. While excessive stress has deleterious effects – beyond the psychological and physical suffering and risk of burn-out it generates in individuals. It can also have harmful consequences on companies by impairing productivity and creativity (*e.g.,* Indhumati and Thirumakkal, 2015), by increasing absenteeism (*e.g.*, Brunner et al., 2019), and by increasing the risk of making errors through reduced attention (*e.g.,* Sänger et al., 2014). According to the American Psychological Association, workplace stress is estimated to cost more than $500 billion dollars to the U.S. economy. According to surveys conducted by the European Foundation for the Improvement of Living and Working Conditions (2010), more than 22% of European workers reported suffering from stress, muscular pain, and fatigue.

interventions at the workplace, finding some positive results. However, these previous studies very rarely tested decision-making in games and, from a methodological perspective, they typically involved self-reported and non-incentivized behavioral measures, with small sample sizes, often no control groups, and only using cross-sectional data.[9,10,11] While these are informative and insightful, one cannot exclude the possibility that the individuals who exerted high effort to attend these programs and exercised between training sessions tend to exaggerate their positive effects to self-justify the cost of effort.

A meta-analysis in the *British Medical Bulletin* (Zhang et al., 2021) cites the need for more high-quality studies with larger sample sizes and long-term follow-up. Ours is one of the first studies to ever measure incentivized performance in conjunction with any mindfulness technique; we do have a control treatment, and we also follow up three months after the training period. We take mindfulness to the laboratory-in-the-field, using the techniques of experimental economics.

Our approach is complementary to the previous literature, as we study the impact of mindfulness and positive-psychology interventions at work on both the quality of decision-making in standard incentivized economic tasks that require focus and attention and on self-assessed perceived stress, resistance to change, cognitive flexibility and mindfulness.[12] Since two crucial dimensions of mindfulness training are the effects on conscious attention and cognitive control (*e.g.*, Brefczynski-Lewis et al., 2007; Tang et al., 2007; Lutz et al., 2009; MacLean et al., 2010; Mrazek et al., 2013; Good et al., 2016) and the regulation of emotions, we conjectured that receiving such training would increase individuals' performance in our tasks compared to a sample of non-trained participants, in relation to improved attention, concentration, and self-control.

Studies in psychology (*e.g.,* Friese et al., 2012; Teper and Inzlicht, 2013; Teper et al., 2013; Flook et al., 2010) have explored the effects of mindfulness on executive function. Wenk-Sormaz (2005) and Moore and Malinowski (2009) found that meditation practice improves performance on the Stroop task. Participants improved at attentional switching after completing a 10-day intensive meditation retreat (Chambers et al., 2008). These studies reflect various aspects of executive functioning, providing evidence on the connection between mindfulness practices and executive function. Mindfulness has also been found to encourage divergent thinking and creativity (Meier et al., 2020; Montani et al., 2020). However, Hafenbrack and Vohs (2018) find at most limited support for the effectiveness of mindfulness training on performance although they present persuasive evidence to support that mindfulness might diminish one's motivation.[13]

Considerable research has indicated that mindfulness-based approaches can reduce chronic stress (Teasdale et al., 2000; Ma and Teasdale, 2004) and anxiety (Hofmann et al., 2010). Carmody and Baer (2008), Carmody et al. (2009), and Epel et al. (2009) found reductions in perceived stress that persisted at one- to three-month follow-ups after the end of participation. Krusche et al. (2013) found significant reductions in perceived stress, anxiety, and depression upon completion of an online mindfulness course, as well as a further decline at a one-month follow-up. However, this study had no control group and involved individuals who paid for the course, so there could be serious selection bias. In fact, a careful review of meditation programs (Goyal et al., 2014), including 47 randomized clinical trials with active controls, found little evidence of the effects of mindfulness on eating habits, sleep, and stress reduction, suggesting that improved study designs are needed to identify any impact of mindfulness programs on stress and anxiety. Allen et al. (2015), Good et al. (2016), Eby et al. (2019), Vonderlin et al. (2020), and Zhang et al. (2021) also provide useful surveys and meta-analyses.[14]

When focusing on work environments, training in mindfulness has been found to regulate emotions in professions with high occupational stress (*e.g.*, Glomb et al., 2011; Hülsheger et al., 2013; Chin et al., 2019; Grupe et al., 2019). While some studies identified a positive effect on work relationships (Reb et al., 2014; West et al., 2014; Liang et al., 2016), Roeser et al. (2013) found it reduces the

[9] The targeted outcome of the programs was stress and mental strain in 81% of the studies reviewed by Eby et al. (2019); most investigations focus on the measurement of perceived stress and affect, and sometimes attention to response tasks, or working memory. There are studies investigating individual and social decision making in experimental settings but not when mindfulness training was implemented at a company (*e.g.,* Hafenbrack et al., 2014, on the sunk cost bias; Kirk et al., 2011, on decisions in the ultimatum game; Lakey et al., 2007, and Zhang, Chen et al., 2021, on risk taking and impulsive gambling).

[10] Behavioral measures are completed in some studies with physical (*e.g.*, weight, blood pressure), electro- and psychophysiological measures (*e.g.*, skin conductance, heartbeat). Neuro-imaging has been used to investigate the neural effects of meditation practice, notably on cognitive control and emotional regulation in the brain.

[11] Only 27 of the 67 experimental studies assessing training programs based on mindfulness for employees reviewed by Eby et al. (2019) involved more than 50 participants; only 13 had more than 100 participants. Only 61% of these studies include a control group and 65% include only one observation point in time after completion of the training. 52% of the 172 studies reviewed by Goldberg et al. (2018) use no treatment comparison groups. Vonderlin et al.'s (2020) meta-analysis that includes 56 RCTs shows that most studies are cross-sectional and many focus on specific occupational groups (*e.g.,* health care professionals, teachers).

[12] One important difference between our study and the non-economic previous literature is that people were paid to attend (training was provided during working time) and for their performance on the tasks. Together with the social pressure stemming from participating in our study with colleagues, this had the effect of reducing attrition, which otherwise tends to be a serious problem. For example, attrition rates of 22.7% for health-care professionals and 34.9% for mental-health professionals who took a mindfulness course have been reported (Shapiro et al., 2005). Another report, a meta-analysis of mindfulness studies in nonclinical populations (Khoury et al., 2015), found attrition rates of 3% to 51%. In nonclinical populations, higher attendance rates are more common among college students who may have protected time and tangible incentives for attendance. But if a study has major attrition, this most likely leads to an upwardly-biased estimate of the effectiveness of the intervention.

[13] Hafenbrach and Vohs (2018) take this view of their results: "When combined with mindfulness's demotivating effects, these results help explain why mindfulness does not alter performance." In fact, their Figure 7 summarizes results for 14 tests; the only significant one is positive. Nine of the 14 show positive effects and five show negative effects, evidence suggesting an overall positive but insignificant improvement.

[14] For example, Vonderlin et al. (2020) found medium-to-high effects of interventions on reduced stress and improved job satisfaction, but it rejected any significant effect of interventions on productivity when outliers are removed from the sample, partly due to the small number of studies on performance, most based on self-reports.

risk of burn-out, therefore improving job satisfaction and productivity (Shonin et al., 2014; Lomas et al., 2017; Kersemaekers et al., 2018). Regarding positive-psychology interventions in a work environment, meta-analyses have identified a significant positive impact on well-being but ambiguous results on negative emotional states (Meyers et al., 2013). The meta-analysis of Donaldson et al. (2019) identified a small to moderate positive effect on job well-being and engagement, but also a small-to-moderate negative effect on stress and other undesirable outcomes at work. Very few studies evaluated interventions that combine mindfulness and positive-psychology training (Ivtzan et al., 2016; Giraud et al., 2022).[15]

Previous interventional studies in economics have identified heterogenous effects of monetarily incentivized programs on habits and health. Charness and Gneezy (2009) paid students to attend the school gym twice a week for a month, finding strong effects (including biometric data) on attendance rates after the intervention period. Charness and Jahnke (2012) introduced wellness practices in eight weekly sessions, finding improvements in pulse rates. Giné et al. (2010) considered cigarette cessation, with mixed results (smoking resumed after the intervention ended). Woerner et al. (2021) showed that monetary incentives to attend a daily meditation program increased meditation frequency, but that offering people a choice of incentive schemes decreased adherence. In any case, research on the impact of non-monetary health interventions on performance has not yet attracted the same interest.

Although economists have been interested in the role of attention in the allocation of scarce cognitive resources since the works of Herbert Simon (1957),[16] there have been very few studies on the impact of mindfulness training on cognitive performance in economically-relevant tasks requiring focus and attention. Alem et al. (2021) explored the effects of a four-week online mindfulness training on students' risk taking and intertemporal decisions, using economic games and both self-reported and revealed measures of health-related behavior. Training reduced perceived stress but had little impact on risk preferences, patience, present bias, cortisol levels, sleep, smoking, or alcohol consumption. Shreekumar and Vautrey (2021) evaluated the impact of a four-week experiment offering access to a meditation app. They found improvement in a proofreading task in the medium term (but also short-lived adverse effects on performance when negative emotions were involved) not only in terms of reduction of anxiety but also in terms of productivity, suggesting that the program helps in reducing the interference between emotional states and attention.[17]

Finally, Cassar et al. (2022) showed evidence of a large positive effect (0.4 standard deviation) of a standard eight-weeks MBSR[18] meditation training on students' academic performance in the long-run (eight months later) but a marginal negative short-term effect on grades just after the end of the program. The negative short-term effect is linked to more sleeping and relaxing, while the positive long-term effect is linked to more practice after the program's end. More evidence is needed to draw firm conclusions, and we present the evidence from our study in this spirit.

3. Experimental design and procedures

We first present the sample of participants. We briefly describe the training program and then the tests used in our study. Finally, we describe the procedures.

3.1. The sample

The experiment was conducted at three large companies located in France.[19] In total, 149 participants enrolled in the study – 77 in the Training group and 72 in the Control group. Details of the number of participants, by company and by stages, are provided in Appendix Table C1. All participants were volunteers who were actually or potentially motivated to engage in a mindfulness training program

[15] Like our study, Giraud et al. (2022) aimed at evaluating the impact of mindfulness and positive-psychology training on performance and well-being in a work environment, and the French Institute of Positive Leadership was involved in both studies. However, the two studies used different tasks (only the perceived Stress Scale is used in both), a different statistical methodology (a DiD analysis in our study), and different subjects.

[16] See, for example, Camerer and Johnson (2004), Falkinger (2008), Brocas et al. (2014), Rubinstein (2016), and Avoyan and Schotter (2020).

[17] See also the RCT study of Ash et al. (2021) that showed that a short mindfulness intervention can reduce the avoidance of information that may cause worries and regret.

