Journal of Economic Behavior and Organization 188 (2021) 322–336

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

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

A local community course that raises wellbeing and pro-sociality: Evidence from a randomised controlled trial

Christian Krekel[a,d,*], Jan-Emmanuel De Neve[b,d], Daisy Fancourt[c], Richard Layard[d]

[a] London School of Economics, United Kingdom
[b] University of Oxford, United Kingdom
[c] University College London, United Kingdom
[d] Centre for Economic Performance (CEP)

ARTICLE INFO

Article history:
Received 26 August 2020
Revised 20 April 2021
Accepted 17 May 2021

ABSTRACT

Despite a wealth of research on its correlates, relatively little is known about how to effectively raise wellbeing in local communities by means of intervention. Can we teach people to live happier lives, cost-effectively and at scale? We conducted a randomised controlled trial of a scalable social-psychological intervention rooted in self-determination theory and aimed at raising the wellbeing and pro-sociality of the general adult population. The manualised course ("Exploring What Matters") is run by non-expert volunteers (laypeople) in their local communities and to date has been conducted in more than 26 countries around the world. We found that it has strong, positive causal effects on participants' subjective wellbeing and pro-sociality (compassion and social trust) while lowering measures of mental ill health. The impacts of the course are sustained for at least two months post-treatment. We compare treatment to other wellbeing interventions and discuss limitations and implications for intervention design, as well as implications for the use of wellbeing as an outcome for public policy more generally.

1. Introduction

For decades, enormous academic effort has been put into exploring the causes and consequences of wellbeing (Diener et al., 1999; Layard et al., 2014). Health (especially mental health), being partnered, and social relationships account for more than three quarters of the explained variance in adult people's life satisfaction (Clark et al., 2018). At the same time, there is growing evidence showing that wellbeing is a significant predictor of important life and economic outcomes, including health and longevity (Danner et al., 2001; Steptoe and Wardle, 2011; Graham and Pinto, 2019), productivity and income (De Neve and Oswald, 2012; Oswald et al., 2015; Bellet et al., 2020), voting (Liberini et al., 2017), and even compliance with lockdown measures during Covid-19 (Krekel et al., 2020).

Yet, we know little about how to effectively improve the wellbeing of the general adult population. Can we teach people to live happier lives? Can we do this by means of intervention, cost-effectively and at scale? Are impacts sustained over

* Corresponding author.
E-mail addresses: c.krekel@lse.ac.uk (C. Krekel), jan-emmanuel.deneve@sbs.ox.ac.uk (J.-E. De Neve), d.fancourt@ucl.ac.uk (D. Fancourt), r.layard@lse.ac.uk (R. Layard).

https://doi.org/10.1016/j.jebo.2021.05.021

time? Answering these questions has profound implications: if wellbeing is not fixed and can be taught, it can be used as a meaningful indicator to measure societal progress, and help direct public policy attention towards areas that are malleable and where there is room for improvement.

The answers to these questions, however, are not *ex-ante* clear. A prominent view argues that there exists a set point of wellbeing around which individuals fluctuate (Brickman and Campbell, 1971). According to this view, individuals largely adapt to various changes in life circumstances, driven by withdrawal of attention to these changes, so that their wellbeing remains largely unchanged over time (Frederick and Loewenstein, 1999; Kahneman, 2000). Hedonic adaptation has been used to explain phenomena such as why life satisfaction has been stagnant in many developed countries over the past decades while economic living standards have increased substantially (Easterlin, 1974; Easterlin et al. 2010). There is now an established body of evidence on hedonic adaptation to various positive or negative changes in life circumstances, including changes in marital status (Lucas, 2005; Lucas and Clark, 2006; Oswald and Gardner, 2006; Stutzer and Frey, 2006), disability (Menzel et al., 2002; Oswald and Powdthavee, 2008), or income (Di Tella et al., 2010; Kuhn et al., 2011). According to this view then, wellbeing is less malleable and significant increases in average population wellbeing may be limited in societies with already high economic living standards.

Another point of view, in line with expectancy-value theory in psychology (Battle, 1965), suggests that familiarising people with evidence on what could make them happier may lead to an update in their beliefs, which, in turn, may lead to a change in their behaviour. Expectancies refer to the subjective probabilities of becoming happier which are attached to certain behaviours, whereas values refer to the magnitudes of happiness changes resulting from these behaviours. To the extent that this change in their behaviour may improve people's wellbeing and thereby reinforce their beliefs, people may uphold that behaviour, leading to permanent (as opposed to temporary) wellbeing change. This mechanism may be especially effective when it comes to behaviours in life domains which are important for wellbeing and, at the same time, are less prone to hedonic adaptation, such as time spent on social relationships (Powdthavee, 2008), experiences (Carter and Gilovich, 2010), or pro-social action (Dunn et al., 2008; Aknin et al., 2013; Drouvelis and Grosskopf, 2016). According to this interventionist view, wellbeing is malleable and significant increases in average population wellbeing may be possible, even in economically affluent societies.

Interventions that aim to improve wellbeing directly have typically been narrow in focus, looking at specific, often clinical target groups or at-risk populations (as opposed to healthy adults in the general population), often including people suffering from depression and anxiety (see Taylor et al. (2017), for example) or bodily pain (see Hausman et al. (2014), for example).[1] A notable exception is Heintzelman et al., 2020: the authors evaluated the impact of ENHANCE, a 12-week wellbeing course targeted at the general adult population in their local communities which has been trialled in hybrid (i.e. ten sessions online and two sessions offline) and face-to-face (i.e. twelve sessions offline) delivery. When delivered face-to-face, it is led by graduate-level trained clinicians. Similar to the intervention presented in this paper, it focuses primarily on positive habits, skills, and attitudes. During the course, a new skill is introduced every week, participants practice that skill, and then write about their experiences. The authors found that it had strong, positive causal effects on participants' wellbeing up to six months after the main intervention has ended and up to three months after an extended following-up period.

We studied the impact of a similar course – "Exploring What Matters" – which is a local community intervention aimed at raising the wellbeing and pro-sociality of the general adult population. Besides contents, it differs from existing interventions in at least two critical aspects: first, the course is manualised and led by non-expert volunteers (laypeople) rather than trained clinicians, making it highly cost-effective. Second, due to its cost-effectiveness, it is highly scalable and can be delivered face-to-face in the local communities of course leaders and participants. Cost-effectiveness and scalability have important implications for the feasibility of social prescribing in health economics, i.e. the referral by GPs to non-medical community interventions to address the wider determinants of health and to help patients improve health-related behaviours (see NHS Long Term Plan (2019), for example). As of August 2020, 431 courses have been completed, with a total of 5621 participants, yielding an average course size of 15 (13 course participants plus two volunteers leading the course). Most courses have been conducted in the UK (343), with a further 88 courses run in 25 countries. "Exploring What Matters" is run by Action for Happiness, a registered charity in England, which was launched in 2011. Its patron is the Dalai Lama, who helped to launch the course in London in 2015.

Using a randomised controlled trial, we studied the impacts of six of these courses which took place in London between August 2016 and December 2017: two during autumn 2016, two during spring 2017, and two during autumn 2017. In what follows, we first describe the intervention, derive hypotheses on wellbeing change, and illustrate the study design, before turning to our findings on self-reported outcomes and biomarkers. We then present the results of a replication exercise using before-after data from the universe of courses conducted to date. Finally, we calculate the cost-effectiveness of the course in raising wellbeing, compare it to other interventions in the literature, and discuss shortcomings, implications, and avenues for future research in the field, as well as implications for the use of wellbeing as an outcome for public policy more generally.

[1] See Sin and Lyubomirsky (2009) and Bolier et al. (2013) for meta-analyses.

2. The intervention

The "Exploring What Matters" course brings together participants in face-to-face groups to discuss what matters for a happy, meaningful, and virtuous life. Participants span a wide range of ages and socio-economic backgrounds but can be broadly classified, as per their self-reports, into two categories: people who are unhappy and looking for ways to improve their lives; and people who are interested in wellbeing more generally and want to learn more, or want to share these ideas with others.

The intervention is manualised: each course is led by two volunteers as facilitators on an unpaid basis.[2] Recruitment of course leaders follows a documented, standardised process: each candidate completes a *Leader Registration* process sharing their motivation and experience and is given instructions on what is required. Once potential course leaders have a co-leader, venue, and dates in mind, they complete a *Course Application* process. Action for Happiness reviews this application and, if certain criteria are met, arranges a call to discuss next steps.[3] Once a course is approved, course leaders receive guidance and structured resources to facilitate course delivery. The Supplementary Materials II include a link to the complete documentation of the recruitment process of course leaders.

Participants sign up online, and when doing so, are asked to make a donation; donations aim to cover the implementation costs of the course (implementation costs are about £90 ($113) per participant, including variable costs for course materials as well as allocated fixed costs).[4] Donations are voluntary and participants can take part without donating. The function of donations is to make the course scalable and accessible to people regardless of their financial situation. Besides that, they aim at raising course attendance, by exploiting the notion of sunk costs.[5] The course consists of eight consecutive weekly sessions lasting between two and 2.5h each. Each of these sessions builds on a thematic question, for example, what matters in life, how to find meaning at work, or how to build happier communities. Each of these questions is discussed against the background of scientific evidence on subjective wellbeing, mental health, and pro-sociality as well as motivation and group learning.

