European Child & Adolescent Psychiatry
https://doi.org/10.1007/s00787-026-03152-x

RESEARCH

Implementation dosage and depressive symptoms following school-based mindfulness and relaxation: Complier Average Causal Effect estimation in the INSPIRE cluster randomised trial

Emma Thornton[1] · Christopher Knowles[1] · Margarita Panayiotou[1] · Joao Santos[1] · Emma Ashworth[2] · Carla Mason[1] · Daniel Hayes[3,4] · Abigail Thompson[3] · Jessica Deighton[3] · Rosie Mansfield[5] · Anna March[3] · Paul Stallard[6] · Neil Humphrey[1]

Received: 23 September 2025 / Accepted: 28 July 2026

Abstract

Implementation variability is hypothesized to be a key treatment effect modifier of universal, school-based mental health interventions, but the evidence base remains meagre. We used Complier Average Causal Effect (CACE) estimation to determine if dosage variability modified treatment effects of relaxation techniques and mindfulness-based exercises in the INSPIRE cluster randomised trial. Analyses were split by phase of education (mindfulness primary: $k=65$ primary schools, $N=3{,}630$ pupils; mindfulness secondary: $k=29$ secondary schools, $N=3{,}708$ pupils; relaxation primary: $k=66$ primary schools, $N=3{,}800$ pupils; relaxation secondary: $k=26$ secondary schools, $N=3{,}252$ pupils) and considered compliance effects at the 50th (moderate) and 75th (high) percentiles in the distribution of intervention dosage (total minutes delivered). Despite null intent to treat findings, we observed compliance effects for both interventions. In primary schools, moderate (320 min) and high (376 min) compliance to relaxation techniques significantly *reduced* depressive symptoms (d=-1.52 and d=-1.48, respectively). In secondary schools, moderate (207 min) and high (285 min) compliance to mindfulness-based exercises also significantly *reduced* depressive symptoms (d=-1.15 and d=-1.40, respectively). Contrastingly, in primary schools, moderate (325 min) and high (393 min) compliance to mindfulness-based exercises significantly *increased* depressive symptoms ($d=0.32$ and $d=0.48$, respectively). In secondary schools, moderate (180 min) and high (275 min) compliance to relaxation techniques also significantly *increased* depressive symptoms ($d=0.67$ and $d=0.83$, respectively). The effects of universal, school-based mental health interventions promoting relaxation and mindfulness are contingent on the delivery of a minimally effective dose. However, these interventions may not be appropriate in all educational contexts and could result in iatrogenic effects.

Trial registration: Trial registration number: ISRCTN16386254. Trial registration date:04/09/2018.

Keywords Cluster randomised controlled trial · Complier average causal effect analysis · Intervention dosage · Mindfulness · Relaxation · Universal school-based interventions

✉ Emma Thornton
emma.thornton@manchester.ac.uk

[1] Manchester Institute of Education, University of Manchester, Manchester, UK

[2] School of Psychology, Liverpool John Moores University, Liverpool, UK

[3] Evidence Based Practice Unit, University College London and Anna Freud, London, UK

[4] Department of Behavioural Science and Health, University College London, London, UK

[5] Centre for Longitudinal Studies, University College London, London, UK

[6] Department of Health, University of Bath, Bath, UK

Published online: 12 August 2026

Introduction

Approximately 1 in 5 (20.3%) children and young people aged 8–16 in England have a probable mental health disorder (with a further 1 in 8 (12%) exhibiting symptoms indicative of possible mental health disorder) [1], among whom internalising problems such as depression and anxiety are the most prevalent [2, 3]. School-based interventions are seen as increasingly central to efforts to prevent the development, maintenance or escalation of such difficulties among children and young people [4], particularly given data highlighting increased prevalence in recent years [5]. These interventions can vary significantly in terms of both their form (i.e., universal, targeted, indicated) and function (i.e., informational, developmental, environmental) [6]. This Article focuses on two *universal* (i.e., delivered to all, irrespective of need), *developmental* (i.e., promotion of adaptive behaviours, focusing on the development of skills hypothesized to support/underpin mental health) interventions: mindfulness-based exercises; and relaxation practices. These interventions were designed specifically to address internalising symptoms, motivated by the Department for Education's desire to find efficacious school-based approaches to address the growing youth mental health crisis [7].

Work on universal interventions is informed by the prevention paradox, which demonstrates that focusing on the low/moderate risk majority offers greater public health benefit than intervening only with the high-risk minority [8]. Meta-analytic evidence indicates that universal, school-based interventions can produce meaningful reductions in depression and anxiety [9], but a number of treatment effect moderators (e.g., outcome domain, programme content, phase of education) are evident, in addition to concerns about overall research quality [10, 11]. A 2017 national survey of schools and colleges in England indicated that 73% were already implementing universal developmental interventions [12], and indeed this has since become mandatory as part of statutory relationships, sex and health education provision [13]. However, a recent regional survey indicated that the most frequently used programmes in schools had little or no evidence of efficacy [14].

A Mindfulness-Based Intervention (MBI) and Relaxation-Based Practice Intervention (RBPI) were selected by the Department for Education following an internal review process, which suggested they were being utilized in English schools despite little or no evidence of their efficacy [15]. Indeed, with reference to MBIs, Dunning et al. (p.135) conclude in their recent meta-analysis that enthusiasm for them, "has arguably run ahead of the evidence" [16]. Regarding RBPIs, while there is meta-analytic evidence of their efficacy for young people exhibiting prodromal symptoms of mental health difficulties (i.e., in targeted form) [17], little is known about their impact when used as a universal preventive strategy in schools. In this Article, we examine the extent to which implementation dosage (the cumulative amount of an intervention received, expressed in total minutes of exposure) of a MBI and a RBPI modifies treatment effects on primary outcomes that were observed in intention to treat (ITT) analyses in the INSPIRE randomised controlled trial (RCT) [18].