[18] MBSR for Mindfulness-Based Stress Reduction, a program developed by Jon Kabat-Zinn, Professor of Medecine Emeritus at the University of Massachussets Medical School in 1979.

[19] BioMérieux is a global leader in *in vitro* diagnostics, listed on NYSE Euronext Paris. It produces diagnostic solutions (reagents, instruments) to identify the source of diseases and contamination. It is present in 45 countries, employs 14,000 persons and realized €3.6 billion in sales in 2022. Sanofi is a leader in R&D and manufacturing of pharmaceutical drugs, the world's fifth-largest by prescription sales, the world's largest producer of vaccines, and a component of the Euro Stoxx 50 stock market index. It is present in 90 countries, employs 91,000 people and realized €43 billion in sales in 2022. Engie is a multinational electric utility company, operating in electricity generation and distribution, natural gas, nuclear, renewable energy, and the first independent power producer in the world. It is present in 31 countries, employs 96,400 people and realized €93,9 billion in revenues in 2022.

offered by these companies.[20] The participants from the Control group did not attend the training program or received later training.[21,22]

The assignment to the Training or the Control groups was made by the firms, and so we had no control over randomizing the treatment assignment. In principle, this could be problematic. To assess the differences between the two groups, Appendix Table C2 provides descriptive statistics on the characteristics of the participants from both groups, and non-parametric statistics comparing their characteristics at each stage of the study. Reassuringly, Table C2 shows that the composition of the two groups at the beginning of the study did not differ in terms of gender, age, education level, sector of activity in the company, holding a managerial position, having supervisory responsibilities, income category, prior experience of meditation, motivation for the training program, or time budget for mindfulness (we asked participants in the Control group to consider a hypothetical enrollment in the program).[23] The vast majority of all participants were graduates and held a managerial position. Moreover, as detailed in the results section (see the first three columns of Tables 3 and 4), the participants' average performance in incentivized and non-incentivized measures at the beginning of the training program were never significantly different between the two groups. So, based on these observable characteristics, the allocation of individuals to the treatments seem to be balanced.

Substantial attrition rates could readily bias effect sizes (most likely upwards), perhaps a factor in some of the previous positive results reported outside of economics. Ours are relatively low. As detailed later, our experiment comprised five stages. Among the 149 participants in stage 1 (77 and 72 in the Training and the Control groups, respectively), 146 attended stage 2 (76 and 70, respectively), 142 for stage 3 (72 and 70, respectively), and 130 for stage 4 (68 and 62, respectively). The total attrition rate was 4.70 % (7/149) in stage 3 and 12.75 % (19/149) in stage 4 (the two stages that correspond to the end of the training program). In addition, 124 participated in stage 5 (66 and 58, respectively, for a modest 16.78 % (25/149) attrition rate even in stage 5, a full three months after the end of the training program).[24]

As explained below, the most important comparisons are the scores in stages 1 and 3 for incentivized measures (142 comparisons), and the scores in stages 2 and 4 for non-incentivized measures (130 comparisons). This gives enough statistical power to detect a medium-size treatment effect on performance in our tests.[25] The low attrition rate most likely reflects the training having been conducted in-company, with possible pressure from colleagues and repeat reminders (sent to all) to take part in all stages. Participants were also informed their earnings from the tasks would be paid at the end of stage 5, conditional on participating in all stages.

Appendix Table C2 shows the limited consequences of attrition on the comparability between the two groups over time. The mean characteristics of the Training and Control groups are almost never significantly different in any stage. The only exception (but significant only at the 10 % level) is the motivation for the program in stage 3. A complementary way to assess attrition consists of comparing the characteristics of those who quit in the Training group and those who quit in the Control group, since this could be a potential source of bias. Table C3 in Appendix C indicates that the effects of attrition across groups are quite modest and mainly related to education, and that this is driven by very few observations (two versus five observations in stage 3). Therefore, we believe that attrition does not impair the comparability of the groups, and we conduct our main data analysis on the observations for which we have a full data set. Nevertheless, we report in Appendix Table C5 additional difference-in-differences analyses that account conservatively for attrition.

[20] We did not use the terms meditation, mindfulness, or positive psychology in the invitation letter, the consent form, or the instructions used in the different sessions. We always talked about participating in a "scientific study on decision-making".

[21] The Control group participated in the same experimental tasks as the Treatment group. They played exactly the same games and responded to the same questionnaires at the same moment although not in the same sessions. The only difference is that they did not participate (or not yet) to the training program. They were informed about the existence of the training program before the beginning of the experiment and most of them were volunteers to participate in the training program in the future. A small fraction of them actually attended later on because the COVID pandemic started just after the end of the intervention in the third company.

[22] Note that it would have been interesting to include an active Control group in the study to limit unobserved differences between the control and the treated groups. For example, another control group could have received an intervention keeping them out of work for the same number of sessions as our treated group, at at the same dates, but using neither mindfulness nor positive psychology training, and not intending to improve cognitive health. However, it would have been difficult to justify to the companies and to design an acceptable neutral program.

[23] We do acknowledge that this possibly non-random assignment of the participants to the two groups is a limitation. We take this selection into account in our Difference-in-Differences analysis (see section 5).

[24] The attrition rates are close in the two groups. In stage 3, they are 6.49% (5/77) in the Training group and 2.78% (2/72) in the Control group. In stage 4, the rates are, respectively, 11.69% (9/77) and 13.89% (10/72), and in stage 5, 14.29% (11/77) and 19.44% (14/72).

[25] We defined the required sample size such that we expected to achieve significance at 5% and power at 95% to detect a medium-sized effect on scores ($d = 0.5$) in one-tail tests, using Mann-Whitney U tests and taking each individual as a single observation. This determined an initial objective of approximately 100 participants in each group to account for attrition. However, because of the pandemic of COVID-19 starting in 2020, the in-company mindfulness training programs had to stop, and we have not been able to reach this target. Post-hoc calculations show that based on the number of participants still present in stage 3, we have a power of 89% to detect an effect of 0.5 in our incentivized measures.

3.2. The training program

The training program comprised eight sessions of two hours on average at the three companies' headquarters, one every two or three weeks. These in-company sessions were organized during work time.[26] The training was led by senior trainers, experienced positive psychologists who had been practicing and teaching mindfulness meditation for several years.

The training program was based on mindfulness meditation and positive psychology. It included didactic teaching, training with positive-psychology tools, training in mindfulness meditation (including two to three guided mindfulness meditative exercises per session), and individual and collective exploratory dialogues (about feelings, perceptions of the practice, Q&A), supervised and led by the trainers. Daily individual at-home and at-work meditation and positive-psychology practices, as well as practice in pairs, were encouraged between sessions during the training period and afterward. Educational materials were provided to participants (including handouts, meditation audio recordings, training guide, and theoretical inputs).

The content of the training program was the same in the three companies. The covered topics included attention to self, attention to others, mobilization of character strengths and positive emotions, vulnerability and mutual aid, gratitude, self-compassion, resilience, and emotional flexibility (see details of each training session in Appendix A). Three dimensions of the training were expected to have an impact on the quality of decision-making. First, by increasing awareness to one's negative mental patterns, mind wandering, and cognitive biases, the program tried to cultivate the quality of attention, attentional control, and self-awareness. Second, the program aimed at strengthening cognitive flexibility through openness to alternative options, disconnecting automatic reactions to stimuli, promotion of divergent thinking, and by training various techniques (such as Respond versus React, and focus of attention to mind, emotions, and body sensations in the present moment). Third, the program tried to develop pro-social behaviors, by cultivating gratitude and mutual aid, thereby strengthening the importance of taking others into account in one's reasoning.

3.3. The tasks

Our study consisted of a set of computer-based tasks, programmed in Java. Participants had to perform eight financially-incentivized tasks and to fill out five questionnaires. Details of the instructions are provided in Appendix B and screenshots of the main tasks are in Appendix D. For the incentivized tasks, we selected a battery of tasks that engaged the participants' cognitive sophistication and allowed us to study decision-making in settings involving strategic reasoning, attention, and/or concentration. There is of course some arbitrariness in how we composed this bundle. Our objective was to use tasks for which participants were likely to have had limited previous experience to avoid an uncontrolled source of differences between subjects. Another objective was to use various tasks that we deemed relevant to test if the training program improved cognitive performance in general or some facets of it (cognitive reasoning vs. attention). For the non-incentivized tasks, we selected questionnaires that could measure dimensions that were targeted by the training program (for example, a questionnaire allowing us to measure a reduction of stress). As far as we are aware, none of our incentivized tasks and only a few of the non-incentivized tasks (for example, the perceived stress scale and mindfulness scale) have been used in previous studies on the effects of mindfulness training.

3.3.1. Financially-Incentivized tasks

These tasks include the Cognitive Reflection Test (Frederick, 2005), the Tower of Hanoi puzzle, the Race-to-60 game (Dufwenberg et al., 2010; Gneezy et al., 2010; Bosch-Rosa et al., 2018), Raven matrices used in tests of fluid intelligence (Carpenter et al., 1990) and self-confidence in this task, a test of Theory-of-Mind (the Reading-In-the-Eyes test of Baron-Cohen et al., 2001), sums of numbers to find in matrices (Mazar et al., 2008), and a simple spot-the-difference game of attention. These included four cognitive-sophistication tasks.

First, participants had to answer the three-item Frederick (2005) Cognitive Reflection Test, in which one must toss out an intuitive-but-wrong answer and engage in non-automatic thinking to find the correct answer. Each correct answer paid €2. Performance in this task has been found to correlate with IQ. Second, participants attempted a non-verbal logic puzzle (the Tower of Hanoi, invented by Édouard Lucas in the 19th century). The computer screen shows three rods and three disks of various diameters. Initially, the discs were stacked on the leftmost rod with the largest disk at the bottom and the smallest one at the top. The objective was to move the whole stack onto the far-right rod in less than three minutes. A larger disk could not be slid over a smaller disk, that is, the discs could only be moved, one at a time, onto an empty rod or on top of a larger disk. Solving the puzzle paid €5. The minimal number of moves to solve the puzzle is seven. Participants could make as many moves as they liked but each move entailed a cost of €0.30 that was deducted from these €5 (but earnings could not be negative). This task also engages cognitive reflection.

Third, strategic sophistication was measured in six rounds of the Race-to-60 game (on race games, see, e.g., Dufwenberg et al., 2010 and Gneezy et al., 2010; Bosch-Rosa et al., 2018) played sequentially against the computer with one minute per round. The participant had to choose a number between 1 and 10, inclusive, and then the computer also chose a number in this range. This alternating process continued and a running count was kept on the sum of all chosen numbers. The first to reach at least 60 won; participants were paid €1 for each game won. This game is solvable by backward induction and has a simple dominant strategy with steps of 60, 49, 38, 27, 16

[26] It could be argued that the effects we are measuring are driven by being away from work for a few hours and not by the content of the program itself. Nevertheless, most participants are paid for doing their job and not for working a given number of hours. If any, the effect of the training program was to create additional pressure on the workload, which would go in the opposite direction of the expected effect of the intervention on notably stress.

and 5. Unknown to the participants, the computer picked numbers randomly in the first three rounds, except that the computer picked the number to reach 60 if the sum was at least 50; in the last three rounds, the computer's sophistication was increased to also make the sum 49 if the previous sum was between 39 and 48. This game helps identify if players are able to plan in a sequential choice problem.