Courses are advertised both online and offline in local communities, and potential participants must register online. Online advertising is done via emails to people who have previously registered with Action for Happiness and live nearby and to new people via targeted local Facebook advertising. Offline advertising is done via local course leaders using word-of-mouth and, to a lesser extent, local promotion (for example, through notice boards or local press).

2.1. Hypotheses

Course design and delivery are rooted in psychological self-determination theory (Deci and Ryan, 1985), which states that autonomy, relatedness, and competence are fundamental human needs that enable people to achieve wellbeing. The course aims at building *(i)* autonomy by enabling participants to discover for themselves what matters for their lives, using a weekly mindfulness exercise, gratitude exercise, and personal reflection, supported by a "Did You Know?" section that introduces scientific evidence on that week's theme; *(ii)* relatedness by facilitating interpersonal connections and social trust, within the gathering of people in their local communities; and *(iii)* competence by enabling participants to experience for themselves how behavioural changes to daily routines can make differences to their and other people's wellbeing, using goal-setting and social commitment tools to help translate motivation into action. The Supplementary Materials II include links to the complete course materials of both course participants and course leaders.

There is an established evidence base linking psychological self-determination theory to wellbeing (Ryan and Deci, 2000), across life domains and different cultural contexts (Milyavskaya and Koester, 2011; Church et al., 2013), including its constituent elements (Brown et al., 2003; Chirkov et al., 2003; Guardia et al., 2000). Likewise, there is evidence from systematic reviews and meta-analyses linking certain elements of the course curriculum, in particular mindfulness, meditation, and related self-regulation strategies, to positive outcomes in non-clinical populations, including wellbeing, depression, and anxiety, with medium to strong effect sizes (see Sedlmeier et al. (2012), Gu et al. (2015), or Querstret et al. (2020), for example).

We therefore hypothesise that, first, the course has positive impacts on wellbeing. Second, we hypothesise that – to the extent that it fosters interpersonal connections between strangers and encourages pro-social action-taking – the course has positive impacts on pro-social attitudes. Third, we hypothesise that – to the extent that it changes beliefs about behaviours in life domains that are important for wellbeing and that are less prone to hedonic adaptation – the course may have sustained impacts.

---

[2] Although the intervention is manualised, some degree of adaptability is possible. For example, course leaders may choose the most appropriate venue or allow for more group discussion time. However, they are encouraged to stick closely to the course guide.

[3] Course leaders have a similar demographic profile as course participants, with a slightly higher average age. 58% are female. 58% are between 31 and 50 years old, 25% between 18 and 30, and 17% between 51 and 70. They tend to have higher than average levels of life satisfaction and social trust (both about 7.9 on zero-to-ten scales).

[4] Converted using an exchange rate of 1:1.25 as of July 16, 2020.

[5] Unfortunately, we did not have data on the donation amount per participant, and hence could not study heterogeneity of course outcomes depending on donations.

3. Methods

We conducted a randomised controlled trial which focused on six courses that took place in London between August 2016 and December 2017, including a total of 146 participants. These were informed about the study, both during online registration and on site, and written consent was taken.[6] Following power calculations based on the historical number of course participants (about 13 per course), this sample size was determined before data collection and analysis.[7]

Course participants were self-selected. To study the extent to which they differed from the general adult population, we compared our estimation sample, pre-treatment, with a sample from the nationally representative UK Household Longitudinal Survey ("Understanding Society"), restricted to London and to the same age span as our participants. We found that there were little, quantitatively relevant differences in the age distribution between course participants and the general population. Participants were, however, significantly more likely to be female in our sample (83% vs. 45%). Moreover, they were significantly less likely to be married (20% vs. 53%) and more likely to be in a domestic partnership (25% vs. less than one percent). This difference, however, is likely to be an artefact arising from survey design: Understanding Society does not ask about a "domestic" (as our survey did) but about a "civil" partnership. When it comes to income, we found again little, quantitatively relevant differences, except for the highest income category: our sample included significantly less individuals earning £75,000 ($94,000) or more and was somewhat more skewed towards lower incomes. Finally, participants reported, on average, a lower level of life satisfaction (by about 47% of a standard deviation), pre-treatment, than the general population.[8]

3.1. Randomised controlled trial

To account for self-selection of participants into the course, we employed a waitlist randomisation protocol: after registering for the course online, participants (who reported that they were able to attend the course on either one of two sets of pre-specified upcoming dates, two months apart) were randomly allocated to one of the two sets, unaware of how these related to treatment and control group. Participants in the earlier set of dates were in the treatment group, those in the later set in the waitlisted control group. They were then invited to arrive on the same date to have their first data collected. The event started with a brief introductory session which explained to participants that they were required to fill in surveys and provide saliva samples. This was when participants read the project information sheet and signed written consent forms. After written consent had been obtained, the data were collected. After data collection had finished, the brief introductory session was over and participants in the treatment group started their course immediately. Participants in the control group would start their course eight weeks later, after the treatment group would have finished, and left the premises. Treatment and control group were kept separate: neither group knew anything about the other, and the two groups did not meet on that day.

Note that the choice of the appropriate control group is not trivial: as there exists no natural, credible counterfactual that could lend itself as a business-as-usual scenario in our intervention context, choosing a waitlisted control group comprised of those who initially selected into the intervention seems most appropriate for adhering as closely as possible to evidence-based practice. Note that our control group does not include a placebo: arguably, a placebo could help to better isolate and identify the active ingredients of the intervention. At the same time, however, it raises the question of what precisely the (neutral) placebo can be, whether one control group with one placebo is actually enough, and whether or not elements like socialising are active parts of the intervention package and should thus be accounted for as such. We will return to the issue of choosing the appropriate control group in more detail later on in the discussion section.

3.2. Data collection

Data were collected at three points in time: at $t = 0$, right before the course started; at $t = 1$, right before it ended, which was eight weeks after $t = 0$; and at $t = 2$, eight weeks after $t = 1$. At each point in time, data were collected at the same hour of day (circa 6pm in the evening). Fig. 1 illustrates our randomised controlled trial and data collection process.

Our estimation sample (exploiting data points at $t = 0$ and $t = 1$) consisted of 146 respondents (279 observations), of which 73 were in the treatment (136 observations) and 73 (143 observations) were in the control group. As can be seen, in our estimation sample, we have an attrition rate of about 5%.[9] We will test the sensitivity of our results regarding attrition later in our robustness section. To look at treatment effect persistence, we exploited data points at $t = 2$ in an extended sample. As all respondents had been treated at $t = 2$, results are exploratory.

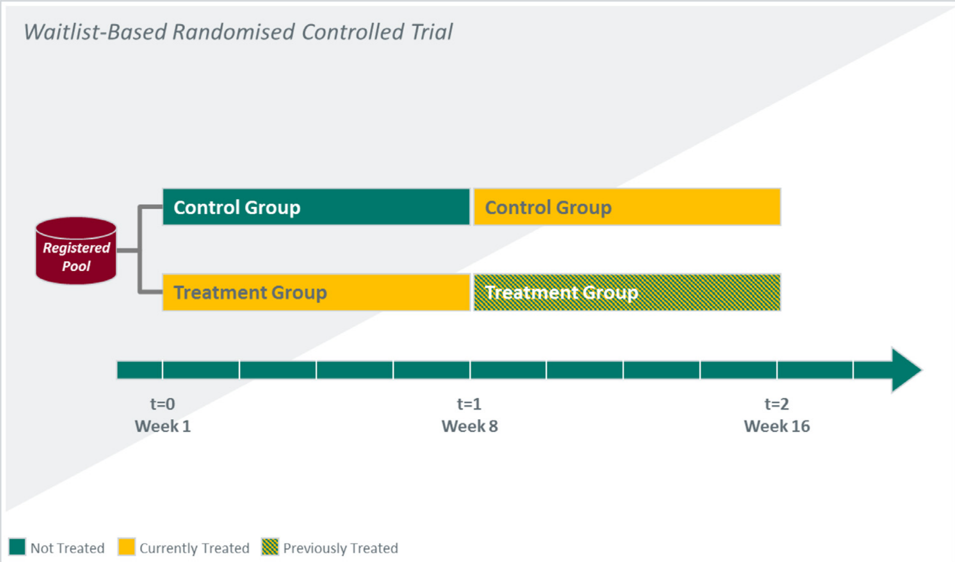
Importantly, data at $t = 0$ and $t = 1$ were collected right before the start of the first and the last session, respectively, at the back of the meeting room. Collecting data before the start of the respective session reduced measurement error which may have resulted from participants' euphoria of having started or finished the course being mixed up with actual outcomes. Note that, during data collection at $t = 0$ and $t = 1$, the atmosphere was deliberately kept neutral, and participants were

[6] This study passed the Internal Review Board of the Research Ethics Division at the London School of Economics (Reference: 00507).

[7] A power of 0.8, alpha of 0.05 two-tailed, and an assumed effect size of 0.5 yielded at least N = 128 individuals, with 64 per experimental group.

[8] See Supplementary Materials Table 1a for this analysis.