Dosage is an obvious candidate as a modifier of treatment effects of universal school-based interventions, given the basic logic that effective skill development likely requires a minimally effective level of exposure that may not be reached in all settings, as a consequence of the inevitable variability in implementation due to issues such as time constraints and/or competing priorities [19]. However, it has largely been neglected, both at individual study and meta-analytic levels. Exceptions illustrate its importance. For example, a recent RCT showed that total minutes of exposure to the Good Behaviour Game modified its impact on children's externalizing behaviour problems (from a null effect in intention to treat (ITT) analysis to a substantial reduction when exposure exceeded 1,030 min over the two year trial period) [20]. Mirroring this, a recent meta-analysis of universal social and emotional learning interventions found that implementation moderated treatment effects on externalizing behaviour problems [21].

We used Complier Average Causal Effect (CACE) estimation to produce treatment effect estimates of MBI and RBPI that account for variability in dosage. In cluster RCTs, which require multilevel analyses, CACE estimation involves fitting mixture models (latent class analysis) where two latent classes are estimated to identify compliers and non-compliers. In the case of an intervention where there is no clear theoretical or empirical threshold for what might constitute compliance, and the cumulative amount of an intervention received in a specified implementation period can vary between 0 and X minutes (as in MBI and RBPI, which are non-curricular and instead involve recommended practices which may vary significantly in both frequency and duration – see Procedures in Method), standard practice is to use cut-points (e.g., 50th, 75th percentiles) in the empirical distribution of dosage data [20].

For pupils randomised to the intervention arms, compliance status is directly observed using implementer-reported dosage data. In contrast, for those allocated to the control condition, compliance status is unobserved and must be inferred. Under maximum likelihood estimation, these pupils are probabilistically classified as either compliers or non-compliers, with control compliers representing a counterfactual group who would have adhered to the intervention had they been assigned to receive it. Whereas

intention-to-treat (ITT) analyses estimate the effect of assignment to the intervention (implicitly assuming successful implementation), CACE quantifies the causal effect of compliance by comparing compliers in the intervention arm with their counterparts in the control condition, thereby providing a robust estimate of intervention effectiveness under conditions of adequate programme delivery [22, 23].

The analyses reported in this Article, undertaken as part of the INSPIRE trial, make significant contributions to the literature. CACE estimation provides an important complement to ITT analyses (which may underestimate treatment effects when there is considerable variability in implementation) by establishing effects for those receiving a hypothesized minimally effective dose of an intervention as opposed to those merely offered it (as in ITT) [23]. We predicted that for each intervention (MBI, RBPI), significantly improved primary outcomes would be observed, on average, for children and young people in intervention arm schools that complied (moderate compliance: 50th percentile; high compliance: 75th percentile), when compared to 'would-be' compliers in control arm schools. We further hypothesized that the magnitude of treatment effects would be larger in the context of high compliance (as opposed to moderate compliance) for these interventions.

Methods

Trial design and participants

Full details of the INSPIRE RCT are provided in its study protocol and main impact paper [18, 24]. In short, trials were delivered in three waves spanning 2018–2024 (though COVID-19 disruption and school closure led to wave 2 being abandoned). Young people were eligible to take part if their school was eligible [18], they were not opted out of the trial by their parent/guardian, and they provided written assent after reading a participant information sheet. This article focuses on the MBI and RBPI components of the INSPIRE trial; CACE analyses for the third INSPIRE intervention (a mental health curriculum based intervention called Strategies for Safety and Wellbeing) are reported elsewhere, due to conceptual differences in programme content and how dosage was captured [25]. Importantly, each school was assigned to a single intervention arm eliminating risk that interventions would confound one another. Both interventions were delivered in primary (Year groups 4 and 5) and secondary (Year groups 7 and 8) schools, with a roughly equal split across year groups in MBI, RBPI, and control schools (range 49–51% for any given year group).

The CACE analyses reported in this Article included all pupils participating in wave 1 (2018) and wave 3 (2022) of INSPIRE who were not missing data on covariates and/or the outcome *and* compliance data (see Statistical Analysis section below). Follow up data for pupils participating in wave 2 is missing due to the COVID-19 pandemic; these pupils are therefore excluded from analyses. The analytical samples for each CACE analysis were as follows: MBI primary ($n=65$ schools; $n=3{,}630$ pupils); MBI secondary ($n=29$ schools; $n=3{,}708$ pupils); RBPI primary ($n=66$ schools; $n=3{,}800$ pupils); RBPI secondary ($n=26$ schools; $n=3{,}252$ pupils) and included trial usual practice schools (INSPIRE primary or INSPIRE secondary, accordingly).

Allocation and masking

The flow of primary and secondary schools/pupils through the trial (exclusive of those recruited during the abandoned wave 2) is illustrated in Figs. 1 and 2, respectively. CACE analysis requires knowledge of treatment assignment and thus, it is not possible to blind analysts for this type of model estimation.

Procedures

The MBI and RBPI were compared with a control group in which the usual school curriculum was delivered. All interventions were implemented after baseline assessment and random allocation, which took place in the autumn term of a given school year. Follow-up assessment occurred after the January to March (inclusive) implementation period.

The MBI was developed by the Anna Freud Schools Division and comprised activities such as mindful breathing and exercises focusing on self-awareness of emotions, thoughts and sensations. Exercises focussing on the mind (for example, breath counting), body (for example, a body scan exercise), and the world (for example, mindful seeing) were included. Guidance materials were provided, with a total of 26 different exercises comprising five mindful breathing exercises, and seven activities for each of the body, mind, and world. There was overlap in activities for primary and secondary schools (for example, both featured Balancing as a body exercise and Thoughts in The Sky as a mind exercise), but some were distinct in order to reflect developmental differences (for example, primary school exercises focused on the mind placed emphasis on imagining distracting thoughts as waves or bubbles, whereas secondary school activities placed emphasis on acknowledging these thoughts). Manuals containing all exercises for primary and secondary schools are available online [26]. The MBI was implemented by trained classroom teachers; recommended frequency and duration was 5 min of exercises each school day during the implementation period, starting with the five mindful breathing exercises in the first week,