Fourth, participants had six minutes to solve six advanced progressive Raven matrices (two simple and four of medium difficulty), a non-verbal test usually used to test fluid intelligence (see, *e.g.*, Carpenter et al., 1990). We used these matrices as a cognitive task without any intention to determine the IQ of the participants which would have required many matrices (typically 48). For each question, a series of figures were displayed on the participant's screen; the task was to identify which of eight possible figures suggested on the screen logically followed. Each accurate answer about the missing figure paid €1 but no feedback was given to the participants on whether they had provided a correct answer. Then, participants estimated their absolute performance and their relative performance (their rank compared to the performance of six other randomly selected participants). Each accurate guess paid €1. The objective was to measure the individuals' level of self-confidence. Overconfidence in the estimation of one's score constitutes our fifth incentivized measure.

Two other tasks directly involved attention. The Find-10 task that participants played for two minutes is similar to the matrix task of Mazar et al. (2008). Six 3 × 4 tables were successively displayed on their screen, with a number with one decimal place in each cell. Participants had to click on the two numbers in the table whose sum was 10. Each accurate answer paid €1. A table could be passed (with no option to return later) and for each table passed, participants earned a fixed payoff of €0.20. Finally, in the Spot-the-Differences task, the participants' screen displayed two identical figures with contained 100 similar items, except for ten differences. Participants had to click on as many differences as possible in 90 s. Each difference spotted correctly paid €0.50. Performance in these two tasks mainly requires concentration and attention.

The eighth incentivized task was the Reading-in-the-Eyes task (Baron-Cohen et al., 2001; Prevost et al., 2014, for the French version), which is a standard measure of differences in Theory-of-Mind capabilities, that is the ability to infer others' mental states and intentions. Thirty-six images showing pairs of eyes were successively displayed on the screen, with four response options describing the target person's emotions or intentions. Participants had to guess what emotion was expressed in each eye gaze, by choosing which of the four options best described the emotion that the eye gaze was expressing. Each accurate answer paid €0.20.

These cognitive and attention tasks allowed us to test whether the training could improve performance by increasing concentration, self-regulation, attention to self and others, emotional stability and, in the last task, by increasing social attention and cognitive empathy.

3.3.2. Non-Incentivized tasks

In addition to these eight monetarily-incentivized tasks, participants filled out five non-incentivized questionnaires. One questionnaire elicited the participants' time preferences. We used both the qualitative measure of patience and the staircase questionnaire from the Falk et al. (2018) Global Preferences Survey. The qualitative measure of patience was based on the response to the question "To what extent are you willing to give up something that would be beneficial for you today in order to benefit more in the future?" on an 11-point scale. In the staircase method, participants made five binary decisions between an immediate reward and a larger reward in twelve months. The immediate hypothetical payment was kept constant in each of the five situations, but the delayed payment increased or decreased in each situation, depending on the previous choices. At any step, participants had the option to make no choice. The method allowed us to identify the indifference point between an immediate-but-smaller reward and a delayed-but-larger reward. We included this dimension because patience usually correlates with a range of life outcomes (*e.g.*, Chabris et al., 2009).

Four surveys were psychological questionnaires. We measured perceived stress by adding the scores on the Cohen et al. (1983) 10-item 5-point scale Perceived Stress Scale (PSS). This assesses how situations that happened in one's life in the past month are appraised as stressful (for example, "In the last month, how often have you found that you could not cope with all the things that you had to do?"). This includes two sub-scales: perceived self-efficacy and perceived helplessness.

We added a test of cognitive flexibility to collect a measure of the mental ability to adapt to a new environment or new actions, to shift attention to different tasks, and to control one's thoughts. Cognitive flexibility was measured by means of the 12-item 6-point scale test of Martin and Rubin (1995). This test assesses the individual's awareness of alternative solutions, willingness to adapt to a new situation, and self-efficacy in being flexible (for example, "I can find workable solutions to seemingly unsolvable problems"). We included this dimension because most managerial tasks request such flexibility, and some studies (*e.g.*, Moore and Malinowski, 2009) have suggested that mindfulness could improve this capacity.[27]

Mindfulness (quality of presence) was measured by means of the 39-item 5-point scale Five Facet Mindfulness Questionnaire (Baer et al., 2006; French version by Heeren et al., 2011). These facets are: observing and attending to thoughts and feelings, describing and labeling with words, acting with awareness and concentration, non-judgment of inner experience, and non-reactivity to inner experience. This test measures the ability of an individual to focus his or her attention to experiences in the present moment, in a non-judgmental and accepting way (for example, "It seems I am 'running on automatic' without much awareness of what I'm doing"). Finally, we administered a 6-item 6-point scale version of the Resistance to Change test (Oreg, 2003). This test assesses behavioral, affective, and cognitive aspects of an individual's resistance to change, such as routine seeking, emotional reaction to changes, short-term thinking and cognitive rigidity (for example, "When I am informed of a change of plans, I tense up a bit").

[27] Alternatively, we could have used an online creativity test but we would have lost control of the conditions in which subjects performed the test. Another possibility would have been to use the Wisconsin card sorting test or the Stroop test. However, we tried to limit the cognitive load of the participants in an online session and thus, a simple 6-item questionnaire was chosen.

3.3.3. Additional information

At the beginning and at the end of stages 1, 3 and 5, we used Self-Assessment-Manikins (Bradley and Lang, 1994) to elicit the participants' feelings of happiness and nervousness, with five figures varying in affective valence and intensity. This was only for exploratory motives. Participants filled out a brief socio-demographic questionnaire at the end of stage 1. In stages 3 and 5, we also asked questions about the practice of mindfulness exercises outside of the training sessions (see details in Appendix B).

Participants received no feedback on their scores and payoffs until the very end of the experiment, except in the Race-to-60 and the Spot-the-Difference tasks in which they could directly observe their success or failure.

3.4. Timeline and procedures

The experiment comprised five stages, two carried out in-person at the companies' premises (stages 1 and 3) and three conducted online (stages 2, 4 and 5).[28] Measures were collected for both Training and Control groups at three times: pre-intervention (stages 1 and 2), the completion of intervention about three months later (stages 3 and 4), and three months after the end of the intervention (stage 5).[29] An exception occurred in the first company in which the Control group started to receive the same training program just after stage 4.[30] Table 1 describes the content of each stage.

The experiment was computerized and programmed in Web language. The in-person sessions were introduced by a presentation of the general rules and timeline of the experiment (including anonymity, confidentiality of the decisions, payment of a fixed 7€ participation fee for each of the five stages, and about a variable payoff depending on the individual decisions made in each stage).[31] At the beginning of stage 1, all participants signed an informed consent. We asked participants not to speak with anyone about their responses in any stages, to preserve confidentiality and independence of observations. Next, participants received a tablet on which we presented the instructions for each part of the experiment, each at a time. Each participant was able to progress at his or her own pace, with the parts following each other automatically. Confidentiality was ensured by using our GATE-Lab mobile lab, setting up partitions separating each workstation (see pictures in Appendix E). Moreover, participants had to create a personal identifier (not linked to one's identity) for the whole experiment. As shown in Table 1, the content of stages 1 and 3 was identical and mainly comprised the incentivized measures. These stages lasted on average 45 min.

Participants could participate in the three online stages at their preferred time, day, and place in a time window of a maximum of four weeks (unannounced), although we insisted that they do it the week following the receipt of the invitation message (in the Training group this was a requirement in stage 2 to validate the data). Participants received several reminders to the groups in each company until (almost) everyone had participated. As shown in Table 1, the content of stages 2 and 4 was identical and mainly comprised the survey questionnaires (only the Theory-of-Mind test was monetarily incentivized). Stage 5 reproduced most of the tests in stages 2 and 4 (except the Theory-of-Mind test, in order to save participants' time) and it repeated one cognitive task (the Race-to-60) and one attention task (Spot-the-Differences) presented in stages 1 and 3. Each stage lasted 25 min on average.

The distribution of earnings was organized in two steps. At the end of each of stages 1 and 3 each participant received a €20 payment in cash or in gift coupons, which corresponded to the fixed participation fee of, respectively, stages 1 and 2, and 3 and 4, plus a small advance on variables earnings. The final payment was done after everyone had completed stage 5 at the companies' premises, staggered over several days to try to avoid earnings comparisons. To receive their earnings, participants presented their identifiers to the experimenter in charge of payment in exchange for a closed envelope containing their personal earnings. In total, participants earned on average €90.57 (S.D.= 13.27).

4. Conjectures

Before we discuss our conjectures and the experimental results, a remark about our ex-ante predictions, our analysis, and pre-registration is in order. We did not pre-register our experiment largely due to oversight. Both co-authors primarily conduct experiments in laboratory settings, where pre-registration is not customary. As well, the practice of pre-registration has not generalized in applied economics, and it has even been called into question by a number of researchers. For example, Coffman and Niederle (2015) make the point that pre-analysis plans have limited upside. John List, who has argued strongly for methodological improvements such as multiple-hypothesis corrections, acknowledges that pre-registration has not worked out the way that was expected.[32]

[28] In the third company, stage 3 had to be conducted online as well, but on the same day and at the same times for all the participants of the Training group or the Control group, because it was scheduled the first day of the lockdown imposed by the French government against the COVID-19 pandemic (March 16, 2020). The sessions were scheduled well before we were aware of the sudden social distancing regulations.

[29] These time lags differed slightly across companies for practical reasons. At the first company, the experiment took place between November 2018 and June 2019, at the second company between March 2019 and October 2019, and at the third company between November 2019 and June 2020.

[30] This means that this sub-sample cannot serve as a control for the long-term effect of the intervention.

[31] In the instructions, we repeated several times that the participants' responses would be anonymous throughout the experiment. We also mentioned that their data would never be communicated to their company. The objective was to limit the risk that the responses were motivated by the willingness to please the employer (or even the trainers).

[32] In his then-editorial role at the *Journal of Political Economy*, he wrote: "I am sympathetic to this overall movement, but my own work tells me that it has not been well executed in our profession, and authors who have papers now should just be upfront and place the results in the proper context." And that is what we do.

Table 1
Timeline and content of the experiment.