[9] That is, $(279 / (146 \times 2) - 1) \times (-1) = \sim 0.05$.



Sources: Own illustration.

Fig. 1. Randomised controlled trial and data collection. Source: own illustration.

asked to complete surveys and give biomarker samples *before* they had a chance to meet other participants in the main room. To be consistent, the same protocol regarding neutrality of atmosphere that applied to data collection at $t = 0$ and $t = 1$ also applied to data collection at $t = 2$. Attending data collection at $t = 2$ had been communicated as mandatory beforehand. To avoid creating emotional arousal about attending this additional session, participants did not know what content and format it involved. Finally, neither course participants nor volunteers leading the course knew whether they were in the treatment or control group during data collection at $t = 0$. Participants' group allocation was announced only *after* data collection at $t = 0$ had finished.

3.3. Outcomes

We collected data on two categories of outcomes: self-reported outcomes came from survey data, which included items on subjective wellbeing, mental health, and pro-sociality. Biomarkers were collected through saliva samples, which included cortisol – a steroid hormone responsive to stress – and a range of cytokines-immune proteins involved in inflammatory response. Activation of the inflammatory response system has been shown to be bidirectionally associated with mental ill health and depressive symptoms (Dowlati et al., 2010; Miller and Raison, 2016). The Supplementary Materials III contain the project information sheet, written consent form, and the survey instruments used in the study, including surveys at $t = 0$, $t = 1$, and $t = 2$.

Items on subjective wellbeing covered evaluative (life satisfaction), experiential (happiness and anxiousness), and eudemonic (worthwhileness) dimensions. They were measured on eleven-point single-item Likert scales whereby zero denoted the lowest possible level and ten the highest. Items on mental health covered frequently used screening measures to detect depression (the three-point nine-item Patient-Health Questionnaire, PHQ-9) and anxiety (the three-point seven-item Generalised-Anxiety-Disorder Questionnaire, GAD-7). PHQ-9 scores from zero to four imply minimal, from five to nine mild, from ten to fourteen medium, and from fifteen to 27 strong depression symptomatology. GAD-7 scores have a similar interpretation but are cut off at 21. Respondents in our sample could thus be characterised as, on average, mildly depressed ($M = 6.4$, SD $= 4.5$) and anxious ($M = 6.1$, SD $= 4.6$). Distributions were, however, highly skewed: in case of depression, for example, we found that 24 out of 133 respondents for whom we had data at $t = 1$ (about 18%) showed medium or strong depressive symptomatology. When these were omitted, the remaining respondents could be characterised as only minimally depressed ($M = 4.4$, SD $= 2.7$), not much different from PHQ-9 scores typically found at the general adult population level, which range from $M = 3.0$, SD $= 4.3$ for 30 to 39 year-olds to $M = 3.7$, SD $= 5.1$ for 50 to 59 year-olds in the US, for example (Tomitaka et al., 2018). Items on pro-sociality included the Santa Clara Brief Compassion Scale, a composite score running from five to 35 which measures pro-sociality by asking respondents about their readiness to help others-and eleven-point single-item Likert scales on social trust and gratitude. We standardised self-reported outcomes to have mean zero and standard deviation one, using the course-set-specific control group mean and standard deviation.

Biomarkers included, besides cortisol, pro-inflammatory cytokines IL-1β and IL-6, anti-inflammatory cytokine IL-10, interferon IFN-γ, and chemokine IL-8. These markers have been shown to be responsive to both short-term and long-term

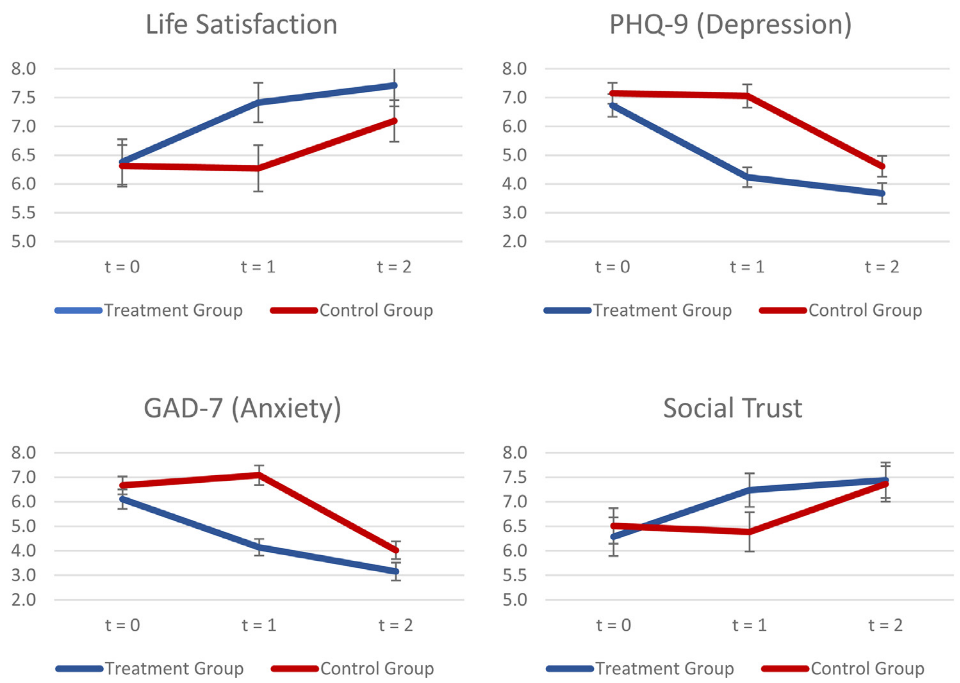


Fig. 2. Average scores of groups at different points in time. *Notes:* A waitlist randomisation design was applied: between $t = 0$ and $t = 1$, the treatment group received treatment; between $t = 1$ and $t = 2$, the control group received treatment. Scores are in natural units. Life satisfaction and social trust were measured on scales from zero to ten, PHQ-9 for depression on a scale from zero to 27, and GAD-7 for anxiety on a scale from zero to 21. $N = 383$ (146 at $t = 0$, 133 at $t = 1$, and 104 at $t = 2$). Confidence intervals are 95%. *Source:* Own data collection, own calculations.

psychosocial interventions (Fancourt et al., 2016). They were collected by means of a saliva sample right after the surveys with self-reported outcomes had been completed. We applied passive drool method of sample collection using low protein-bind collection cryovials. Samples were analysed-three times independently at the Institute for Interdisciplinary Salivary Bioscience Research at the University of California at Irvine-using multiplex immunoassays. Cortisol was measured in µg/dL, cytokines in pg/mL. We took means across the three analyses run for each biomarker, removed outliers, and log-transformed and standardised the data.

3.4. Controls

We collected survey data on socio-demographic characteristics of respondents, including age, gender, marital status, education, employment, income, religion, religious practice, preference for meeting new people and making new friends, health (including pregnancy), and health-related behaviours (including smoking and medication usage), to control for potential differences between treatment and control group over time. All controls were measured pre-treatment. Table 1b in the Supplementary Materials shows variable definitions and descriptive statistics, Table 1c balancing properties between treatment and control group: there was little evidence for significant mean differences in outcomes and controls between groups prior to course start. Similarly, Table 4 in the Supplementary Materials shows that there was little evidence for significant differences for the control group between $t = 0$ and $t = 1$, pointing towards the absence of time trends or waitlist effects. There were no known confounding events during the study period.

3.5. Descriptive evidence

Before turning to our empirical model and results, we first look at descriptive evidence on subjective wellbeing, mental health, and pro-sociality. Fig. 2 plots the raw means of four of our self-reported outcomes-life satisfaction, mental health (PHQ-9 for depression and GAD-7 for anxiety), and social trust-during the observation period.[10]

We make three observations: first, between points $t = 0$ and $t = 1$, the course improved the scores of the treatment group, in line with our first and second hypotheses, whereas those of the control group remained constant. Second, between points $t = 1$ and $t = 2$, the course improved the scores of the control group (which received treated during that period) in a

[10] Figures for other self-reported outcomes are available upon request.

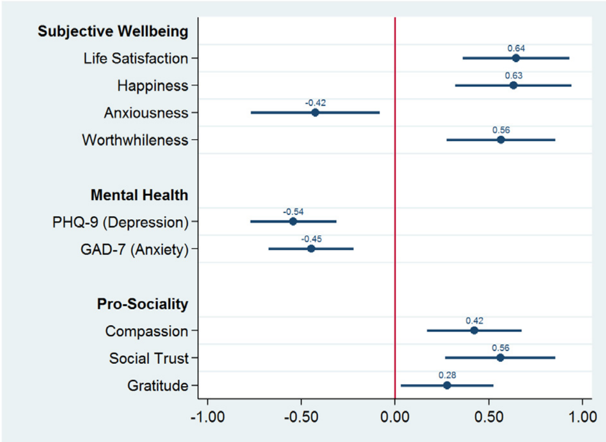


Fig. 3. Impacts on self-reported outcomes: subjective wellbeing, mental health, and pro-sociality. *Notes:* Outcomes have been standardised prior to running regressions (i.e. transformed to z-scores with mean of zero and standard deviation of one, using the control group mean and standard deviation). See Supplementary Materials Table 2a for the corresponding regression table with controls. Robust standard errors were clustered at the participant level. $N = 279$ (146 respondents, of which 73 were in treatment and 73 in control). Confidence bands are 95%. *Source:* Own data collection, own calculations.

similar fashion, whereas those of the treatment group were sustained or even continued to improve, in line with our third hypothesis.