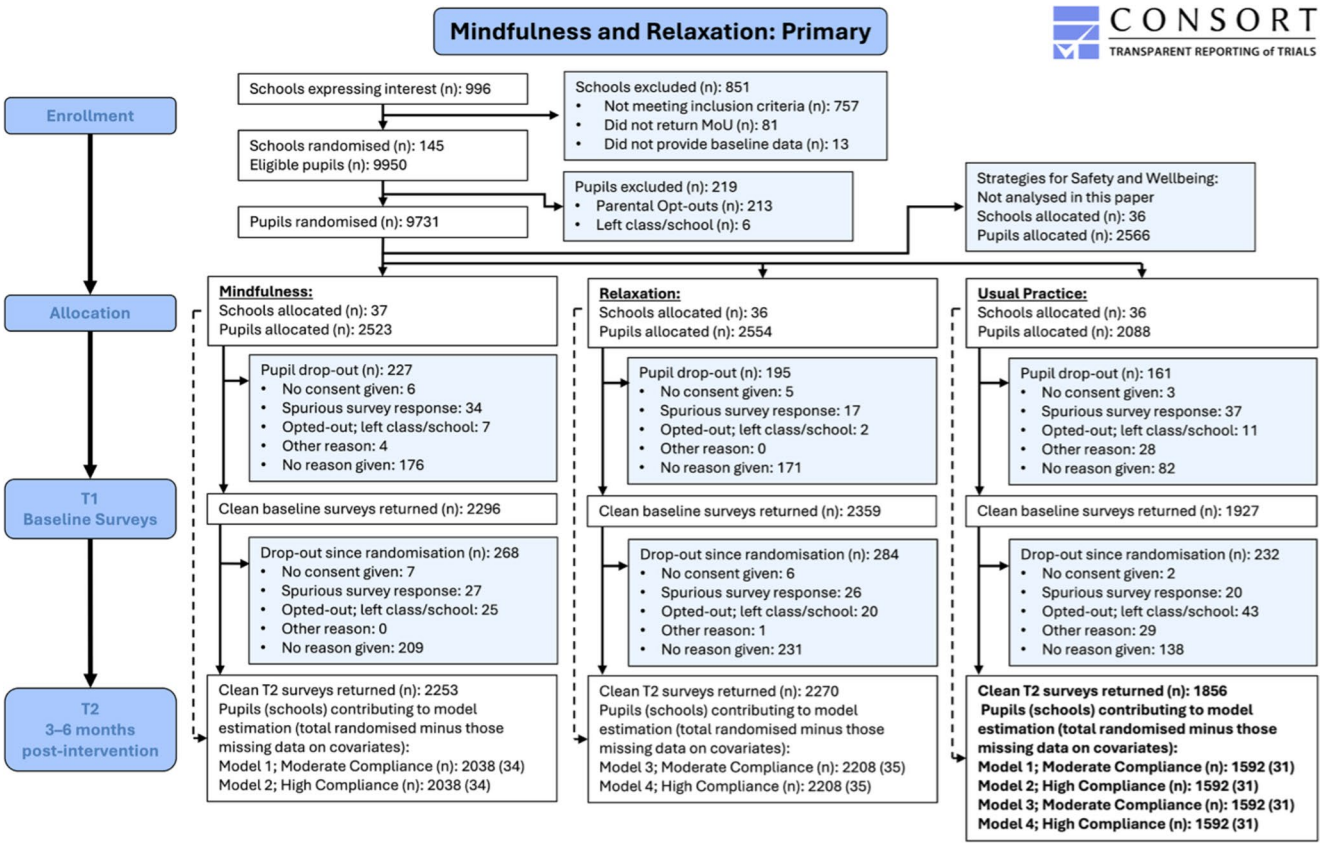

Fig. 1 CONSORT diagram depicting flow of (primary) schools and pupils through the INSPIRE trial

followed by the freedom to choose from the remaining exercises thereafter.

The RBPI was also developed by the Anna Freud Schools Division and included a range of exercises involving deep breathing (for example, a general deep breathing practice) and progressive muscle relaxation (for example, shoulder shrugs). As above, guidance materials were provided, with a total of 20 different exercises (10 deep breathing and 10 progressive muscle relaxation). There was some overlap in activities for primary and secondary schools (for example, both included the Square Breathing exercise and the Hakka Chant progressive muscle relaxation), however, most activities differed between primary and secondary schools to make them age appropriate (for example, one of the primary school breathing exercises asked children to imagine they were a bunny in a field trying to find food, whereas secondary school activities directed participants to focus on their breathing). Manuals containing all exercises for primary and secondary schools are available online [26]. The RBPI was implemented by trained classroom teachers; recommended frequency and duration was 5 min of exercises each school day during the implementation period, alternating weekly between deep breathing and progressive muscle relaxation.

Control schools were asked to continue usual mental health provision, with no restrictions or alterations. A 'current mental health provision' survey was completed at baseline to assess existing staff mental health training, universal school provision (e.g., support skills sessions/workshops), and targeted provision (e.g., cognitive behavioural therapy). A summary percentage score across the range of activities for each domain was derived to quantify the extent of provision at the school level (see the Education for Wellbeing Technical Report for details on the data reduction approach taken) [26]. INSPIRE control primary and secondary schools offered a degree of staff training, universal provision, and targeted provision at baseline (see Table 1).

Measures

Covariates and predictors of compliance

Covariates were selected on the basis of theoretical plausibility (for example, wave of trial was used as a proxy to indicate pre- and post-pandemic delivery) and/or empirical evidence (for example, baseline emotional and behavioural difficulties were found to predict compliance in a CACE analysis of the PATHS curriculum) [27]. For obtaining the