Timeline	Stage 1 Pre-intervention	Stage 2 Pre-intervention	Stage 3 End of intervention	Stage 4 End of intervention	Stage 5 3 months after intervention
Type	*In-person*	*On-line*	*In-person*	*On-line*	*On-line*
1. SAM Happiness - Nervousness	X		X		X
2. CRT test	X		X		
3. Hanoi Tower	X		X		
4. Race-to-60	X		X		X
5. Patience	X		X		
6. Raven matrices	X		X		
7. Confidence in score in Raven	X		X		
8. Find-10	X		X		
9. Spot-the-difference	X		X		X
10. Socio-demographic survey	X		X		
11. SAM Happiness - Nervousness	X		X		X
12. Reading-in-the-Eyes (ToM)		X		X	
13. Perceived Stress Scale		X		X	X
14. Cognitive Flexibility		X		X	X
15. Five Facet Mindfulness		X		X	X
16. Resistance to Change		X		X	X
17. Questionnaire on the training and meditation practice		X	X		X

We have always had the obvious hypothesis that mindfulness practice will improve performance on incentivized tasks and lead to better psychological outcomes. However, we had no initial plan for analysis other than testing for each improvement in each task. As shown in the results section, we also formulated indices for aggregating the results for overall performance. This was done post-hoc, after seeing the data. But we report the results honestly and in full detail, so that readers can draw their own conclusions. We most certainly welcome replications.

The literature has shown that mindfulness affects human functioning primarily through the stability, control, and efficiency of attention (see, *e.g.*, Lutz et al., 2009, and the Good et al., 2016, survey). Mindfulness training has been shown to also improve cognitive capacity (including working memory) and cognitive flexibility, as well as to reduce automaticity of behavior by fostering awareness of automatic actions (Good et al., 2016). By reducing mind wandering and improving attentional control (*e.g.*, Mrazek et al., 2013), the trained participants are expected to make better choices and fewer decision errors in the incentivized games, particularly in the cognitive sophistication and the attention tasks.

Also, since the mindfulness literature has shown beneficial effects for reducing stress and encouraging introspection, we expected that attending the training would decrease overconfidence for overconfident individuals.[33] Finally, since the literature has shown that mindfulness programs tend to strengthen other-orientation and reduce egocentric tendencies,[34] we might also expect better performance for trained participants in the Theory-of-Mind test, possibly through an increased focus of mindful attention on the interpretation of the emotions of others. This leads to our first general conjecture.

C1: Mindfulness training improves performance on the incentivized tasks.

Conjecture 1 can be decomposed into its different components.

Training improves performance in: C1a) the cognitive sophistication tasks (CRT test, Hanoi Tower, Race-to-60, Raven matrices), C1b) the attention tasks (Find-10 and Spot-the-Differences), C1c) reducing the overconfidence of individuals in terms of performance in the Raven matrices, and C1d) the Theory-of-Mind test.

Our second conjecture incorporates the psychology studies (Section 2) that showed positive effects of mindfulness training on stress, flexibility, presence, and acceptance of change. We therefore suspected that the training would improve the scores in these psychological questionnaires. We also expected a positive impact of training on patience. Mindfulness training might increase patience and reduce impulsiveness by improving executive function and by increasing attention to the relative price of sooner vs. later rewards.

C2: Mindfulness training improves performance in the non-incentivized tasks.

Conjecture 2 can also be decomposed into its different components:

Training decreases perceived stress (C2a) and resistance to change (C2b), and increases cognitive flexibility (C2c), mindfulness (C2d), and patience (C2e).

[33] We ignored under-confidence as we expected to observe very few cases, which was indeed confirmed in the data.

[34] Note that there is no consensus in the literature. For example, recent studies have called into question the extent to which mindfulness strengthens other-regarding orientation (Hafenbrack et al., 2022) or reduces egocentric tendencies (Gebauer et al., 2018; Nichols et al., 2018). Its effect may depend on a person's indeoendent vs. interdependent self-construals (Poulin et al., 2021).

5. Results

To analyze our data, we proceeded in four steps. First, we analyzed performance on each task taken separately. Second, we tested our two general conjectures at the aggregate level. In each case, our analysis is based on both non-parametric statistics (each participant being one observation) and difference-in-differences analysis. Third, we conducted a regression analysis to estimate the average treatment effects on the treated, aiming at measuring the causal effect of the training program. Finally, we examined the longer-term effects of the training program.

5.1. *Differences-in-Differences between the training and the control groups by task*

Table 2 reports statistics for each of the eight incentivized tasks. Columns (1) and (2) report the mean scores and standard deviations in parentheses in the Training group and the Control group prior to the intervention.

We report tests of the randomization in the left panel. Since there is no conjecture concerning pre-intervention characteristics, we report two-tailed tests in column (3). Reassuringly, the allocation of employees to the treatments seem to be balanced since there is no difference significant at the 5 % level (one is significant at the 10 % level).

We also report the mean differences in scores before and after the intervention in Columns (4) and (5), with standard deviations in parentheses. Column (6) reports *p*-values from DiD tests (directional one-tailed Mann-Whitney rank-sum tests since we do predict a greater increase in scores in the Training group than in the Control group). It also indicates whether the DiD comparison goes in the expected direction (+) or not (-). When analyzing each incentivized task separately, we found no significant differences between the Training and the Control groups prior to the (non-) intervention (except for overconfidence). The analysis in Table 2 shows significant improvement for the Training group beyond the change in the Control group only in the Race-to-60 ($p = 0.041$). While they did not differ initially, the mean performance in this task increased by 16.14 % in the Training group and only by 2.19 % in the Control group. As can be seen, seven of the eight difference-in-differences show improvement.[35]

Table 3 reports a similar analysis for the five non-incentivized measures. In the left panel, we see that no pre-intervention score differs significantly between the Training and the Control groups (one is significantly different at the 10 % level). In the right panel, the DiD tests in column (6) show that the Training group improved more than the Control group in every measure except patience. The findings on stress, cognitive flexibility, and mindfulness are robust to multiple outcome comparisons ($p = 0.001$ or less), using a bootstrapping approach allowing *p*-values to be correlated to control for family-wise error rate (List et al., 2019), although the resistance to change result is not robust to this. Note that all five DiD measures move in the predicted direction; a one-tailed binomial test gives $p = 0.031$ (one in 32). This includes patience: participants from both groups became more impatient over time, but this is slightly less the case, on average, for those from the Training group.

This analysis supports our conjectures that training decreased perceived stress (C2a) and resistance to change (C2b), while increasing cognitive flexibility (C2c) and mindfulness (C2d). It rejects our conjecture that training would increase patience (C2e). One interpretation is that such a training may change earlier-versus-later preferences by drawing attention to the relative price of an earlier reward (which should increase patience) but on the other hand, it trains individuals to focus their mind on the present experience (which may make the earlier reward more attractive). The two effects, if they exist, may cancel out each other. This is consistent with Alem et al. (2021), who found little or no effect of meditation training on students' patience.

5.2. *Aggregate differences-in-differences between the training and the control groups*

We now turn to an aggregate analysis. To test our conjectures with a common metric, we built composite measures of performance for the incentivized measures and for the non-incentivized measures, combining re-scaled individual scores in the various tasks and questionnaires. Building a composite index allows us to represent most variation in our measures; it is particularly relevant if some of the components are correlated. Since some researchers may quite reasonably question the construction of such indices, it is nevertheless a very useful way of summarizing information. We note that people in financial markets keenly follow the various stock indices, and that price and quality indices are in wide usage in goods markets.

We used two different procedures for constructing indices to have at least some degree of generality. Simply summing raw scores in the various tasks to compute a single measure would not be reliable because the maximum responses in some tasks would drive the outcome. Therefore, to give equal weighting to each component, we computed an index based on standardized *z*-scores. We standardized each score such that each variable has a mean of 0 and a standard deviation of 1: we subtracted the mean score from each individual score and divided this difference by the standard deviation. Then, we summed the standardized scores obtained in the various tasks to compute this index. As a complement, we built a second composite index based on the classification of the score in each considered task on a 1–4 scale. We started by dividing the performance data in stage 1 or 2 (depending on the task) into quartiles.[36] We sometimes had to adjust the definition of the four categories according to the dispersion of scores in the task. We detail and justify the cut-off points used for each component in Appendix E. After re-scaling the data, we summed the obtained scores in the various tasks to obtain the index.

[35] Or a smaller decrease in performance, such as for the Raven matrices. Indeed, we (unintentionally) selected harder matrices in the second test.

[36] This results in unequal score intervals for a given task. Alternatively, we could have divided scores in equal intervals, but this was not always meaningful because the various components had different standard deviations. Thus, some categories for some tasks would have been almost empty.

Table 2
Comparisons of Scores in Incentivized Tasks.

Tasks	Pre-intervention scores			Difference in Post – Pre-intervention scores (DiD)		
	T (1)	C (2)	p-value (3)	T (4)	C (5)	p-value (6)
CRT	1.51	1.46	0.750	0.39	0.23	0.163
(0–3)	(1.15)	(1.02)		(1.19)	(0.97)	(+)
N	*72*	*70*		*72*	*70*	
Hanoi Tower	1.26	1.03	0.284	0.17	0.42	0.917
(0–2.9)	(1.26)	(1.14)		(1.31)	(1.18)	(-)
N	*72*	*70*		*72*	*70*	
Race-to-60	2.85	2.74	0.793	0.46	0.06	0.041**
(0–6)	(1.47)	(1.47)		(1.76)	(1.45)	(+)
N	*72*	*70*		*72*	*70*	
Raven Matrices	3.82	3.73	0.809	−0.30	−0.49	0.355
(0–6)	(1.54)	(1.67)		(1.44)	(1.60)	(+)
N	*72*	*70*		*72*	*70*	
Overconfidence	1.21	0.94	0.086*	−0.43	−0.25	0.115
(0–6)	(1.06)	(1.07)		(1.45)	(1.41)	(+)
N	*72*	*67*		*72*	*67*	
Find-10	3.82	3.79	0.810	0.35	0.27	0.497
(0–6)	(1.76)	(1.53)		(1.70)	(1.62)	(+)
N	*72*	*70*		*72*	*70*	
Spot-the-Differences	7.27	6.91	0.256	0.91	0.90	0.448
(0–10)	(1.68)	(1.90)		(1.88)	(2.36)	(+)
N	*70*	*70*		*70*	*70*	
Theory-of-Mind	27.04	26.84	0.632	0.03	−0.43	0.344
(0–36)	(2.93)	(2.60)		(3.22)	(2.97)	(+)
N	*68*	*62*		*68*	*62*	

Notes: T for Training group and C for Control group. Below the name of each test are the minimum and maximum possible scores in the tasks. Columns (1) and (2) report mean scores prior the intervention and standard deviations in parentheses. Column (3) indicates the *p*-values from two-tailed Mann-Whitney rank-sum tests comparing the two groups. Columns (4) and (5) report the mean differences in scores 'after' compared to 'before' the intervention. Column (6) reports *p*-values from DiD tests (one-tailed Mann-Whitney rank-sum tests since our conjecture was that the Training group's performance increases more than did that of the Control group). The signs in parentheses indicate whether the DiD is going in the expected direction (+) or not (-). Overconfidence is the average difference between the expected score in the Raven matrices and the actual performance. For the tests administered in stages 1 and 3, the number of observations is 72 in the Training group and 70 in the Control group. In some cases, observations are missing, due to technical problems during the experiment.
** $p < 0.05$.
* $p < 0.10$.