3.6. Empirical model

We now turn to our empirical model. Our baseline model is a difference-in-differences specification that compared the evolution of course outcomes between groups over time:[11]

$$y_{it} = \beta_0 + \beta_1 \text{Treatment}_i * \text{Post}_t + \beta_2 \text{Treatment}_i + \beta_3 \text{Post}_t + \beta_4{}' X_{it} + \mu_s + \varepsilon_{it} \quad \text{with } t = \{0,\ 1\} \tag{1}$$

where y_{it} is the outcome of respondent i at time t; Treatment_i is a dummy equal to one if the respondent belonged to the treatment group, and zero else; Post_t is a dummy equal to one at $t = 1$, and zero else; X_{it} is a vector of controls; and μ_s is a course-set-specific fixed effect. In what follows, we present coefficients obtained from estimating Eq. (1) without controlling for X_{it}, and relegate those obtained from estimating the equation with controls to the Supplementary Materials. If randomisation was successful and treatment was exogenous, controlling for X_{it} should not make any difference, and this is precisely what we will show.

Our model was estimated using OLS, with robust standard errors clustered at the participant level. β_1 is the causal effect (i.e. the average treatment effect on the treated) of course participation. Note that our model could not exploit data points at $t = 2$ because there was no credible control group anymore.

Taken together, we tested fifteen hypotheses in our main analysis (i.e. four outcomes related to subjective wellbeing, two outcomes related to mental health, three outcomes related to pro-sociality, plus six biomarkers). To account for multiple hypotheses testing, we used the stepdown multiple testing procedure suggested by Romano and Wolf (2005a, 2005b), with the four-step algorithm outlined in Romano and Wolf (2016). In essence, the algorithm constructs a null distribution for each of our fifteen hypotheses tests based on a set of null resampling test statistics (in our case, using a bootstrap with 100 repetitions and cluster-robust standard errors at the participant level in both the original regression and during the resampling procedure). We find that our stepdown adjusted P values (corresponding to the significance of a hypothesis test where fifteen tests were implemented) continue to indicate significance at conventional levels for all our coefficient estimates (where our original P values indicated significance).[12]

4. Results

4.1. Impacts on subjective wellbeing, mental health, and pro-sociality

Fig. 3 plots the coefficient estimates of our self-reported outcomes. We again confirmed our first and second hypotheses on positive impacts on wellbeing and pro-social attitudes.

[11] Alternatively, one could regress the post-treatment on the pre-treatment outcome and a treatment dummy (which enforces a balanced panel). Results were qualitatively the same.

[12] See Supplementary Materials Tables 5a and 5b for these results.

In terms of subjective wellbeing, the course increased life satisfaction by about 64% of a standard deviation, happiness by about 63%, and worthwhileness by about 56%. Anxiousness, on the contrary, was decreased by about 42%. Impacts were large: for life satisfaction, for example, the effect size corresponds to an increase of about one point on a zero-to-ten scale. Impacts were significant at the 5% level.

In terms of mental health, the course decreased both PHQ-9 and GAD-7 scores, respectively, by about 54% and 45% of a standard deviation (impacts did not significantly differ from each other). Impacts were again large: participants, prior to taking the course, reported mean PHQ-9 and GAD-7 scores of about 6.7 and 6.1, respectively, which corresponds to a clinical symptomatology of mild depression and anxiety. The course improved scores to, on average, 4.3 points for PHQ-9 and 3.7 for GAD-7, which corresponds to minimal depression and anxiety. Impacts were again significant at the 5% level.

Although strong, impacts on mental health were clearly weaker than those found in trials based on cognitive behavioural therapy (CBT). For example, the Improving Access to Psychological Therapies (IAPT) scheme in the UK has been found to reduce PHQ-9 and GAD-7 scores, on average, by about eight and seven (Clark et al., 2009). The Cognitive Behavioural Therapy as an Adjunct to Pharmacotherapy (CoBalT) trial has been found to reduce PHQ-9 and GAD-7 scores, on average, by about 7.1 and 4.7 (Wiles et al., 2016). However, these trials were targeted specifically at individuals who suffer from depression and anxiety, rather than the general adult population.

In terms of pro-sociality, we found that the course significantly increased both compassion and social trust at the 5% level, respectively, by about 42 and 56% of a standard deviation (about 0.6 and 1.1 points). The impact on gratitude, however, was lower and only marginally significant.

Next, we ran a series of regressions to look into the importance of social context, potential mechanisms behind our average treatment effects, and heterogeneous effects. To do so, we first re-estimated our baseline model with controls (Supplementary Materials Table 2a), and then selectively exploited these controls in these subsequent analyses. Note that including controls has little impact on our identified effects (compare Fig. 3 with Supplementary Materials Table 2a), which suggests that randomisation was successful and treatment was exogenous.

To study the importance of social context, we note that whether or not we control for social context, measured as participants' preference for socialising, has made little difference to our findings.[13] Next, we ran two additional regressions. First, we re-estimated our model without controlling for participants' preference for socialising but controlling for all other covariates: coefficient estimates were slightly attenuated yet continued to be strong, suggesting that socialising may play a role but only partially explains impacts. We then split our sample by the mean pre-treatment value of this variable: again, we did not find that impacts were systematically stronger for respondents who had a higher preference for socialising, pre-treatment, and *vice versa*. Thus, it does not seem that participants who had a higher preference for socialising benefited more from the course than others, or the other way around.[14]

To explore potential mechanisms, we collected data on two categories of additional outcomes: information and behaviour. The former included measures that relate to knowledge of what contributes to one's own and other people's wellbeing. The latter included measures that relate to frequencies of behaviours in various social domains, including the private sphere, close relationships, and other people.[15] Items on information and behaviour also served as manipulation checks, as the course explicitly aims at changing both.

Re-estimating our baseline model with standardised measures of information as outcomes, we indeed found that participants reported to feel more knowledgeable of what contributes to a happy and meaningful life, to know more what matters to them personally, and to feel more able to do things to improve their own, and to a somewhat lesser extent, the wellbeing of other people. When it comes to standardised measures of behaviour, the course increased the frequency in which participants reported to practice mindfulness or meditation, to treat themselves in a kind way, to connect with other people, and to do something kind or helpful for others. Effect sizes ranged between 50% and 80% of a standard deviation-comparable to our main outcomes.[16] Impacts of mindfulness on wellbeing and the importance of social relationships and pro-social behaviour for wellbeing are well-documented in the literature (see Bohlmeijer et al. (2010), Godfrin and van Heeringen (2010), or Gu et al. (2015) for mindfulness; Powthavee (2008) for social relationships; or Borgonovi (2008), Meier and Stutzer (2008), or Dolan et al. (2021) for pro-social behaviour, for example).

To shed light on whether some participants benefited more than others, we conducted a heterogeneity analysis, running separate regressions for participants in different terciles of the respective self-reported outcome distribution, pre-treatment.[17] Fig. 1 in the Supplementary Materials shows our findings: only in case of PHQ-9 scores did differences between terciles turn out to be significant. Impacts on participants in the first tercile of PHQ-9 scores (who were more depressed) were almost seven times larger than for those in the bottom tercile (who were less); the difference was significant at the 5% level. Besides that, we did not find much evidence for heterogeneous effects.

[13] We found similar results regardless of whether a stated-preference (i.e. importance for meeting new people and making new friends) or a revealed-preference item (i.e. frequency of meeting in local clubs) was used to measure the importance of social context to participants.

[14] Results are available upon request.

[15] Data on these additional outcomes had only been collected at a later stage (starting from t = 1).

[16] See Supplementary Materials Tables 3a and 3b for these findings.

[17] The choice of terciles was motivated by sample size.

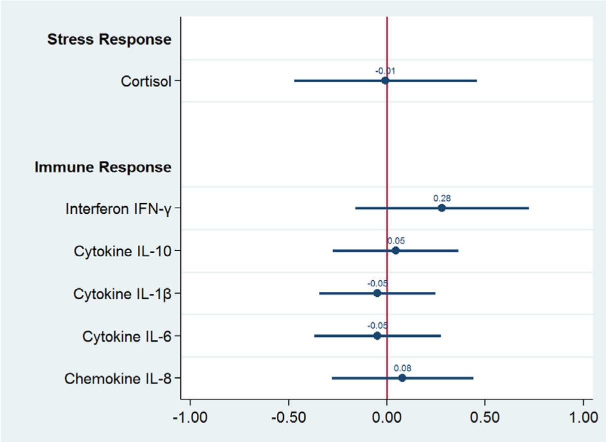


Fig. 4. Impacts on biomarkers: cortisol and cytokines. *Notes:* Outcomes have been standardised prior to running regressions (i.e. transformed to z-scores with mean of zero and standard deviation of one, using the control group mean and standard deviation). See Supplementary Materials Table 2b for the corresponding regression table with controls. Robust standard errors were clustered at the participant level. N between 236 and 275 depending on biomarker due to removal of outliers. Confidence bands are 95%. *Source:* Own data collection, own calculations.