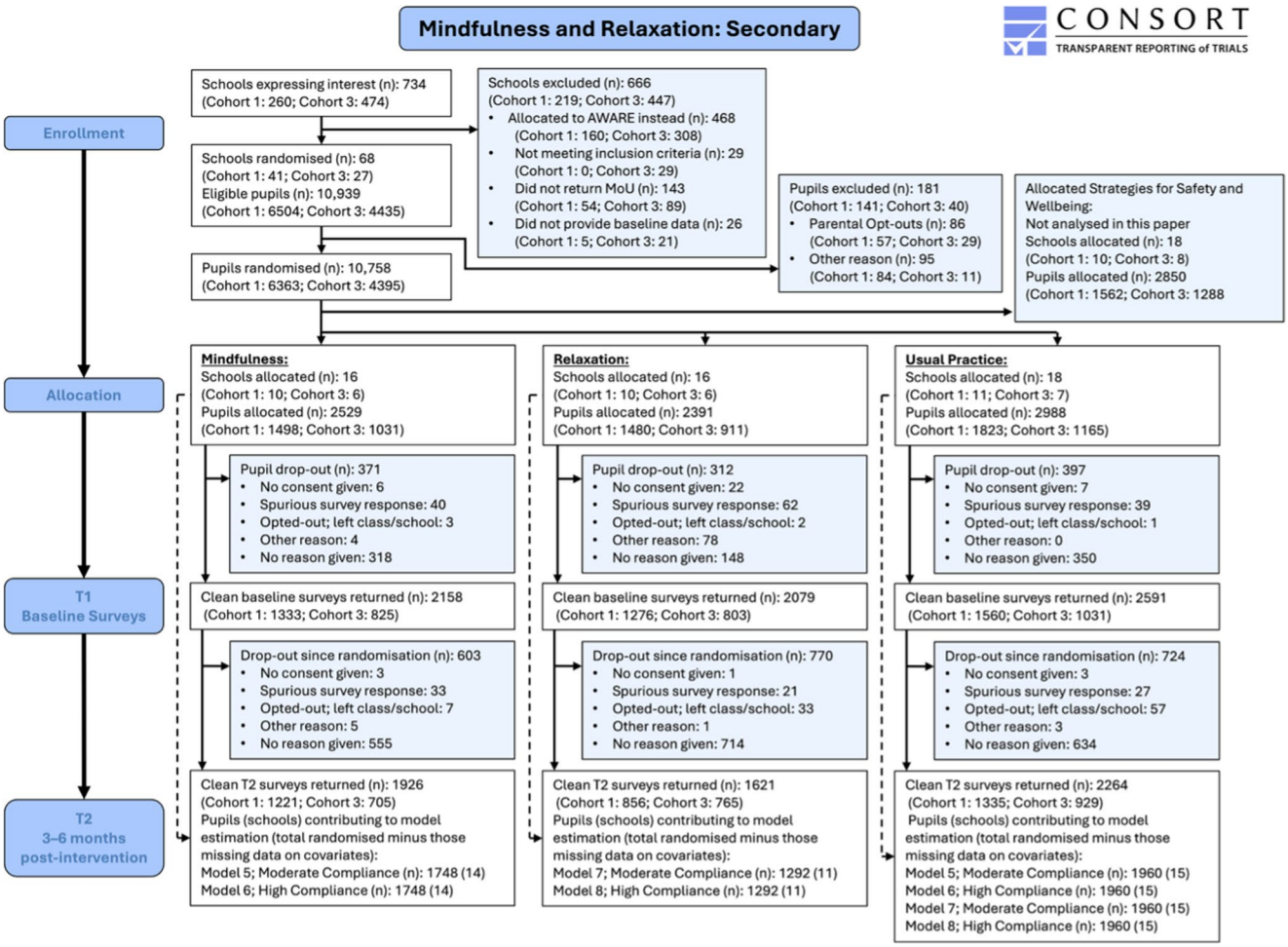

Fig. 2 CONSORT diagram depicting flow of (secondary) schools and pupils through the INSPIRE trial

Table 1 Extent of mental health provision in usual practice (control) schools at baseline

	Trial	
	INSPIRE primary schools	INSPIRE secondary schools
Staff Training	10% (±8.48); range = 0-36.9%	17.7% (±11.1); range = 0-34.5%
Universal Provision	11.9% (±5.62); range = 1.43–28.6%	21.5% (±12.2), Range 5.71–47.1%
Targeted Provision	9.48% (±6.76); range = 0-31.7%	18% (±13.3); range = 3.17–58.7%

CACE estimate, covariates modelled at the within (pupil) level were sex and baseline depressive symptoms (Short Mood and Feelings Questionnaire (SMFQ), see 'Outcome' section below), and covariates modelled at the between (school) level were existing level of mental health provision (yes/no), staff mental health training (see above), trial wave, and whether or not the school was in the top (i.e., most deprived) within-sample tertile according to the percentage of its pupils eligible for free school meals.

Predictors of compliance included the aforementioned covariates, plus baseline behavioural difficulties. Behavioural difficulties were measured using the behavioural difficulties subscale of the self-reported Me & My Feelings Questionnaire (e.g., "I get very angry") [28]. Scores range between 0 and 12, with a higher score indicating more behavioural difficulties,

Implementation data

Compliance data were derived from surveys administered to implementing school staff, who were asked to report (i) start and end dates of implementation; (ii) number of sessions per week; and (iii) average length of sessions. Intervention dosage was modelled as cumulative intervention intensity [29], defined as the total number of minutes delivered across the implementation period, using the following formula: number of weeks*sessions per week*average length of sessions. Although compliance was based on intervention dosage (due to the appropriateness of this dimension for CACE), data pertaining to additional dimensions of implementation

were also collected as part of the INSPIRE trial (fidelity, reach, pupil responsiveness). Descriptive statistics for all implementation dimensions are reported at the outset of the Results section.

Outcome

The primary outcome for both interventions was depressive symptoms, as assessed by the SMFQ, in which participants report the extent to which 13 statements (e.g., "I felt miserable or unhappy") reflect the way that they have felt in the last two weeks, on a three-point scale (not true, sometimes, true). A sum score is calculated across the 13 items, ranging between 0 and 26. Higher scores indicate greater depressive symptoms [30]. The rationale for this choice of primary outcome measure is noted in the main impact paper for the INSPIRE trial [18].