We computed these two indices separately for the incentivized tasks (index-I) and for the non-incentivized tasks (index-NI). The indices for the incentivized tasks include the re-scaled scores in the CRT, the Hanoi tower, Raven matrices, overconfidence on score in the Raven matrices, the Find-10 task, the Spot-the-differences task, the Race-to-60 game, and the Theory-of-Mind test. The indices for the non-incentivized measures include the re-scaled scores for perceived stress, cognitive flexibility, mindfulness, resistance to change, patience from the staircase method. We assigned a reverse-re-scaled score to the actual overconfidence scores, stress, and resistance to change, so that a higher re-scaled value could be assigned to lower actual scores in these variables; for computing the z-scores, we simply reverse coded these variables. This allowed us to sum the re-scaled scores with those of the other variables to compute the indices. We only compared the observations of participants whose scores were measured for the included tasks in the two stages.[37]

Appendix Table C4 displays the correlation coefficients between the scores in the various tasks. Scores in the CRT and Raven matrices correlate with most other measures, while Theory-of-Mind and patience correlate with none. There is a much higher correlation for the scores in the various non-incentivized measures.

Table 4 summarizes the *p*-values from Mann-Whitney rank-sum tests comparing the indices (I and NI, based on *z*-scores or on categories) between the Training and the Control groups (1) before the intervention (stages 1 or 2), and (2) after the intervention (stages 3 or 4). However, the post-intervention comparisons between the Training and the Control groups are naïve since they do not control for possible differences in unobservable characteristics across groups. So, column (3) gives *p*-values from Mann-Whitney rank-sum tests comparing the difference in the composite indices across the two stages of observation in the Training versus the Control groups. This difference-in-difference (DiD) approach controls for heterogeneity in unobservable characteristics and sheds light on the evolution of performance over time in the Training group compared to the Control group.

[37] There are two exceptions. The index for the incentivized measures mainly includes tasks performed in stages 1 and 3, and one task performed in stages 2 and 4 (the test of Theory-of-Mind). To keep the highest number of observations when comparing the evolution of this index over time, we also included the participants who did not attend stage 4 in the computation of the index; for these subjects we computed their pre- and post-intervention indices without including their score in the test of the Theory-of-Mind. We proceeded similarly for the computation of the non-incentivized indices: we included the participants whose patience score is stage 1 or 3 was missing and we computed their pre- and post-intervention indices without including their score in patience.

Table 3
Comparisons of Scores in Non-incentivized Tasks.

Questionnaires	Pre-intervention scores			Difference in Post – Pre-intervention scores (DiD)		
	T (1)	C (2)	p-value (3)	T (4)	C (5)	p-value (6)
Stress	18.19	16.95	0.315	−3.51	−0.16	<0.001***
(0–40)	(6.64)	(6.16)		(6.38)	(4.90)	(+)
N	*68*	*62*		*68*	*62*	
Cognitive flexibility	49.56	49.03	0.649	3.10	−0.06	0.001***
(12–72)	(6.96)	(6.74)		(6.07)	(4.35)	(+)
N	*68*	*62*		*68*	*62*	
Mindfulness	106.44	110.32	0.303	14.19	−1.03	<0.001***
(39–195)	(20.57)	(20.52)		(15.93)	(10.29)	(+)
N	*68*	*62*		*68*	*62*	
Resistance to Change	15.56	16.82	0.073*	−0.47	0.19	0.044**
(6–36)	(4.15)	(4.14)		(2.99)	(2.72)	(+)
N	*68*	*62*		*68*	*62*	
Patience	24.14	24.41	0.969	−1.21	−1.35	0.223
(1–32)	(9.15)	(8.79)		(9.48)	(8.23)	(+)
N	*71*	*66*		*71*	*66*	

Notes: T for Training group and C for Control group. Below the name of each questionnaire are the minimum and maximum possible scores. Columns (1) and (2) report mean scores in each questionnaire prior the intervention and standard deviations in parentheses. Column (3) indicates the *p*-values from two-tailed Mann-Whitney rank-sum tests comparing the two groups. Columns (4) and (5) report the mean differences in scores 'after' compared to 'before' the intervention. Column (6) reports *p*-values from DiD tests (one-tailed Mann-Whitney rank-sum tests since our conjecture was that the Training group's performance increases more than did that of the Control group). The signs in parentheses indicate whether the DiD is going in the expected direction (+) or not (-). The number of observations in the patience measure is lower than the number of participants present in stages 1 and 3 because we exclude individuals who did not answer in one step of the questionnaire, making the computation of the patience index impossible.

*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.10$.

Table 4
Index Comparisons.

	Pre-intervention Training vs. Control (1)	Post-intervention Training vs. Control (2)	DiD Training vs. Control (3)
Index built on 1–4 scale categories			
– Index-I	0.350	0.012**	0.049**
– Index-NI	0.940	0.001***	< 0.001***
Index built on z-scores			
– Index-I	0.380	0.013**	0.079*
– Index-NI	0.929	0.002***	< 0.001***

Notes: The Table reports *p*-values from Mann-Whitney rank-sum tests (two-tailed for column (1) and one-tailed for columns (2) and (3), given our directional hypotheses.). The first two rows use the index built on the 1–4 scale categories, and the last two rows use the index built on z-scores. Index-I include the eight incentivized measures (CRT, Hanoi tower, Raven matrices, overconfidence on score in the Raven matrices, Find-10, Spot-the-differences, Race-to-60, and the Theory-of-Mind test). Index-NI include the five non-incentivized measures (perceived stress, cognitive flexibility, mindfulness, resistance to change, patience). For the incentivized measures, $N = 70$ in the Training group and $N = 67$ in the Control group. For the non-incentivized measures, $N = 68$ in the Training group and $N = 62$ in the Control group. DiD is for Difference-in-Differences.

*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.10$.

Siegel (1956), perhaps the most venerable and respected work on statistics in the social sciences, argues strongly for the use of one-tailed tests when one has an *ex-ante* hypothesis. We follow this policy in our analysis. We report one-tailed tests in columns (2) and (3) since we predicted that the training program would improve performance on the tasks.

We can see that the two indices give generally consistent results. When we compare the two groups, we find no significant differences between the two groups prior to the intervention, whereas the indices significantly differ across groups after the intervention, regardless of whether we consider the incentivized or the non-incentivized measures. Using DiD, our directional one-tailed tests give *p*-values of 0.079 for the index built on *z*-scores and 0.049 for the index built on categories, showing a marginal effect of training at the aggregate level on performance in the incentivized tasks. Regarding non-incentivized measures, the DiD tests reveal that both indices increased significantly more in the Training group than in the Control group, indicating a positive effect of training on performance in these measures.

Overall, this analysis at the aggregate level strongly supports our second conjecture (on the non-incentivized measures) and offers modest support for the first conjecture (on the incentivized measures). Note that the differences found in column (2) are all highly significant, so that a more naïve analysis would draw stronger conclusions. Nevertheless, we acknowledge limitations in this exercise. A composite index cannot fully substitute for an analysis of the difference-in-difference in each task taken in isolation, as reported in the next sub-section.

5.3. Estimation of the causal effect of training in the training group

To identify the causal effect of training on performance, accounting for the non-random assignment of individuals to the treatment and individual traits, we also estimated linear models with robust standard errors, formalized as follows:

$$Y_{it} = \alpha + \gamma Treated_i + \delta Stage + \beta(Treated_i * Stage) + \theta x_j + \psi z_i + \varepsilon_{it} \quad (1)$$

In each regression, the outcome variable, Y_{it}, is the score of the individual in either the pre-intervention stage (1 or 2) or the post-intervention stage (3 or 4). Since we have two observations for each person, we clustered the standard errors at the individual level. The independent variables include a dummy variable for being in the Training group *("Treated$_i$")*, a dummy variable for being in the post-intervention stage *("Stage",* capturing the effect of time on score in both groups), and an interaction term for belonging to the Training group and being in the post-intervention stage *("Treated$_i$*Stage")*. The β parameter captures the Average Treatment Effect on the Treated (ATET), that is, the causal effect of the training program on the trainees. We include two company dummies, x_j, and an individual dummy for a regular practice of meditation, z_i, with θ and ψ the respective associated parameters.[38] ε_{it} represents a random error term.

Table 5 summarizes the estimated regression coefficients of the causal effect of the treatment on the Training group for the composite indices, based on the 4-scale categories or on z-scores, and for each incentivized task (model (1)) and non-incentivized task (model (2)). We report one-tailed tests to reflect our conjectures for a significant improvement of performance in the Training group compared to the Control group. The details of each regression are reported in Appendix G. To correct for attrition, Table C5 in Appendix C reports a similar analysis, assigning the average performances in Stage 3 of the individuals who stayed to the missing participants in Stage 3 who hold similar characteristics in terms of education.[39]

Table 5 shows that training had a causal positive effect on the trainees' performance at the aggregate level for both types of tasks (this is also seen in Table C5, except for the index based on z-scores in model (1)).[40] For example, using the scores based on categories, attending the training program increased the performance index intervention by 1.14 (1.94) points for the incentivized (non-incentivized) tasks. For the index based on z-scores, we found an effect of 0.87 points (only marginally significant) and 1.81 points, respectively. The task-by-task analysis shows no significant causal effect of the training in model (1), except for the Race-to-60 task ($p = 0.070$) in which training marginally increased performance by 0.40 points over the baseline.

Model (2) in Table 5 reveals a significant causal effect of the training on the trained participants' scores over time in four of five non-incentivized measures. It decreased the reported stress levels by 3.35 points and resistance to change by 0.66 points (the latter is only marginally significant), and it increased the cognitive flexibility score by 3.17 points and the mindfulness score by 15.22 points.[41] In contrast, we found no causal effect of training on patience, although the sign of the ATET is in the expected direction (patience tended to decrease if no treated individuals had been trained but increased due to training). Table C5 replicates these findings, although the marginal improvement in resistance-to-change is no longer significant when we correct for attrition.