4.2. Impacts on cortisol and cytokines

We next look at biomarkers – cortisol as a stress response hormone and a range of cytokines as immune response proteins associated with mental ill health and depressive symptoms. Fig. 4 shows coefficient estimates.

We did not find that the course had significant impacts on biomarkers at conventional levels. However, we found that cytokines consistently moved into the hypothesised direction: pro-inflammatory cytokines IL-1β and IL-6, which correlate positively with depressive symptoms, decreased, whereas anti-inflammatory cytokine IL-10, interferon IFN-γ, and chemokine IL-8 (which correlate negatively) increased. Compared to self-reported outcomes, biomarkers were noisier and impacts smaller in size.[18]

4.3. Robustness

To the extent that out-of-sample selection was not random and correlated with outcomes (for example, unhappier people may have been more likely to drop out of the study), or differed by group, it would have biased our identified effects. We looked at attrition by regressing the number of periods on each outcome alongside course-set-specific fixed effects, using robust standard errors clustered at the individual level. We found little evidence that outcomes were significant predictors of the number of periods participants remained in the programme, neither on average nor by group.[19] We take this as evidence that out-of-sample selection was rather random. Note that only about 5% of participants dropped out between $t = 0$ and $t = 1$, and a slightly larger proportion (22%) between $t = 1$ and $t = 2$. Finally, compliance was high: on average, participants attended seven out of eight sessions.

5. Replication

Since its launch in 2015, 431 courses have been completed worldwide, totalling 5621 participants. From the beginning, the charity running the courses – Action for Happiness – has been collecting data on course outcomes at the participant level. Participants are sent a link to the survey at $t = 0$ after registering online for the course. Completing the online survey is mandatory for course participation. After the course has finished, they are again sent a link to the survey at $t = 1$, whereby completion is incentivised by a voucher for a free, one-year subscription to a mindfulness app.

In particular, by means of online surveys, data on course participants' life satisfaction, mental wellbeing, compassion, and social trust have been collected. Mental wellbeing is measured using the Short Warwick-Edinburgh Mental Well-being Scale, which asks respondents to report the frequency of several experiences related to their mental wellbeing during the past two weeks. The item is bound between seven and 35, whereby higher scores indicate higher mental wellbeing.

Although a before-after comparison of these measures does not yield causal effects of course participation on course outcomes, we can still use these online surveys, which are high-powered and widely spread across geographical regions and

[18] As with our self-reported outcomes, we ran separate regressions for participants in different terciles of the respective biomarker distribution, pre-treatment. Fig. 2 in the Supplementary Materials plots coefficient estimates: we found again little systematic evidence that the course had significant impacts by tercile at conventional levels.

[19] Results are available upon request.

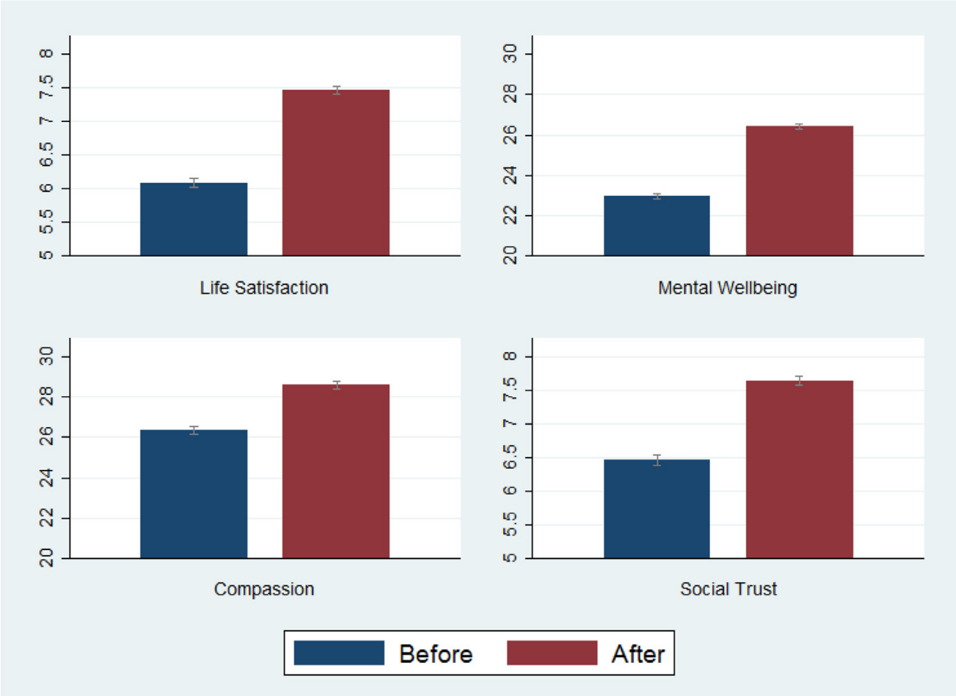


Fig. 5. Impacts on self-reported outcomes in online surveys: life satisfaction, mental wellbeing, compassion, and social trust. *Notes:* data at $t = 0$ and $t = 1$ from online surveys on the universe of courses during the period 2015 to 2019. Scores are in natural units. Life satisfaction and social trust were measured on scales from zero to ten; mental wellbeing by means of the Short Warwick-Edinburgh Mental Well-being Scale, which runs from seven to 35; and compassion by means of the Santa Clara Brief Compassion Scale, which runs from five to 35. Confidence intervals are 95%. *Source:* Own data collection, own calculations.

over time, to check the external validity of our main findings, which were based on six courses in London between 2016 and 2017. Fig. 5 shows the results of this before-after comparison of course outcomes collected via online surveys, restricted to respondents for whom we had both data at $t = 0$ and $t = 1$, amounting to about 5600 individuals (about 2300 observations before and 2300 after) for comparison.

Similar to the findings in our trial, the before-after comparison showed strong, positive associations between course completion and life satisfaction, mental wellbeing, compassion, and social trust.

Associations were, however, larger: for life satisfaction, for example, we found a mean difference of about 1.4 points on a zero-to-ten scale (pre-mean of 6.1, post-mean of 7.5). Larger associations could be driven by three factors: first, our before-after comparison did not account for general trends in wellbeing. Second, larger associations could, in part, be driven by attrition in online surveys: whereas attrition was low in our trial (only about 5% of participants dropped out between $t = 0$ and $t = 1$), attrition in online surveys was much higher, at about 36%. Finally, larger impacts could be explained by the timing of surveys at $t = 1$: the link to the survey is sent out shortly after the course has finished, whereas in our trial data at $t = 1$ had been collected *before* the last session started. It is therefore possible that participants' euphoria of having finished the course was mixed up with actual course outcomes in online surveys.

6. Discussion

Using a randomised controlled trial, we found that the "Exploring What Matters" course had strong, positive causal effects on participants' self-reported subjective wellbeing and mental health. It also induced a shift in participants' attitudes towards more pro-sociality. These impacts seemed to be sustained at $t = 2$ two months post-treatment. An analysis of the mechanisms of wellbeing change suggested that effects on participants may have come about through changes in knowledge of wellbeing and behaviour in areas that have been shown to be important for wellbeing and in which there is little hedonic adaptation, including mindfulness, social relationships, and pro-social behaviour. Biomarkers collected through saliva samples, including cortisol and a range of cytokines involved in inflammatory response, moved consistently into the hypothesised direction yet failed to reach statistical significance at conventional levels.

One explanation for why we did not find significant effects on biomarkers may be power issues combined with relatively noisy measures. Another, related explanation may be the composition of our sample: high levels of pro-inflammatory cytokines have been found for major depression. Respondents in our sample, however, reported only mild depressive symptomatology on average, pre-treatment. In fact, we found that only eight out of 133 respondents for whom we had data at

$t = 1$ (about 6%) reported strong symptomatology, as indicated by PHQ-9 scores of fifteen or higher. Moreover, even amongst these, only about a third showed associated elevated inflammation (Wium-Andersen and Nielsen, 2013). For cortisol, individual differences and timing of measurement matter: it has been found to be a rather short-term measure for stress (Miller et al., 2007). Another, complementary explanation is that the course improves participants' positivity towards life more generally, which is initially captured by self-reported outcomes and may manifest itself in impacts on biomarkers only in the long-run. Indeed, there is some indication in the literature that tangible health outcomes of wellbeing interventions are attainable only in the longer term, especially if participants are motivated to sustain the behaviour promoted during the intervention afterwards, possibly over a period of months (see Steptoe (2019) for a review). While effects on biomarkers turned out insignificant, the fact that they consistently moved into the hypothesised direction still suggests a promising avenue for future exploration amongst individuals specifically with higher levels of depressive symptoms at $t = 0$, and in particular, for long-run follow-up measurement.

Compared to the literature, impacts on self-reported outcomes were large: the course increased participants' life satisfaction on a zero-to-ten scale by about one point, which is more than being partnered as opposed to being single (+0.6) (Clark et al., 2018). Impacts were stronger than those found in trials funded by the UK Big Lottery Fund, which financed a wide range of wellbeing programmes (fourteen portfolios, each consisting of three to 34 actual trials) from 2008 to 2015 at a volume of £200 ($251) million. Trials typically included community-based activities such as community gardens or sports events. As a conservative estimate, they increased life satisfaction on a zero-to-ten scale by, on average, 0.5 points for six months post-treatment (New Economics Foundation-Centre for Local Economic Strategies, 2013). Different from the "Exploring What Matters" course, however, these trials targeted specific groups with mental health needs.