Statistical analysis

Robust CACE estimation relies on several assumptions. First, the stable unit treatment value assumes that the potential outcome of each participant is distinct from the treatment assignment of others. In other words, control participants do not receive the intervention (there is no contamination). Second, CACE assumes participants are randomly allocated to the treatment condition. The third assumption is that there are no defiers (those who do the opposite of what is expected; monotonicity) or always-takers (those who receive treatment regardless of intervention arm). The fourth and final assumption is that there is no effect of the intervention for noncompliers (i.e., exclusion restriction) [22, 23]. Allocation took place at the school level, and control schools did not have access to intervention materials, therefore we were confident the first three assumptions were met [23, 31]. However, the exclusion restriction assumption may be violated in school-based intervention research. Although the resulting bias can be counteracted by the inclusion of strong compliance predictors, candidate variables for compliance prediction are typically collected for other purposes, meaning the selection of available data is often limited. Analyses were performed using Mplus version 8.8 and MplusAutomation in R Studio [32, 33] using two-level models with pupils nested within schools, consistent with the analyse as you randomise principle. Although dosage data were collected from the facilitating member of staff at the classroom-level, Mplus currently supports a maximum of two levels for this type of analysis. Accordingly, clustering was specified at the school level. This approach captures the primary source of dependency introduced by the study design, while variation in implementation at the classroom level is partially accounted for through staff-reported dosage measures disaggregated to the pupil level.

Intervention allocation was defined by a binary variable (*IV*) (0 = control; 1 = intervention). For the identification of the latent compliance variable, a binary indicator of compliance (*U*) is necessary (0 = non-compliers, 1 = compliers) [34]. However, dosage was measured continuously for both the MBI and RPBI interventions (i.e., total minutes of exposure). In the absence of verified compliance thresholds for MBI and RPBI, we used the distribution of the dosage variable to define compliance, adopting the predominant approach in the CACE literature, whereby the 50th and 75th percentiles are used to indicate moderate and high compliance, respectively [20, 27, 35]. Compliance data were collected at the teacher level and disaggregated to the pupil level, so that each pupil had a compliance value based on their classroom membership. As is standard in CACE, compliance for those in control schools was labelled as missing.

A series of multi-level (level 1 = pupils; level 2 = schools) mixture models were fitted using moderate and high compliance thresholds for primary and secondary schools. This resulted in the following CACE models: (1) RPBI moderate compliance in primary schools; (2) RPBI high compliance in primary schools; (3) RPBI moderate compliance in secondary schools; (4) RPBI high compliance in secondary schools; (5) MBI moderate compliance in primary schools; (6) MBI high compliance in primary schools; (7) MBI moderate compliance in secondary schools; (8) MBI high compliance in secondary schools. For each model, two latent classes were identified wherein the probability that $U = 1$ was fixed at 0% (i.e., a non-complier class) and 100% (a complier class), respectively. Depressive symptoms scores were linearly regressed on the binary treatment variable (*IV*) and all covariates simultaneously to obtain the CACE estimate (i.e., the association between the outcome and intervention assignment in the complier class). This association was fixed at zero in the non-complier class, in line with the exclusion restriction assumption. Where statistically significant CACE estimates were observed, we used Tymms' [36] formula to calculate effect sizes (Δ) akin to *d* (i.e., standardized mean difference). Predictors of compliance were established using a multinomial logistic regression. Effects of within (pupil) and between (school) level covariates are reported as logistic and linear regression coefficients, respectively.

School level clustering was accounted for using a sandwich estimator, using type = complex. Missing data on the outcome and compliance variables were handled using full information maximum likelihood (FIML). However, mirroring the approach taken in previous CACE analyses [35], individuals missing any covariate data or both outcome *and* compliance data were deleted listwise from the analysis, as

using FIML on these variables resulted in increased computational demands and model non-convergence.

Results

Descriptive statistics

Baseline characteristics based on measures used in this Article can be found in Table 2; extended baseline demographic information for the INSPIRE sample can be found in the main impact paper [18]. Descriptive statistics for the different dimensions of implementation (fidelity, reach, pupil responsiveness, dosage, feasibility, and acceptability) at the classroom level can be found in Table 3. Once dosage information was disaggregated to the pupil level (for use in the CACE analyses), in primary MBI schools, 1,115 pupils (60.5%) were moderate compliers (≥ 325 min), and 465 pupils (25.2%) were high compliers (≥ 393 min). In secondary MBI schools, 788 pupils (50.2%) were moderate compliers (≥ 207 min) and 367 pupils (23.4%) were high compliers (≥ 285 min). In primary RBPI schools, 1,183 pupils (57.9%) were moderate compliers (≥ 320 min) and 534 pupils were high compliers (26.1%) (≥ 376 min). Finally, in secondary RBPI schools, 573 pupils (57.4%) were moderate compliers (≥ 180 min) and 226 pupils (22.6%) were high compliers (≥ 275 min). Note that the number of compliers reported here is descriptive and based on the reported dosage information for intervention schools. The number of compliers reported hereafter is based on the model-estimated figures and includes would-be compliers in control schools.

MBI primary

Findings are reported in Table 4. Data for N = 3,630 pupils from 65 MBI and control schools were analysed, of whom n = 2,095 (57.7%) and n = 1,076 (29.6%) were estimated by the CACE model to be moderate and high compliers, respectively. Moderate compliance (≥ 325 min) led to a significant *increase* in SMFQ scores (β = 1.72, CIs 1.39 to 2.04; Δ = 0.32, CIs 0.20 to 0.44). High compliance (≥ 393 min) led to a further increase in SMFQ scores (β = 1.80, CIs 1.47 to 2.13; Δ = 0.48, CIs 0.28 to 0.69). These findings contrast with no main intervention effect (ITT estimate: effect size = -0.00, CIs − 0.08 to 0.08) [18] and indicates that when delivered at moderate or high dosages, the MBI had iatrogenic effects. None of our predictor variables influenced the likelihood of moderate or high compliance.