In an exploratory analysis, we estimated the same models as those in Table 5, augmented with gender covariates (see Appendix Table C6). Gender was included both as a dummy variable and in interaction with each of the other independent variables (Treated, Stage and Treated*Stage). In model (1), the common causal effect of the treatment on the index based on categories became significant at the 1 % level and the effect on the index based on z-scores at the 5 % level; the causal effect of the treatment on the two indices that is specific to the males who received the treatment was negative and significant at the 5 % level. This suggests that overall,

[38] The dummy variable for a regular practice of meditation is coded 1 for all individuals who reported to practice meditation every day or every week, and 0 otherwise. We also include fixed effects for fims. Note that clustering the standard errors at the company level would not provide reliable estimates because having too few clusters (it would be three in our study) would bias the estimates.

[39] This criterion was selected because it is significantly different between those who quit before a given stage in the Training group and in the Control group (see Table C3). Thus, those who quit from a given group and were (were not) graduates received the mean performance of those from the very same group who were (were not) graduates.

[40] We explored alternative ways to build the aggregate indices for the incentivized measures. In particular, we created two indices based respectively on the aggregated measures of cognitive sophistication (CRT, Hanoi tower, Race to 60 and Raven matrices) and on the measures of attention (Find-10 and Spot-the-differences tasks). We proceeded similarly for the indices based on z-scores. The same DiD regressions as with the full indices were run with the two new indices. We also used alternatively a Principal Component Analysis to construct the indices and conducted the same DiD analysis. However, most of the lower levels of aggregation produced non-significant coefficients (available upon request). This indicates that we need a sufficiently high number of tests whose results point in the same direction to identify a positive effect of the intervention at the aggregate level.

[41] The mindfulness score aggregates five factors. It could be argued that the last three factors (act with awareness, non-judgment, and non-reactivity to inner experience) are those that are the most related to the effect of the intervention, whereas the first two factors (observe, describe) are more related to the program itself. We replicated the same analysis for each of the five factors taken separately. The causal effect of training is significant at the 1% level on each of the five factors. In all the other regressions reported in the paper, using the aggregate measures or each factor separately delivers qualitatively the same results with the same level of significance. Details are available from the authors upon request.

Table 5
Average Treatment Effects on the Treated.

Dependent Variable: Index or Score	*Incentivized Tasks (1)*	*Non-Incentivized Tasks (2)*	
Analysis at the aggregate level			
Index based on categories	1.138**	Index based on categories	1.945***
	(0.612)		(0.442)
N	*274*	*N*	*260*
Index based on z-scores	0.866*	Index based on z-scores	1.812***
	(0.604)		(0.371)
N	*274*	*N*	*260*
Analysis at the task level			
CRT	0.160	Stress	−3.353***
	(0.183)		(1.001)
N	*284*	*N*	*260*
Hanoi Tower	−0.253	Cognitive Flexibility	3.167***
	(0.210)		(0.928)
N	*284*	*N*	*260*
Race-to-60	0.401*	Mindfulness	15.223***
	(0.272)		(2.351)
N	*284*	*N*	*260*
Raven Matrices	0.180	Resistance to Change	−0.664*
	(0.258)		(0.505)
N	*284*	*N*	*260*
Overconfidence	−0.177	Patience	0.137
	(0.244)		(1.525)
N	*278*	*N*	*274*
Find-10	0.076		–
	(0.281)		
N	*284*		
Spot-the-Differences	0.014		–
	(0.363)		
N	*280*		
Theory-of-Mind	0.465		–
	(0.548)		
N	*260*		

Notes: The table summarizes the estimates of the Average Treatment Effects on the Treated -ATET- (the β parameter in Eq. (1)) from separate linear regressions with robust standard errors clustered at the individual level. The dependent variable is the index or the score (pre-intervention score in stage 1 or 2, or post-intervention score in stage 3 or 4, depending on the task). Company dummies and a dummy for a regular (weekly or daily) practice of meditation are included. Model (1) is for the incentivized tasks and model (2) for the non-incentivized tasks. N is the number of observations.

*** $p < 0.01$.
** $p < 0.05$.
* $p < 0.10$ in one-tailed tests.

the females' performance in the incentivized tasks benefited more from the treatment than did males' performance. The common ATET on the Race-to-60 score became significant at the 5 % level, with no significant gender effect. No gender effect on the ATET was detected in model (2) for the non-incentivized measures: both genders benefited from the training to the same extent.

This analysis confirms the results presented in the previous subsections. Summarizing:

Result 1: There is a causal positive impact of the mindfulness training program on the treated individuals' aggregate indices based on incentivized measures. However, at the individual task level, there is only significant evidence of a causal impact in the Race-to-60 game (a); no significant impact was found (b) on the other cognitive-sophistication tasks (CRT test, Hanoi Tower, Raven matrices), (c) the attention tasks (Find-10 and Spot-the-Difference), (d) over-confidence, and (e) Theory-of-Mind.

Result 2: There is a causal positive impact of the mindfulness training program on the treated individuals' aggregate indices based on non-incentivized measures related to mental health. A robust causal impact is identified at the individual task level for (a) perceived stress, (b) cognitive flexibility, (c) mindfulness, and some evidence is found of a causal effect on (d) resistance to change. No such effect is found for individuals' patience.

5.4. Estimation of the causal longer-term effect of training

Finally, we explored the causal effect of training on the treated three months after the end of the intervention. In stage 5, participants again played the Race-to-60 game and the Spot-the-differences task, and they answered the psychological questionnaires. Table 6 reports the ATET from linear regressions with robust standard errors clustered at the individual level. The dependent variable is the score of the individual in either the pre-intervention stage (stage 1 or 2, according to the task) or three months after the end of the intervention (stage 5).

Table 6
Average Treatment Effects on the Treated in the Longer Term.

Dependent Variable: Index or Score	
Analysis at the aggregate level	
Index based on category	2.609***
	(0.577)
Index based on z-scores	2.566***
	(0.497)
Analysis at the task level	
Race-to-60	0.503*
	(0.390)
Spot-the-Differences	0.827*
	(0.619)
Stress	−4.828***
	(1.244)
Cognitive Flexibility	4.671***
	(0.894)
Mindfulness	16.755***
	(3.168)
Resistance to Change	−1.039*
	(0.702)
Number of observations	*188*
Number of clusters	*94*

Notes: The table summarizes the estimates of the Average Treatment Effects on the Treated (the β parameter in Eq. (1)) from separate linear regressions with robust standard errors clustered at the individual level. The dependent variable is the index or the score (the pre-intervention score in stage 1 or 2, depending on the task, or the score three months after the end of the intervention, in stage 5). The composite indices are computed based on the four non-incentivized tasks performed in stage 5. Company dummies and a dummy for a regular (weekly or daily) practice of meditation are included. The regressions exclude participants from the Control group who received training just after stage 4. This exclusion leads us to recalculate the z-scores on the restricted sample. ** $p < 0.05$;.

*** $p < 0.01$.

* $p < 0.10$ for the justified one-tailed tests.

The analysis is conducted at the aggregate level for the four non-incentivized measures collected in stage 5, using our two composite indices, and at the task level for the six tasks performed in this stage. 30 participants from the Control group were excluded from the analysis, having received the same training just after stage 4 (as explained earlier). Appendix Table C7 reports complementary specifications. There model (1) excludes the participants of the Training group from the same company; model (2), instead, includes all participants. As in Table 5, these regressions include company dummies and a dummy for a regular practice of meditation.

Tables 6 and C7 show that, regardless of the exclusions considered, we can still identify a significant and beneficial effect of the training at the aggregate level for the non-incentivized measures, and on perceived stress, cognitive flexibility, and mindfulness, each taken in isolation, three months after the end of the intervention. The magnitude of the effects is naturally smaller in model (2) in Table C7. Excluding the participants from the Control group who received training also reveals a significant long-term effect of the training on the treated participants' scores in the Race-to-60 game, the Spot-the-Differences task, and reduced resistance to change (see Table 6), but this is only marginally significant (and no longer significant in Table C7).

This analysis leads to our last result:

Result 3: The causal positive impact of the training program on the treated individuals' aggregate indices and scores in the non-incentivized measures related to mental health persists three months after the end of the intervention. The longer-term causal impact of training on scores in incentivized tasks is conditionally observed, depending on the specification.

6. Discussion and conclusion

Mindfulness has been considered to involve being more focused on one's immediate experience through improved attention, awareness, and control. Practicing mindfulness may improve the well-being of individuals and may also have corresponding benefits for the performance of firms and for public health. We studied the impact of mindfulness and positive-psychology training at three large companies with both non-incentivized questions and incentivized tasks with control groups. There have been previous studies on the effects of mindfulness practice, but these have typically been conducted without incentives, and they used different tasks than most of ours. Conducting this study in-house at three firms also contributed to a much lower attrition rate. We do find strong evidence for the effect on mental health, with more nuanced support for the effect of mindfulness training on cognitive performance.

The questionnaire results are easy to interpret, with highly-significant results for the aggregated non-incentivized tasks ($p < 0.001$). Regarding cognitive performance, although our study offers one of the toughest possible tests, we found indications that this program led to some beneficial effects, with nearly every measure moving in the expected direction for those who took the training relative to the control group. The average treatment effect on the treated for the aggregated incentivized tasks in two tests is at least weakly significant ($p = 0.032$ and $p = 0.076$), while it is highly significant for the measures more directly related to mental health.

Why would the results differ depending on whether incentives were used for the tasks? As mentioned earlier, there are several reasons why one might observe stronger effects for the non-incentivized than for the incentivized measures (despite the fact that most of the verbal measures were developed in English and translated into French, which could have increased the error variance). One is that by introducing monetary incentives in some tasks, the level of attention of all the participants (in both the control and treatment groups) increased, thereby perhaps clouding differences between the two groups in terms of attention capability. Another possible reason is a declarative bias – individuals who spent a lot of energy to attend the training program (increasing time pressure on their regular workload) and to exercise between training sessions may persuade themselves that it was a worthy investment, perhaps exaggerating the conscious (or unconscious) benefits when this is costless. Such a bias might also result from a willingness to please the company that paid for this training.

A third possibility, related to the content of two types of tasks, is that it is more difficult to improve cognitive abilities (mainly captured by the incentivized tasks) than psychological states and well-being (mainly captured by the non-incentivized measures). Finally, training may have had positive impact on performance in all tasks, but this was counteracted by a negative effect on effort in the incentivized tasks. Studies in psychology have found that mindfulness can increase subjective well-being by reducing the financial desires discrepancies (Brown et al., 2009). Furthermore, studies using functional magnetic resonance imaging have found that mindfulness meditators are less susceptible to monetary rewards and exhibit reduced neural activation in reward-related areas of the brain during incentive anticipation compared to non-meditators (Kirk et al., 2015).[42]

We believe that previous results on incentivized performance are consistent with our data. Alem et al. (2021) feature one task, choosing a low- or high-calorie item for consumption. They state: "We find indicative evidence that … eating habits may have improved." Shreekumar and Vautrey (2021) find a 1.9 % increase in productivity in a proofreading task by users of a meditation app in the U.S., based on a large scale RCT study online. They found no effect on sustained app usage once the incentives expired. We also saw little deterioration in our measures three months after the end of our intervention. This is an important issue that requires more study, applying to all interventions involving incentives. Finally, the field experiment conducted by Cassar et al. (2022) in an educational setting also reveals positive effects of mindfulness meditation on students' mental health but complex effects on academic performance. While a negative effect was identified in the short run through, notably, increased sleeping time, a positive effect on performance was observed six months after the end of the training program, driven by those who continued practicing meditation.