Finally, impacts were highly comparable to those of ENHANCE (for life satisfaction, about one point in "Exploring What Matters" vs. 1.1 points in ENHANCE), a 12-week wellbeing course focusing primarily on positive habits, skills, and attitudes, which is the most comparable intervention and which can be delivered both offline and online, with little reported differences between both delivery modes (Heintzelman et al., 2020).[20] The authors were able to provide evidence of positive impacts over an even longer period of six months post-treatment. Note that this six-month post-treatment period includes a three-month sub-period in which participants who had finished the course were repeatedly followed up: in the offline version, this included an alternating series of bi-weekly phone calls (of ten to fifteen minutes duration each) and in-person group sessions (of two hours duration each) during these three months; participants in the online version received six bi-weekly e-mails during this period.

"Exploring What Matters" differs from ENHANCE in several, aspects. We limit our comparisons to the offline version of ENHANCE because there exists, to date, no online version of "Exploring What Matters".[21] Different from ENHANCE, "Exploring What Matters" is led exclusively by non-expert volunteers (essentially laypeople), whereas ENHANCE relies on graduate-level trained clinicians. This is interesting, because it shows that laypeople without any specific academic background can be effectively utilised to systematically improve the wellbeing and pro-sociality of others. In fact, the manualisation of the "Exploring What Matters" course and its reliance on volunteer laypeople as course leaders make it highly cost-effective for face-to-face settings: costing only £90 ($113) per WELLBY (a one-point increase in life satisfaction on a zero-to-ten scale for one individual for one year), it is well above the advocated wellbeing cost-effectiveness threshold of about £2500 ($3139) derived from marginal National Health Service (NHS) spending in the UK (Clark et al., 2018), and well below the individual willingness to pay for one WELLBY of about £9000 ($11,300) derived from marginal health improvements (Huang et al., 2018).[22,23] Another difference between the two courses is the period after the course has ended. Different from ENHANCE, "Exploring What Matters" includes no labour-intensive maintenance period (i.e. bi-weekly alternating phone calls and group sessions), which from a costing point of view should be seen as part of the intervention package and which has implications for cost-effectiveness. Such a period may not be necessary, considering the similarity in outcomes between the two courses.

Regardless of these differences, ENHANCE and the "Exploring What Matters" course show remarkable similarities in terms of impacts and demonstrate that the wellbeing of healthy adults in the general population can be effectively improved by means of intervention. An important, policy-relevant question is how average people can be motivated to take up wellbeing interventions (for example, by targeting their expectancies or subjective valuations of interventions), especially if they may not believe *ex-ante* in their effectiveness and interventions may therefore represent credence goods, i.e. goods of which the value only becomes apparent upon consumption.

---

[20] The impact of this course has been studied using a waitlist randomisation design, as in our paper, and the authors found an impact of about 0.5 between baseline and posttest on life satisfaction measured on a one-to-five multi-item summed scale (the Satisfaction With Life Scale) (Heintzelman et al., 2020, Table 3). With the caveat that both measures of life satisfaction are not perfectly comparable, rescaling this item to a zero-to-ten scale yields an impact of about $0.5 \times (11/5) = 1.1$.

[21] In light of Covid-19, Action for Happiness, the charity running the "Exploring What Matters" course, has developed a new version of the course optimised for online delivery, due to be launched in 2021. During the pandemic, over 100 local groups have conducted the course online using Zoom and over 5,000 participants have been involved.

[22] The wellbeing cost-effectiveness threshold of about £2,500 derived from marginal NHS spending can be calculated as follows: the NHS approves treatment if the QALY per cost ratio is 1 / £25,000. Since QALYs are measured on a scale from zero-to-one and life satisfaction is measured on a scale from zero-to-ten, the translated advocated wellbeing cost-effectiveness threshold becomes $(1 / £25{,}000) \times 10$. See Layard (2016), Clark et al. (2018), and Frijters et al. (2020) for the concept of WELLBY and Frijters and Krekel (2021) for a discussion of wellbeing cost-effectiveness analysis.

[23] We made the assumption that impacts are sustained for at least one year. If we assume that they are sustained for two months only, for which we have suggestive evidence, the course would cost £540 ($678) per WELLBY.

Our study has several shortcomings. The most important one is that significant effects on self-reported outcomes were not mirrored by biomarkers. Impacts at $t = 1$ may thus have reflected participants' euphoria of having finished the course, placebo effects, or social desirability if course participants tried to please course leaders. Although none of these can be ruled out for sure, we argue that it is unlikely that our impacts were primarily driven by these artefacts. First, recall that the atmosphere during data collection (including $t = 0$, $t = 1$, and $t = 2$) was kept strictly neutral according to protocol, and that participants could meet and chat to each other only *after* data collection had finished. Second, there was evidence for sustained impacts: it is unlikely that placebo effects were sustained two months post-treatment. Moreover, impacts at $t = 2$ were similar (if not stronger) than at $t = 1$: it is unlikely that, two months after having completed the last survey, participants perfectly recalled their previous responses. Likewise, the fact that different types of self-reported outcomes, particularly, PHQ-9 and GAD-7, point into the same direction makes us more confident in that our identified treatment effects are not driven exclusively by demand effects. Arguably, PHQ-9 and GAD-7 should be less susceptible to such effects, because they *(a)* are multi-item summary scales (and hence relatively less prone to them), *(b)* ask about actual experiences during the past two weeks (for example, trouble falling or staying asleep, or sleeping too much), and *(c)* are not framed around the notion of "happiness" (which the course is advertised to promote). Finally, data collection was strictly anonymous, and there was little incentive for participants to answer in a strategic or socially desirable way. Likewise, anonymous online surveys from the universe of courses conducted showed similar impacts. They also point against observer effects: for participants who completed online surveys, no field experiment was salient.

Despite these protocols, two other types of placebo effects are thinkable: first, participants self-selected into the intervention (i.e. knowing that it aims at increasing their happiness) and were likely to be actively looking to improve their lives. A question then arises as to whether our identified treatment effects are due to placebo effects (i.e. motivated cognition) to which self-selected participants may be especially susceptible. Alternatively, one might argue that self-selected participants may be especially motivated to "work hard" in order to improve their lives, i.e. pure motivational effects. Unfortunately, our study design does not allow us to disentangle these effects, but the literature provides evidence on their relative importance. Lyubomirsky et al. (2011) show that self-selection strengthens treatment effects, but only when interacted with a wellbeing-enhancing treatment (as opposed to a neutral control). Moreover, the authors show that self-selected participants put more effort into treatment compared to non-self-selected participants. Hence, self-selection seems to matter, but not so much because of motivated cognition. Rather, it seems that self-selected participants bring with them more positive behavioural attitudes towards treatment.

A second placebo effect may arise from the upfront donation (£90) that participants may make in order to cover costs: one could argue that, because participants paid upfront, they may report a higher wellbeing *ex-post* due to cognitive dissonance. Although we cannot fully exclude this possibility, the combination of *(i)* the rather small amount (i.e. between £90 / (8×2) = £5.6 to £90 / (10×2) = £4.5 per course hour for a course duration of between 16h and 20h); *(ii)* the relatively long duration between payment and outcome measurement of more than two months; and *(iii)* the fact that course participants were not primarily from the lower end of the income distribution reduces the likelihood of significant placebo effects from the possibility of making an upfront donation.

Another shortcoming was the waitlist randomisation design: the choice of this design was motivated by the fact that – in our non-clinical, general adult population, and local community intervention context – there exists no natural, credible control group that could lend itself as a counterfactual business-as-usual. At the same time, alternate double-blind impact study designs with placebo control groups are difficult to implement in the context of course-based social-psychological interventions (Herbert and Gaudiano, 2005). On the one hand, a placebo (for example, having meetings at the same time as the treatment group but in an unstructured format without delivering course contents) could have helped to better isolate and identify the active ingredients of the intervention (for example, specific course contents *versus* socialising or disrupting the daily routine), beyond the self-reported changes in information and behaviour that we document. On the other hand, a placebo that eliminates (ideally) one specific channel is difficult to find and implement, especially in case of in-person courses involving several sessions over a long period of time. Ideally, one would want to work with multiple control groups and placebos, which can easily become quite complex. At a conceptual level, this raises the question of whether or not elements like socialising or disrupting the daily routine are themselves active parts of the intervention package.