MBI secondary

Findings are reported in Table 4. Data from N = 3,708 pupils from 29 MBI and control schools were analysed, of whom n = 1,109 (29.9%) and n = 455 (12.3%) were estimated by the

Table 2 Sample characteristics

		Primary Schools			Secondary Schools		
		MBI[a]	RBPI [b]	Control	MBI[a]	RBPI [b]	Control
Pupil-level factors							
Pupils	*N*	2523	2554	2,088	2529	2391	2,988
Sex	Female *n* (%)	1256 (49.9)	1235 (48.4)	996 (47.7)	1330 (52.6)	1029 (43.1)	1,600 (53.6)
	Male *n* (%)	1262 (50.1)	1317 (51.6)	1,091 (52.3)	1198 (47.4)	1361 (56.9)	1,387 (46.4)
Behavioural Difficulties (scale range: 0–12)	T1 Mean (SD)	2.88 (2.11)	3.14 (2.24)	3.19 (2.30)	3.24 (2.30)	3.53 (2.53)	3.21 (2.22)
Depressive Symptoms (scale range 0–26)	T1 Mean (SD)	5.92 (4.68)	6.31 (4.89)	6.62 (4.93)	6.17 (5.66)	6.36 (5.74)	6.11 (5.33)
	T2 Mean (SD)	5.84 (5.05)	5.91 (5.07)	6.14 (5.33)	6.67 (6.07)	6.57 (5.97)	6.83 (6.10)
School-level factors							
Schools	*N*	37	36	36	16	16	18
Wave	1 *n* (%)	37 (100)	36 (100)	36 (100)	10 (62.5)	10 (62.5)	11 (61.1)
	3 *n* (%)	0 (0)	0 (0)	0 (0)	6 (37.5)	6 (37.5)	7 (38.9)
Top tertile percentage free school meal eligibility	*n* (%)	12 (32.4)	12 (33.3)	13 (36.1)	5 (31.2)	6 (37.5)	5 (27.8)
Staff Training	% (SD)	9.98 (6.70)	7.76 (7.82)	10.0 (8.48)	14.9 (7.78)	12.9 (5.92)	17.7 (11.1)
Prior Interventions	Yes *n* (%)	16 (43.2)	15 (41.7)	15 (41.7)	9 (56.2)	8 (50)	9 (50.0)

[a] Mindfulness-Based Intervention

[b] Relaxation-Based Practice Intervention

[*] *Note: descriptive data reflects the sample proportions for pupils allocated to each trial arm prior to model estimation*

Table 3 Descriptive statistics for implementation dimensions at the classroom level

	MBI [a]		RBPI [b]	
	Primary School	Secondary School	Primary School	Secondary School
Fidelity[c] %(SD)	89% (±16)	82.6% (±19.8)	89.2% (±14.6)	90% (±13)
Reach[d] % (SD)	97.60%	91.20%	98.90%	80.80%
Pupil Responsiveness[e] Mean (SD)	3.85(±0.65)	3.26 (±0.77)	4.03 (±0.63)	3.47 (±0.79)
Feasibility[f] Mean (SD)	4.16 (±0.72)	3.88 (±0.74)	4.32 (±0.61)	3.82(±0.65)
Acceptability[f] Mean (SD)	4.36 (±0.77)	4.14 (±0.78)	4.5 (±0.59)	4.36 (±0.64)
Appropriateness[f] Mean (SD)	4.24 (±0.78)	4.08 (±0.73)	4.35 (±0.66)	4.03 (±0.64)
Dosage[g] Mean (SD)	359 (±170)	234 (±172)	329 (±147)	246 (±153)

Note. Data for all dimensions of implementation were collected from the facilitating member of staff via surveys.

[a] *Mindfulness-Based Intervention*

[b] *Relaxation-Based Practice Intervention*

[c] *The extent to which teachers followed the guidance provided in the intervention manual (recorded as a percentage)*

[d] *The proportion of classes where teachers reported 80% of pupils were present for session delivery*

[e] *The mean score across five 5-point likert-type items (completely disagree (1) to completely agree (5)) measuring the extent to which pupils actively participated, engaged with, were interested in, enthusiastic about, and enjoyed the intervention (scored 1–5)*

[f] *Each measured as the mean of four 5-point likert-type items scored completely disagree (1) to completely agree (5)*

[g] *Mean minutes received*

CACE model to be moderate and high compliers, respectively. Moderate compliance (≥ 207 min) led to a significant reduction in SMFQ scores (β = -1.98, CIs − 2.48 to -1.47; Δ = -1.15, CIs − 1.57 to -0.74). High compliance (≥ 285 min) led to a further reduction in SMFQ scores (β = -1.97, CIs − 2.48 to -1.46; Δ = -1.40, CIs − 1.73 to -1.07). This indicates that receiving a greater dose of MBI increases its impact on depressive symptoms. Findings contrast with no main intervention effect (ITT estimate: effect size = -0.07, CIs − 0.18 to 0.05) [18] and indicate that the MBI can lead to reduced depressive symptoms among secondary school pupils, contingent upon the delivery of a minimally effective dose. Pupils in schools that did not have prior mental health interventions were more likely to be moderate compliers (see Table 4). However, no predictor variables significantly influenced the likelihood of high compliance.

RBPI primary

Findings are reported in Table 5. Data from $N = 3{,}800$ pupils from 66 RBPI and control schools were analysed, of whom n = 1,365 (35.9%) and n = 695 (18.3%) were estimated by the CACE model to be moderate and high compliers, respectively. Moderate compliance (≥ 320 min) led to a significant reduction in SMFQ scores (β = -1.98, CIs − 2.32 to -1.65; Δ = -1.52, CIs − 1.90 to -1.13). High compliance (receiving ≥ 376 min) also led to a significant reduction in SMFQ scores (β = -1.93, CIs − 2.26 to -1.60; Δ = -1.48, CIs − 1.83 to -1.14) but the similar effect size estimate to that reported for moderate compliance indicates no additional benefit accrued from the increased dosage. Findings contrast with no main intervention effect (ITT estimate: effect size = -0.01, CIs − 0.10 to 0.07) [18] and indicate that the RBPI can lead to reduced depressive symptoms among primary school pupils, contingent upon the delivery of a minimally effective dose. None of our included variables significantly influenced the likelihood of moderate compliance, however, lower levels of baseline behavioural difficulties made high compliance more likely (see Table 5).