In any case, taken together, the data from our eight incentivized tasks show improvements in performance in seven of the eight tasks ($p = 0.031$ on a one-tailed binomial test). While it is true that only one comparison for individual tasks shows a significant difference, we nevertheless feel that this is a meaningful pattern overall. Our view is that mindfulness has the potential to improve performance, but that it may be necessary to have longer or more intensive trainings to routinely show this improvement and create a habit.

We view our results on incentivized cognitive performance as exploratory, in that either a larger sample or a more intensive training might well have provided more conclusive results. Indeed, there may be some heterogeneity in resistance and benefit to such training, and in the natural mindfulness of individuals (*e.g.*, Brown et al., 2003). Moreover, by focusing on individual tasks our design excluded potential effects of the intervention on behavior in team settings. For example, the training could help improve communication skills and trust among employees. We also acknowledge some limitations of our study. In particular, the intervention that was introduced was not a standardized MBSR program, as it incorporated components of both mindfulness and positive psychology. This limits the comparability of our results with those of studies using the standard MBSR program. Furthermore, manipulations of the training program would be needed to disentangle the effects of the mindfulness and positive-psychology elements of the intervention.

Nevertheless, we believe that the methodological improvements introduced in our study enhance the credibility of our results. It is particularly crucial to minimize the selection bias that has typically been exacerbated by high levels of attrition. The low attrition rate in our study reflects our conducting the research in-house at firms; this is something to be considered in subsequent studies. Our results offer hope for future work to refine and extend the research on the effects of mindfulness training on health and cognitive abilities. It would be particularly interesting to explore whether this type of intervention at work might reduce the risk of burn-out and improve teams' wellbeing and efficiency. We hope that others join us in this interesting and important quest.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: One of the co-authors is a member of the French Institute of Positive Leadership that designed the training program. It has been decided from the very beginning of the study that this co-author would not have access to the dataset during the study, would not

[42] From that perspective, the level of incentives to perform the cognitive tasks was perhaps too low in our experiment. Indeed, reducing stress may help people perform better in the incentivized tasks when stakes are high, whereas it may disengage people if stakes are low.

contribute to the data analysis, and would not contribute to the writing of the paper. The funding of the experiment to pay the participants was provided exclusively by the University of Lyon, independently from the French Institute of Positive Leadership and from the companies in which the training was implemented. There has been no (and there will be no) financial flows between the researchers and with this institute or the companies.

Data availability

Data will be made available on request.

Acknowledgements

We are particularly grateful to Michele Belot, Alessandra Cassar, Lea Cassar, Uri Gneezy, Catherine Moste (Sanofi), Valérie Leyldé and Elisabeth Hild (BioMérieux), and Nicolas Rolland (ENGIE). We thank Quentin Thévenet for programming the experiment, and Julien Benistant, Ariane Bontems, and Ann Thampi for research assistance. This research has benefited from the financial support of the LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon, within the program Investissements Avenir (ANR-11-IDEX-007) operated by the French National Research Agency (ANR).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jebo.2023.10.027.

References

Alem, Y., Behrendt, H., Belot, M., Bíró, A., 2021. Mind training, stress and behaviour—arandomised experiment. PLoS One 16 (11), e0258172.
Allen, T.D., Eby, L.T., Conley, K.M., Williamson, R.L., Mancini, V.S., Mitchell, M.E., 2015. What do we really know about the effects of mindfulness-based training in the workplace? Ind. Organ. Psychol. 8, 652–661.
Ash, E., Sgroi, D., Tuckwell, A., Zhuo, S., 2021. Mindfulness reduces information avoidance. *Center for Law & Economics Working Paper Series*, 2021-13.
American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders, 5th ed., 2013. https://doi.org/10.1176/appi.books.9780890425596
Avoyan, A., Schotter, A., 2020. Attention in games: an experimental study. Eur. Econ. Rev. 124 (C).
Baer, R.A., Smith, G.T., Hopkins, J., Krietemeyer, J., Toney, L., 2006. Using self-report assessment methods to explore facets of mindfulness. Assessment 13 (1), 27–45.
Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., Plumb, I., 2001. The "Reading the Mind in the Eyes" test revised version: a study with normal adults, and adults with Asperger syndrome or high−functioning autism. J. Child Psychol. Psychiatry 42, 241–251.
Bosch-Rosa, C., Meissner, T., Bosch-Domènech, A., 2018. Cognitive bubbles. Exp. Econ. 21, 132–153.
Bradley, M.M., Lang, P.J., 1994. Measuring emotion: the self-assessment manikin and the semantic differential. J. Behav. Ther. Exp. Psychiatry 25 (1), 49–59.
Brefczynski-Lewis, J., Lutz, A., Schaefer, H., Levinson, D., Davidson, R., 2007. Neural correlates of attentional expertise in long-term meditation practitioners. Proc. Natl. Acad. Sci. U.S.A 104, 11483–11488.
Brocas, I., Carrillo, J.D., Wang, S.W., Camerer, C.F., 2014. Imperfect choice or imperfect attention? Understanding strategic thinking in private information games. Rev. Econ. Stud. 81 (3), 944–970.
Brown, K.W., Kasser, T., Ryan, R.M., Linley, P.A., Orzech, K., 2009. When what one has is enough: mindfulness, financial desire discrepancy, and subjective well-being. J. Res. Pers. 43 (5), 727–736.
Brown, K.W., Ryan, R.M., 2003. The benefits of being present: mindfulness and its role in psychological well-being. J. Pers. Soc. Psychol. 84 (4), 822–848.
Brunner, B., Igic, I., Keller, A.C., Wieser, S., 2019. Who gains the most from improving working conditions? Health-related absenteeism and presenteeism due to stress at work. Eur. J. Health Econ. 20, 1165–1180.
Camerer, C., Mobbs, D., 2017. Differences in behavior and brain activity during hypothetical and real choices. Trends Cogn. Sci. (Regul. Ed.) 21, 46–56.
Camerer, C.F., Hogarth, R.M., 1999. The effects of financial incentives in experiments: a review and capital-labor-production framework. J. Risk Uncertain. 19, 7–42.
Camerer, C.F., Johnson, E.J., 2004. Thinking about attention in games: backward and forward induction. In: Brocas, I., Carrillo, J.D. (Eds.), *The Psychology of Economic Decisions* (vol. 2.): Reasons and Choices. Oxford University Press, Oxford, pp. 111–129.
Carpenter, P., Just, M., Shell, P., 1990. What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices Test. Psychol. Rev. 97 (3), 404–431.
Cassar, L., Fischer, M., Valero, V., 2022. Keep calm and carry on: the short- vs. long-run effects of mindfulness meditation on (academic) performance. CEPR Press Discussion Paper No. 17675. https://cepr.org/publications/dp17675.
Chabris, C.F., Laibson, D., Morris, C.L., et al., 2009. Individual laboratory-measured discount rates predict field behavior. J. Risk Uncertain. 37 (2–3), 237–269.
Charness, G., Gneezy, U., 2009. Incentives to exercise. Econometrica 77, 909–931.
Charness, G., Jahnke, R., 2012. A group-based wellness intervention in the laboratory. B. E. J. Econom. Anal. Policy 12 (1), 7.
Chin, B., Slutsky, J., Raye, J., Creswell, J.D., 2019. Mindfulness training reduces stress at work: a randomized controlled trial. Mindfulness 10, 627–638.
Coffman, L., Niederle, M., 2015. Pre-analysis plans have limited upside especially where replications are feasible. J. Econ. Perspect. 29 (3), 81–98.
Cohen, S., Kamarck, T., Mermelstein, R., 1983. A global measure of perceived stress. J. Health Soc. Behav. 24 (4), 385–396.
Donaldson, S.I., Lee, J.Y., Donaldson, S.I., 2019. Evaluating positive psychology interventions at work: a systematic review and meta-analysis. Int. J. Appl. Posit. Psychol. 4, 113–134.
Dufwenberg, M., Sundaram, R., Butler, D., 2010. Epiphany in the game of 21. J. Econ. Behav. Organ. 75, 132–143.
Eby, L.T., Allen, T.D., Conley, K.M., Williamson, R.L., Henderson, T.G., Mancini, V.S., 2019. Mindfulness-based training interventions for employees: a qualitative review of the literature. Human Resour. Manag. Rev. 29 (2), 156–178.
European Foundation for the Improvement of Living and Working Conditions (2010). Eurofound: work-related stress. https://www.eurofound.europa.eu/publications/report/2010/work-related-stress.
Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., Sunde, U., 2018. Global evidence on economic preferences. Q. J. Econ. 133 (4), 1645–1692.
Falkinger, J., 2008. Limited attention as a scarce resource in information-rich economies. Econ. J. 118 (532), 1596–1620.
Frederick, S., 2005. Cognitive reflection and decision making. J. Econ. Perspect. 19 (4), 25–42.