Our waitlist randomisation, therefore, balanced these challenges while adhering as closely as possible to evidence-based practice in social science. Nevertheless, it has drawbacks. The most important one is that being waitlisted itself could be a treatment. Bias could have gone both ways. We found little evidence for either: between $t = 0$ and $t = 1$, there were little significant differences in outcomes and covariates for the waitlisted control group, except for mindfulness and meditation (which the waitlisted control group seemed to practice more at $t = 0$). Excluding individuals for whom this behavioural change occurred between $t = 0$ and $t = 1$ left our findings unchanged.[24]

Future research may build on and extend the evidence established in this trial, for example, by looking at long-term impacts that go beyond two months post-treatment. Moreover, it may be interesting to look at behavioural spillovers from one life domain to another or wellbeing spillovers between individuals. We found participants who were initially in more mental distress to benefit more from the course. A larger sample size could help stratifying results by demographics and other participant characteristics, providing useful insights into targeting particular groups of people more effectively. It may

[24] Results are available upon request.

also help resolve power issues with biomarkers. Finally, motivated by the growing literature on mentoring and advice-giving in social psychology rooted in self-perception theory and advocacy, studying the causal effect of the course on the wellbeing of facilitators (i.e. the volunteers who lead the course) would be a promising avenue for future research.

7. Conclusion

Our study shows that wellbeing is not fixed but can be changed by means of intervention, cost-effectively and at scale, and that self-reported impacts are sustained over time. In particular, exposing people to the scientific evidence base on what has been found to cause wellbeing (even when presented by non-expert laypeople), jointly discussing this evidence, and committing to make behavioural changes to daily routines can have lasting impacts on wellbeing. This speaks against a set point of wellbeing around which individuals fluctuate and return to by adapting to changes in life circumstances (Brickman and Campbell, 1971). Rather, the evidence presented here speaks for an expectancy-value approach to behaviour change (Battle, 1965), in which individuals – once they update their beliefs about what matters to their wellbeing, change their behaviour initially, and experience an initial increase in wellbeing – may change their behaviour more permanently, with then sustained impacts on wellbeing. To the extent that people do not anticipate or believe in such interventions, these may constitute credence goods and there may be a role to play for policy to accredit their effectiveness and disseminate that information.

This has important implications for economics: apart from wellbeing being a significant predictor of economic behaviour and individual-level outcomes such as productivity and income (De Neve and Oswald, 2012; Oswald et al., 2015; Bellet et al., 2020), health (Graham and Pinto, 2019), voting (Liberini et al., 2017), or organisation-level productivity and profitability (Krekel et al., 2019), there are important implications for measuring societal progress more generally. If wellbeing is not fixed and adaptation is not inevitable (e.g. we know that there is no full adaptation to unemployment, cf. Clark et al., 2008), wellbeing can be used as a meaningful indicator to measure societal progress, and help direct policy attention towards areas in which there may be little adaptation (such as lack of social relationships, unemployment, lack of community cohesion and trust, or mental health), and by the same token, towards more wellbeing-improving activities.

The Easterlin Paradox (Easterlin, 1974, 2019) shows that, despite substantial increases in GDP per capita, wellbeing has been largely stagnant in many developed countries over the past decades, or even declined for some population groups (Stevenson and Wolfers, 2009). The finding that wellbeing can improve when redirected towards certain behaviours, combined with the growing evidence base on its causes and consequences, underlines its usefulness as an indicator for measuring how we are doing as a society, which is a core activity of the economics profession.

Funding

Funding from the John Templeton Foundation and the What Works Centre for Wellbeing's evidence programme is gratefully acknowledged.

Declarations of Competing Interest

Christian Krekel, Jan-Emmanuel De Neve, and Daisy Fancourt: none. Richard Layard is cofounder of the charity “Action for Happiness”, which runs the “Exploring What Matters” course. He has advised in the design of the course but has had no active role in its impact evaluation.

Acknowledgments

We are heavily indebted to Fulvio D'Acquisto and Martin Gross at the William Harvey Research Institute, Queen Mary University of London, for helping us with the logistics of temporarily storing our biomarker samples. We are thankful to Ed Diener, Carol Graham, Paul Frijters, Claryn Kung, and Ashley Whillans, as well as seminar participants at the London School of Economics and various other places, for helpful comments and suggestions. Lucía Macchia and Ekaterina (Katya) Oparina provided excellent research assistance. A special thanks goes to the volunteer course leaders at Action for Happiness, course participants, and to the Action for Happiness staff, in particular Keith Cowley, Alex Nunn, and Mark Williamson. We thank the editor, associate editor, and two anonymous referees for their very helpful comments and suggestions.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jebo.2021.05.021.

References

Aknin, L.B., Barrington-Leigh, C.P., Dunn, E.W., Helliwell, J.F., Burns, J., Biswas-Diener, R., Kemeza, I., Nyende, P., Ashton-James, C.E., Norton, M.I., 2013. Pro-social spending and well-being: cross-cultural evidence for a psychological universal. J. Pers. Soc. Psychol. 104 (4), 635–652.
Battle, E.S., 1965. Motivational determinants of academic task persistence. J. Pers. Soc. Psychol. 2 (2), 209–2018.
Bellet, C.S., De Neve, J.E., Ward, G., 2020. Does employee happiness have an impact on productivity? CEP Discuss. P. 1655.