RBPI secondary

Findings are reported in Table 5. Data from $N = 3{,}252$ pupils from 26 RPBI and control schools were analysed, of whom n = 2,261 (69.5%) and n = 1,565 (48.1%) were estimated by the CACE model to be moderate and high compliers, respectively. Moderate compliance (≥ 180 min) led to a significant increase in SMFQ scores (β = 1.70, CIs 1.19 to 2.20; Δ = 0.67, CIs 0.24 to 1.03). High compliance (≥ 275 min) led to a further increase in SMFQ scores (β = 1.74, CIs 1.33 to 2.15; Δ = 0.83, CIs 0.48 to 1.18). This indicates that receiving a greater dose of RPBI increases its iatrogenic impact on depressive symptoms. These findings contrast with no main intervention effect (ITT estimate: effect size = -0.10, CIs − 0.24 to 0.02 [18]. Higher levels of baseline depressive symptoms made moderate compliance more likely (see Table 5). Higher levels of baseline depressive symptoms also made high compliance more likely, as did offering fewer staff training opportunities (see Table 5).

Discussion

In this Article we reported results of CACE analyses of a mindfulness-based intervention (MBI) and relaxation practice-based intervention (RPBI) undertaken as part of the INSPIRE trial, representing the first robust estimates of their effects that account for implementation variability. In contrast to null ITT effect estimates [18], compliance effects

Table 4 CACE analyses for moderate and high compliers in primary and secondary schools for MBI (mindfulness-based intervention)

	SMFQ (b (SE))							
	Primary Schools (N=3,630)				Secondary Schools (N=3,708)			
	Moderate Compliance		High Compliance		Moderate Compliance		High Compliance	
	Non-Compliers	Compliers	Non-compliers	Compliers	Non-compliers	Compliers	Non-compliers	Compliers
Predictor	(42.3%)	(57.7%)	(70.4%)	(29.6%)	(70.1%)	(29.9%)	(87.7%)	(12.3%)
Pupil level predictors								
Sex (0=Male; 1=Female)	−0.032 (0.049)	0.051 (0.046)	−0.055 (0.033)	0.152 (0.058)**	0.234 (0.033)***	0.285 (0.080)***	0.244 (0.037)***	0.265 (0.096)**
Baseline emotional difficulties	0.470 (0.032)***	0.510 (0.028)***	0.492 (0.024)***	0.504 (0.036)***	0.622 (0.017)***	0.525 (0.028)***	0.606 (0.018)***	0.496 (0.050)***
School level predictors								
MBI vs. Control (CACE)	**-**	**1.715 (0.164)*** **	**-**	**1.80 (0.168)*** **	**-**	**-1.975 (0.257)*** **	**-**	**-1.971 (0.259)*** **
Wave (0=Wave 1; 1=Wave 3)	-	-	-	-	−0.297 (0.517)	−0.039 (0.104)	0.006 (0.613)	0.053 (0.122)
Free School Meals (0=bottom/middle tertiles; 1=top tertile)	0.777 (0.551)	−0.337 (0.228)	0.419 (0.558)	0.015 (0.309)	−0.311 (0.419)	−0.082 (0.088)	0.007 (0.523)	−0.099 (0.120)
Prior Interventions (0=No; 1=Yes)	1.533 (0.401)***	0.408 (0.261)	1.42 (0.306)***	0.320(0.363)	0.198 (0.356)	0.116 (0.094)	0.080 (0.398)	−0.027 (0.091)
Staff Training	0.217 (0.310)	0.045 (0.139)	0.214 (0.218)	−0.126 (0.135)	−0.077 (0.252)	0.044 (0.052)	−0.090 (0.233)	0.031 (0.067)
Entropy	0.666		0.691		0.804		0.899	
Predictors of compliance[a]					Prior Interventions: b = -1.122 (0.561)*			

CACE effects are marked in bold

[a]*Statistically significant predictors of compliance for each model*

*$p < .05$, **$p < .01$, ***$p < .001$

OR, *odds ratio with 95% confidence intervals*

were observed in each CACE analysis, with increased minutes of delivery leading to significant changes in depressive symptoms. Furthermore, high (as opposed to moderate) compliance led to further treatment effect amplification in all cases except RPBI in primary schools, where no additional effects were observed at high compliance. However, depressive symptoms did not change as expected in all cases, yielding an *increase* in depressive symptoms for MBI when delivered in primary schools, and for RPBI when delivered in secondary schools These effects were amplified further in the context of high compliance.

The MBI yielded null effects in the INSPIRE ITT analysis, consistent with another recent large trial of a universal school-based mindfulness intervention in the UK (MYRIAD) [37]. Our CACE analyses in *primary schools* found iatrogenic effects (i.e., when delivered at moderate or high levels of dosage, the MBI was found to *increase* depressive symptoms). This aligns with findings from an instrumental variable analysis undertaken in the MYRIAD trial, in which higher dose was associated with detrimental effects at post-intervention follow-up (though MYRIAD involved only *secondary* schools) [38]. Where the two trials diverge is that our CACE analyses revealed that increased MBI dosage (whether moderate or high compliance) in *secondary schools* led to *reduced* depressive symptoms. This may be the result of differences between the two MBIs and the resulting treatment effect moderation analyses. The MYRIAD MBI involved teacher-led delivery of 10 manualised, structured mindfulness lessons of 30–50 min duration, with pupils encouraged to engage in mindfulness practices

Table 5 CACE analyses for moderate and high compliers in primary and secondary schools for RBPI (relaxation-based practice intervention)