Gebauer, J.E., Nehrlich, A.D., Stahlberg, D., et al., 2018. Mind-body practices and the self: yoga and meditation do not quiet the ego but instead boost self-enhancement. Psychol. Sci. 29 (8), 1299–1308.
Giné, X., Karlan, D., Zinman, J., 2010. Put your money where your butt is: a commitment savings account for smoking cessation. Am. Econ. J.: Appl. Econ. 2 (4), 213–235.
Giraud, L., Frimousse, S., Le Bihan, Y., 2022a. Mindful and Positive leadership Interventions: the impacts on positive experiences, traits, and behaviors. Revue de Gestion des Ressources Humaines (2), 3–23.
Glomb, T.M., Duffy, M.K., Bono, J.E., Yang, T, 2011. Mindfulness at work. In: Martoccio, J., Liao, H., Joshi, A. (Eds.), Research in Personnel and Human Resource Management. Emerald Group Publishing, Bingley, pp. 115–157.
Gneezy, U., Rustichini, A., Vostroknutov, A., 2010. Experience and insight in the Race game. J. Econ. Behav. Organ. 75, 144–155.
Goldberg, S.B., Tucker, R.P., Greene, P.A., Davidson, R.J., Wampold, B.E., Kearney, D.J., Simpson, T.L., 2018. Mindfulness-based interventions for psychiatric disorders: a systematic review and meta-analysis. Clin. Psychol. Rev. 59, 52–60.
Good, D.J., Lyddy, C.J., Glomb, T.M., Bono, J.E., Brown, K.W., Duffy, M.K., et al., 2016. Contemplating mindfulness at work: an integrative review. J. Manag. 42, 114–142.
Gotink, R.A., Chu, P., Busschbach, J.J., Benson, H., Fricchione, G.L., Hunink, M.G., 2015. Standardised mindfulness-based interventions in healthcare: an overview of systematic reviews and meta-analyses of RCTs. PLoS One 10 (4).
Goyal, M., Singh, S., Sibinga, E.M.S., et al., 2014. Meditation programs for psychological stress and well-being: a systematic review and meta-analysis. JAMA Intern. Med. 174 (3), 357–368.
Grupe, D.W., McGehee, C., Smith, C., Francis, A.D., Mumford, J.A., Davidson, R.J., 2019. Mindfulness training reduces PTSD symptoms and improves stress-related health outcomes in police officers. J. Police Crim. Psychol.
Hafenbrack, A.C., Kinias, Z., Barsade, S.G., 2014. Debiasing the mind through meditation: mindfulness and the sunk-cost bias. Psychol. Sci. 25 (2), 369–376.
Hafenbrack, A.C., LaPalme, M.L., Solal, I., 2022. Mindfulness meditation reduces guilt and prosocial reparation. J. Pers. Soc. Psychol. 123 (1), 28–54.
Hafenbrack, A.C., Vohs, K.D., 2018. Mindfulness meditation impair task motivation but not performance. Organ. Behav. Hum. Decis. Process. 147, 1–15.
Heeren, A., Douilliez, C., Peschard, V., Debrauwere, L., Philippot, P., 2011. Cross-cultural consistency of the Five Facets Mindfulness Questionnaire: adaptation and validation in a French sample. Eur. Rev. Appl. Psychol. 61, 147–151.
Hülsheger, U.R., Alberts, H., Feinholdt, A., Lang, J.W.B., 2013. Benefits of mindfulness at work: the role of mindfulness in emotion regulation, emotional exhaustion, and job satisfaction. J. Appl. Psychol. 98, 310–325.
Indhumathi, G., Thirumakkal, G., 2015. A study on role of occupational stress on employees' productivity. Int. J. Manag. 6 (1), 560–572.
Initiative Mindfulness France, 2020. *La pleine conscience (Mindfulness) dans la société française : réalités et perspectives*, 81p.
Ivtzan, I., Young, T., Martman, J., Jeffrey, A., Lomas, T., Hart, R., Eiroa-Orosa, F.J., 2016. Integrating mindfulness into positive psychology: a randomised controlled trial of an online positive mindfulness program. Mindfulness 7 (6), 1396–1407.
Kabat-Zinn, J., 1990. Full Catastrophe living: Using the Wisdom of Your Body and Mind to Face stress, Pain and Illness. Delacorte Press, New York.
Kersemaekers, W., Rupprecht, S., Wittmann, M., Tamdjidi, C., Falke, P., Donders, R., Speckens, A., Kohls, N., 2018. A workplace mindfulness intervention may be associated with improved psychological well-being and productivity. A preliminary field study in a company setting. Front. Psychol. 9, 195.
Khoury, B., Sharma, M., Rush, S.E., Fournier, C., 2015. Mindfulness-based stress reduction for healthy individuals: a meta-analysis. J. Psychosom. Res. 78 (6), 519–528.
Kirk, U., Brown, W., Downar, J., 2015. Adaptive neural reward processing during anticipation and receipt of monetary rewards in mindfulness meditators. Scan 10, 752–759.
Kirk, U., Downar, J., Montague, P.R., 2011. Interoception drives increased rational decision-making in meditators playing the ultimatum game. Front. Neurosci. 5, 1–11.
Lakey, C.E., Campbell, W.K., Brown, K.W., Goodie, A.S., 2007. Dispositional mindfulness as a predictor of the severity of gambling outcomes. Personal. Individual Differences 43, 1698–1710.
List, J.A., Shaikh, A.M., Xy, Y., 2019. Multiple hypothesis testing in experimental economics. Exp. Econ. 22, 773–793.
Lomas, T., Medina, J.C., Ivtzan, I., Rupprecht, S., Hart, R., Eiroa-Orosa, F.J., 2017. The impact of mindfulness on well-being and performance in the workplace: an inclusive systematic review of the empirical literature. Eur. J. Work Organ. Psychol. 26, 492–513.
Lutz, A., Slagter, H.A., Rawlings, N.B., Francis, A.D., Greischar, L.L., Davidson, R.J., 2009. Mental training enhances attentional stability: neural and behavioral evidence. J. Neurosci. 29, 13418–13427.
MacLean, K.A., Ferrer, E., Aichele, S.R., Bridwell, D.A., Zanesco, A.P., Jacobs, T.L., et al., 2010. Intensive meditation training improves perceptual discrimination and sustained attention. Psychol. Sci. 21, 829–839.
Martin, M.M., Rubin, R.B., 1995. A new measure of cognitive flexibility. Psychol. Rep. 76 (2), 623–626.
Mazar, N., Amir, O., Ariely, D., 2008. The dishonesty of honest people: a theory of self-concept maintenance. J. Market. Res. 45 (6), 633–644.
Meier, M., Unternaehrer, E., Schorpp, S.M., Wenzel, M., Benz, A., Bentele, U.U., Dimitroff, S.J., Denk, B., Prüssner, J.C., 2020. The opposite of stress: the relationship between vagal tone, creativity, and divergent thinking. Exp. Psychol. 67 (2), 150–159.
Meyers, M.C., van Woerkom, M., Bakker, A.B., 2013. The added value of the positive: a literature review of positive psychology interventions in organizations. Eur. J. Work Organ. Psychol. 22 (5), 618–632.
Montani, F., Vandenberghe, C., Khedhaouria, A., Courcy, F., 2020. Examining the inverted U-shaped relationship between workload and innovative work behavior: the role of work engagement and mindfulness. Hum. Epidemiol. Anim. Lab. Correl. Chem. Carcinog. 73 (1), 59–93.
Moore, A., Malinowski, P., 2009. Meditation, mindfulness and cognitive flexibility. Conscious. Cogn. 18 (1), 176–186.
Mrazek, M.D., Franklin, M.S., Tarchin Phillips, D., et al., 2013. Mindfulness training improves working memory capacity and GRE performance while reducing mind wandering. Psychol. Sci. 24 (5), 776–781.
National Institute on Aging (2023). "Cognitive health and older adults." https://www.nia.nih.gov/health/cognitive-health-and-older-adults.
Nichols, S., Strohminger, N., Rai, A., Garfield, J., 2018. Death and the self. Cogn. Sci. 42, 314–332.
Ohrnberger, J., Fichera, E., Sutton, M., 2017. The relationship between physical and mental health: a mediation analysis. Soc. Sci. Med. 195, 42–49.
Oreg, S., 2003. Resistance to change: developing an individual differences measure. J. Appl. Psychol. 88 (4), 680–693.
Poulin, M.J., Ministero, L.M., Gabriel, S., Morrison, C.D., Naidu, E., 2021. Minding your own business? Mindfulness decreases prosocial behavior for people with independent self-construals. Psychol. Sci. 32 (11), 1699–1708.
Prevost, M., Carrier, M.E., Chowne, G., Zelkowitz, P., Joseph, L., Gold, I., 2014. The Reading the Mind in the Eyes test: validation of a French version and exploration of cultural variations in a multi-ethnic city. Cogn. Neuropsychiatry 19, 189–204.
Reb, J., Narayanan, J., Chaturvedi, S., 2014. Leading mindfully: two studies on the influence of supervisor trait mindfulness on employee well-being and performance. Mindfulness 5, 36–45.
Roeser, R.W., Schonert-Reichl, K.A., Jha, A., Cullen, M., Wallace, L., Wilensky, R., Oberle, E., Thomson, K., Taylor, C., Harrison, J., 2013. Mindfulness training and reductions in teacher stress and burnout: results from two randomized, waitlist-control field trials. J. Educ. Psychol. 105, 787–804.
Rubinstein, A., 2016. A typology of players: between instinctive and contemplative. Q. J. Econ. 131 (2), 859–890.
Sänger, J., Bechtold, L., Shoofs, D., Blaszkewicz, M., Wascher, E., 2014. The influence of acute stress on attention mechanisms and its electrophysiological correlates. Front. Behav. Neurosci. 8, 353.
Seligman, M.E.P., Csikszentmihalyi, M., 2000. Positive psychology: an introduction. Am. Psychol. 55, 5–14.
Shapiro, S.L., Astin, J.A., Bishop, S.R., Cordova, M., 2005. Mindfulness-based stress reduction for health care professionals: results from a randomized trial. Int. J. Stress Manag. 12 (2), 64–176.
Shonin, E., Van Gordon, W., Dunn, T.J., Singh, N.N., Griffiths, M.D., 2014. Meditation awareness training (MAT) for work-related wellbeing and job performance: a randomised controlled trial. Int. J. Ment. Health Addict. 12, 806–823.

Shreekumar, A., Vautrey, P.L., 2021. Managing emotions: the effects of online mindfulness meditation on mental health and economic behavior. MITEconomics Job Market Paper economics.mit.edu/files/22355.
Siegel, S., 1956. Nonparametric Statistics for The Behavioral Sciences. McGraw Hill, New York.
Simon, H.A., 1957. Models of Man. Wiley, New York.
Smith, V.L., 1976. Experimental economics: induced value theory. Am. Econ. Rev. 66, 274–279.
Tang, Y.Y., Ma, Y., Wang, J., et al., 2007. Short-term meditation training improves attention and self-regulation. Proc. Nat. Acad. Sci. U.S.A. 104 (43), 17152–17156.
Vonderlin, R., Biermann, M., Bohus, M., et al., 2020. Mindfulness-based programs in the workplace: a meta-analysis of randomized controlled trials. Mind. https://doi.org/10.1007/s12671-020-01328-3.
Woerner, A., Romagnoli, G., Probst, B.M., Bartmann, N., Cloughesy, J.N., Lindemans, J.W., 2021. Should Individuals Choose Their Own Incentives? Evidence from a Mindfulness Meditation Intervention. Working Paper. University of Munich.
Zhang, D., Lee, E.K.P., Mak, E.C.W., Ho, C.Y., Wong, S.Y.S., 2021a. Mindfulness-based interventions: an overall review. Br. Med. Bull. https://doi.org/10.1093/bmb/ldab005.
Zhang, Y., Chen, Z.B., Lu, S., Ni, S., 2021b. Are mindful people more risk-averse? Effects of trait and state mindfulness on risk preference in decision-making. Int. J. Psychol. 56 (3), 407–414.