Bohlmeijer, E., Prenger, R., Taal, E., Cuijpers, P., 2010. The effects of mindfulness-based stress reduction therapy on mental health of adults with a chronic medical disease: a meta-analysis. J. Psychosom. Res. 68 (6), 539–544.
Bolier, L., Haverman, M., Westerhof, G.J., Riper, H., Smit, F., Bohlmeijer, E., 2013. Positive psychology interventions: a meta-analysis of randomized controlled studies. BMC Public Health 13, 119.
Borgonovi, F., 2008. Doing well by doing good. The relationship between formal volunteering and self-reported health and happiness. Soc. Sci. Med. 66, 2321–2334.
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, 822–848.
Brickman, P., Campbell, D., 1971. Hedonic relativism and planning the good society. In: Apley, M.H. (Ed.), Adaptation-Level Theory: A Symposium. Academic Press, New York.
Carter, T.J., Gilovich, T., 2010. The relative relativity of material and experiential purchases. J. Pers. Soc. Psychol. 98 (1), 146–159.
Chirkov, V.I., Ryan, R.M., Kim, Y., Kaplan, U., 2003. Differentiating autonomy from individualism and independence: a self-determination theory perspective on internalization of cultural orientations and well-being. J. Pers. Soc. Psychol. 84, 97–110.
Church, A.T., Katigbak, M.S., Locke, K.D., Zhang, H., Shen, J., Jesús de Vargas-Flores, J., Tanaka-Matsumi, J., et al., 2013. Need satisfaction and well-being: testing self-determination theory in eight cultures. J. Cross. Cult. Psychol. 44 (4), 507–534.
Clark, A.E., Diener, E., Georgellis, Y., Lucas, R.E., 2008. Lags and leads in life satisfaction: a test of the baseline hypothesis. Econ. J. 118 (529), F222–F243.
Clark, A.E., Flèche, S., Layard, R., Powdthavee, N., Ward, G., 2018. The Origins of Happiness: The Science of Well-Being Over the Life Course. Princeton University Press, Princeton, NJ.
Clark, D.M., Layard, R., Smithies, R., Richards, D.A., Suckling, R., Wright, B., 2009. Improving access to psychological therapy: initial evaluation of two UK demonstration sites. Behav. Res. Ther. 47 (11), 910–920.
Danner, D.D., Snowdon, D.A., Friesen, W.V., 2001. Positive emotions in early life and longevity: findings from the nun study. J. Pers. Soc. Psychol. 80 (5), 804–813.
Deci, E.L., Ryan, R.M., 1985. Intrinsic Motivation and Self-Determination in Human Behaviour. Plenum, New York.
De Neve, J.E., Oswald, A.J, 2012. Estimating the influence of life satisfaction and positive affect on later income using sibling fixed effects. Proc. Nat. Acad. Sci. 109 (49), 19953–19958.
Di Tella, R., Haisken-DeNew, J., MacCulloch, R., 2010. Happiness adaptation to income and to status in an individual panel. J. Econ. Behav. Organ. 76, 834–852.
Diener, E., Suh, E.M., Lucas, R.E., Smith, H.L., 1999. Subjective well-being: three decades of progress. Psychol. Bull. 125 (2), 276–302.
Dolan, P., Krekel, C., Shreedhar, G., Lee, H., Marshall, C., Smith, A., 2021. Volunteering improves wellbeing: evidence from a nationwide micro-volunteering programme. CEP Discuss. P. 1772.
Dowlati, Y., Herrmann, N., Swardfager, W., Liu, H., Sham, L., Reim, E.K., Lanctôt, K.L., 2010. A meta-analysis of cytokines in major depression. Biol. Psychiatry 67, 446–457.
Drouvelis, M., Grosskopf, B., 2016. The effects of induced emotions on pro-social behaviour. J. Public Econ. 134, 1–8.
Dunn, E.W., Aknin, L.B., Norton, M.I., 2008. Spending money on others promotes happiness. Science 319, 1687–1688.
Easterlin, R., 1974. Does economic growth improve the human lot? Some empirical evidence. In: David, R., Reder, R. (Eds.), Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz. Academic Press, New York.
Easterlin, R.A., Angelescu McVey, L., Switek, M., Sawangfa, O., Smith Zweig, J., 2010. The happiness-income paradox revisited. Proc. Nat. Acad. Sci. 107 (52), 22463–22468.
Fancourt, D., Perkins, R., Ascenso, S., Atkins, L., Kilfeather, S., Carvalho, L., Steptoe, A., Williamon, A., 2016. Group drumming modulates cytokine response in mental health services users: a preliminary study. Psychother. Psychosom. 85 (1), 53–55.
Frederick, S., Loewenstein, G., 1999. Hedonic adaptation. In: Kahneman, D., Diener, E., Schwartz, N. (Eds.), Well-Being: The Foundations of Hedonic Psychology. Russell Sage, New York.
Frijters, P., Krekel, C., 2021. A Handbook for Wellbeing Policy-Making. Oxford University Press, Oxford.
Frijters, P., Clark, A.E., Krekel, C., Layard, R., 2020. A happy choice: wellbeing as the goal of government. Behav. Public Policy 4 (2), 126–165.
Graham, C., Pinto, S., 2019. Unequal hopes and lives in the USA: optimism, race, place, and premature mortality. J. Popul. Econ. 32 (2), 665–733.
Godfrin, K.A., van Heeringen, C., 2010. The effects of mindfulness-based cognitive therapy on recurrence of depressive episodes, mental health and quality of life: a randomized controlled study. Behav. Res. Ther. 48 (8), 738–746.
Gu, J., Strauss, C., Bond, R., Cavanagh, K., 2015. How do mindfulness-based cognitive therapy and mindfulness-based stress reduction improve mental health and wellbeing? A systematic review and meta-analysis of mediation studies. Clin. Psychol. Rev. 37 (1), 1–12.
Hausman, L.R.M., Parks, A., Youk, A.O., Kwoh, C.K, 2014. Reduction of bodily pain in response to an online positive activities intervention. J. Pain 15, 560–567.
Heintzelman, S.J., Kushlev, K., Lutes, L.D., Wirtz, D., Kanippayoor, J.M., Leitner, D., Oishi, S., Diener, E., 2020. ENHANCE: evidence for the efficacy of a comprehensive intervention program to promote subjective well-being. J. Exp. Psychol. Appl. 26 (2), 360–383.
Herbert, J.D., Gaudiano, B.A., 2005. Moving from empirically supported treatment lists to practice guidelines in psychotherapy: the role of the placebo concept. J. Clin. Psychol. 61 (7), 893–908.
Huang, L., Frijters, P., Dalziel, K., Clarke, P., 2018. Life satisfaction, QALYs, and the monetary value of health. Soc. Sci. Med. 211, 131–136.
Kahneman, D., 2000. Experienced utility and objective happiness: a moment-based approach. In: Kahneman, D., Tversky, A. (Eds.), Choice, Values, and Frames. Cambridge University Press, New York.
Krekel, C., Swanke, A., De Neve, J.E., Fancourt, D., 2020. Are happier people more compliant? Global evidence from three large-scale surveys during covid-19 lockdowns. IZA Discuss. P. 13690.
Krekel, C., Ward, G., De Neve, J.E., 2019. Employee wellbeing, productivity and firm performance. CEP Discuss. P. 1605.
Kuhn, P., Kooreman, P., Soetevent, A., Kapteyn, A., 2011. The effects of lottery prizes on winners and their neighbors: evidence from the dutch postcode lottery. Am. Econ. Rev. 101 (5), 2226–2247.
Guardia, La, G., J., Ryan, R.M., Couchman, C.E., Deci, E.L, 2000. Within-person variation in security of attachment: a self-determination theory perspective on attachment, need fulfillment, and well-being. J. Pers. Soc. Psychol. 79 (3), 367–384.
Layard, R., 2016. Wellbeing measurement and cost-effectiveness analysis. CEP Work. P..
Layard, R., Clark, A.E., Cornaglia, F., Powdthavee, N., Vernoit, J., 2014. What predicts a successful life? A life-course model of well-being. Econ. J. 124 (580), F720–F738.
Liberini, F., Redoano, M., Proto, E., 2017. Happy voters. J. Public Econ. 146, 41–57.
Lucas, R., 2005. Time does not heal all wounds: a longitudinal study of reaction and adaptation to divorce. Psychol. Sci. 16 (12), 945–950.
Lucas, R., Clark, A.E., 2006. Do people really adapt to marriage? J. Happiness Stud. 7 (4), 405–426.
Lyubomirsky, S., Dickerhoof, R., Boehm, J.K., Sheldon, K.M., 2011. Becoming happier takes both a will and a proper way: an experimental longitudinal intervention to boost well-being. Emotion 11 (2), 391–402.
Meier, S., Stutzer, A., 2008. Is volunteering rewarding in itself? Economica 75 (297), 39–59.
Menzel, P., Dolan, P.H., Richardson, J., Olsen, J., 2002. The role of adaptation to disability and disease in health state valuation: a preliminary normative analysis. Soc. Sci. Med. 55 (12), 2149–2158.
Milyavskaya, M., Koester, R., 2011. Psychological needs, motivation, and well-being: a test of self-determination theory across multiple domains. Pers. Individ. Dif. 50 (3), 387–391.
Miller, A.H., Raison, C.L., 2016. The role of inflammation in depression: from evolutionary imperative to modern treatment target. Nat. Rev. Immunol. 16 (1), 22.

Miller, G.E., Chen, E., Zhou, E.S., 2007. If it goes up, must it come down? Chronic Stress and the hypothalamic-pituitary-adrenocortical axis in humans. Psychol. Bull. 133 (1), 25–45.
New Economics Foundation – Centre for Local Economic Strategies (2013). Big Lottery Fund National Well-Being Evaluation. Draft Report.
National Health Service (NHS) Long Term Plan, 2019. The NHS Long Term Plan 2019 Online https://www.longtermplan.nhs.uk/wp-content/uploads/2019/08/nhs-long-term-plan-version-1.2.pdf .
Oswald, A.J., Gardner, J., 2006. Do divorcing couples become happier by breaking up? J. R. Stat. Soc. Ser. A Stat. Soc. 169 (2), 319–336.
Oswald, A.J., Powdthavee, N., 2008. Does happiness adapt? A longitudinal study of disability with implications for economists and judges. J. Public Econ. 92 (5–6), 1061–1077.
Oswald, A.J., Proto, E., Sgroi, D., 2015. Happiness and productivity. J. Lab. Econ. 33 (4), 789–822.
Powdthavee, N., 2008. Putting a price tag on friends, relatives, and neighbours: using surveys of life satisfaction to value social relationships. J. Socio Econ. 37 (4), 1459–1480.
Querstret, D., Morison, L., Dickinson, S., Cropley, M., John, M., 2020. Mindfulness-based stress reduction and mindfulness-based cognitive therapy for psychological health and well-being in nonclinical samples: a systematic review and meta-analysis. Int. J. Stress Manag. 27 (4), 394–411.
Romano, J.P., Wolf, M., 2005a. Stepwise multiple testing as formalized data snooping. Econometrica 73 (4), 1237–1282.
Romano, J.P., Wolf, M., 2005b. Exact and approximate stepdown methods for multiple hypothesis testing. J. Am. Stat. Assoc. 469 (100), 94–108.
Romano, J.P., Wolf, M., 2016. Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. Stat. Probab. Lett. 113, 38–40.
Ryan, R.M., Deci, E.L., 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am. Psychol. 55 (1), 68–78.
Sedlmeier, P., Eberth, J., Schwarz, M., Zimmermann, D., Haarig, F., Jaeger, S., Kunze, S., 2012. The psychological effects of meditation: a meta-analysis. Psychol. Bull. 138 (6), 1139–1171.
Sin, N.L., Lyubomirsky, S., 2009. Enhancing well-being and alleviating depressive symptoms with positive psychology interventions: a practice-friendly meta-analysis. J. Clin. Psychol. 65 (5), 467–487.
Steptoe, A., 2019. Happiness and health. Annu. Rev. Public Health 40, 339–359.
Steptoe, A., Wardle, J., 2011. Positive affect measured using ecological momentary assessment and survival in older men and women. Proc. Nat. Acad. Sci. 108 (45), 18244–18248.
Stevenson, B., Wolfers, J., 2009. The paradox of declining female happiness. Am. Econ. J. Econ. Policy 1 (2), 190–225.
Stutzer, A., Frey, B., 2006. Does marriage make people happy, or do happy people get married? J. Socio Econ. 35 (2), 326–347.
Taylor, C.T., Lyubomirsky, S., Stein, M.B., 2017. Upregulating the positive affect system in anxiety and depression: outcomes of a positive activity intervention. Depress. Anxiety 34, 267–280.
Tomitaka, S., Kawasaki, Y., Ide, K., Akutagawa, M., Ono, Y., Furukawa, T.A., 2018. Stability of the distribution of patient health questionnaire-9 scores against age in the general population: data from the national health and nutrition examination survey. Front Psychol. 9, 390.
Wiles, N., Thomas, L., Turner, N., Garfield, K., Kounali, D., Campbell, J., Kessler, D., et al., 2016. Long-term effectiveness and cost-effectiveness of cognitive behavioural therapy as an adjunct to pharmacotherapy for treatment-resistant depression in primary care: follow-up of the CoBalT randomised controlled trial. Lancet Psychiatry 3, 137–144.
Wium-Andersen, M.K., Nielsen, S.F., 2013. Elevated C-reactive protein levels, psychological distress, and depression in 73,131 individuals. JAMA Psychiatry 70 (2), 176–184.