	SMFQ (b (SE))							
	Primary Schools (N=3,800)				Secondary Schools (N=3,252)			
	Moderate Compliance		High Compliance		Moderate Compliance		High Compliance	
	Non-compliers	Compliers	Non-compliers	Compliers	Non- compliers	Compliers	Non-compliers	Compliers
Predictor	(64.1%)	(35.9%)	(81.7%)	(18.3%)	(30.5%)	(69.5%)	(51.9%)	(48.1%)
Pupil level predictors								
Sex (0=Male; 1=Female)	0.059 (0.041)	−0.003 (0.057)	0.042 (0.039)	0.002 (0.117)	0.313 (0.061)***	0.367 (0.055)***	0.360 (0.046)***	0.277 (0.081)**
Baseline emotional difficulties	0.534 (0.017)***	0.544 (0.026)***	0.535 (0.017)***	0.528 (0.029)***	0.452 (0.087)***	0.598 (0.039)|***	0.480 (0.057)***	0.598 (0.040)***
School level predictors								
RBPI vs. Control (CACE)	**-**	**-1.983 (0.170)*****	**-**	**-1.932 (0.168)*****	**-**	**1.698 (0.258)*****	**-**	**1.739 (0.207)*****
Wave (0=Wave 1; 1=Wave 3)	-	-	-	-	0.052 (0.908)	0.407 (0.451)	−0.310 (0.524)	0.255 (0.267)
Free School Meals (0=bottom/middle tertiles; 1=top tertile)	0.915 (0.352)**	0.131 (0.083)	0.749 (0.349)*	0.173 (0.156)	−0.456 (0.871)	−0.132 (0.609)	−0.389 (0.404)	0.158 (0.225)
Prior Interventions (0=No; 1=Yes)	−0.674 (0.359)	0.083 (0.085)	−0.538 (0.345)	0.289 (0.116)*	−0.995 (1.334)	−0.621 (0.792)	−0.813 (0.554)	−0.187 (0.190)
Staff Training	−0.038 (0.165)	−0.008 (0.038)	0.034 (0.173)	0.074 (0.054)	0.731 (0.248)**	0.182 (0.130)	0.715 (0.146)***	0.161 (0.104)
Entropy	0.909		0.915			0.587		0.618
Predictors of compliance[a]			Baseline behavioural difficulties: b=−0.096 (0.044)*; OR = 0.908 (0.833 to 0.990)			Baseline emotional difficulties: b = 0.974 (0.301)**; OR=2.648 (1.467 to 4.782)		Baseline emotional difficulties: b=1.054 (0.365)**; OR=2.870 (1.403 to 5.870) Staff training opportunities: b = -1.514 (0.337)***

CACE effects are marked in bold

[a]*Statistically significant predictors of compliance for each model*

*$p < .05$, **$p < .01$, ***$p < .001$

OR, *odds ratio with 95% confidence intervals*

thereafter [37]. Dosage was measured in terms of lesson exposure and was very high (i.e., mean 9/10), but reported frequency of subsequent pupil engagement in mindfulness practices was very low (i.e., mean score of 1/5) [38]. By contrast, the MBI reported in this Article involved very short (5 min) but frequent (daily) teacher-led mindfulness exercises. It was rated as being both feasible and acceptable by teachers (mean score ~ 4/5 for both dimensions of social validity across both primary and secondary schools), and likely promoted greater frequency of pupil engagement in mindfulness practices, since these were built into the daily delivery, as opposed to being encouraged post-delivery.

Collectively, findings across INSPIRE and MYRIAD may reflect the underpinning processes through which MBIs influence mental health outcomes. In theory, MBIs raise participants' awareness of their thoughts and feelings, including those that are challenging or unpleasant, and teach key adaptive skills (e.g., self-regulation) that support mental

health [37]. It is possible that for younger age groups (as in MBI primary), who may lack the metacognitive skills necessary for MBIs to be fully effective [30], or situations in which engagement in post-intervention mindfulness practices is relatively infrequent (as in MYRIAD), an increased awareness of challenging or unpleasant feelings may inadvertently encourage rumination and subsequently depressive symptoms [39]. It may therefore be important for effective self-regulation to be learned, practiced, and consolidated, such that the negative cycle between ineffective self-regulation and the increase in unwanted feelings that underpin the development of emotional difficulties is disrupted before it becomes habitual [40].

In the RBPI, the reduction of depressive symptoms in primary school pupils was found to be contingent upon delivery of a minimally effective dose. These findings align with those of a recent meta-analysis, in which density of RBPI dosage explained 44% of the variation in overall treatment effect size for anxiety outcomes, although the focus was on adolescence [17]. The notion of a minimally effective dose of the RBPI also fits with underpinning theory; they are hypothesized to improve mental health by replacing default psychological and physiological stress responses (e.g., worry, increased heart rate) with diametrically opposed relaxation responses (e.g., calm, reduced heart rate) [41], and higher levels of repeated exposure are likely required for reinforcement, consolidation, and generalisation of learned behaviours to trigger such change. However, in secondary schools, RBPI *increased* depressive symptoms, with higher compliance exacerbating this effect, mirroring the outcomes of MBI delivery in primary schools. This may reflect a paradoxical phenomenon known as 'Relaxation Induced Anxiety', which occurs when practising relaxation techniques results in an *increase* in psychological and physiological stress responses, and is thought to be linked to increases in rumination, awareness of one's internal state, or fear of losing control when fully relaxed [42]. While typically associated with Generalised Anxiety Disorder [42], subgroup analyses of the current RBPI in secondary schools suggests fewer emotional symptoms among those with higher baseline depressive symptoms [18].

The fact that we did observe a reduction in depressive symptoms among primary school pupils when RBPI was delivered at a minimally effective dose indicates possible developmental differences in responsiveness to relaxation practices. Adolescence, a key life stage characterised by rapid biological changes and cognitive, social and emotional growth, presents unique challenges for intervention. Critics argue that universal interventions inadequately cater to the diverse needs of this age group, who may perceive such activities as irrelevant or a lost opportunity to engage in other preferred and more intrinsically rewarding activities [43]. The amplification of adverse effects at higher levels of compliance supports this perspective. Conversely, a reverse pattern was observed for MBI (i.e., unwanted effects in *primary* school and reduced symptoms in *secondary* school). This indicates that MBIs may be more appropriate for older pupils with more enhanced metacognitive skills, while RBPIs may be better suited to younger pupils, but could become detrimental as they age.

In summary, our analyses indicated that the administration of a minimally effective dose may be critical to unlocking the effects of the universal, school-based mindfulness and relaxation interventions. This was consistent across both interventions and across primary and secondary school settings, although observed dosage effects were not always positive. Null and iatrogenic effects of universal school-based mental health interventions are increasingly common among trials with robust methodology and large sample sizes [44]. That the MBI and RBPI reported in this Article yielded null effects on depressive symptoms when analysed according to the ITT principle [18], and apparent iatrogenic dosage effects in the case of MBI primary and RBPI secondary CACE analyses, suggests that these interventions may be unlikely to reduce depressive symptoms when delivered as universal school based interventions in some age groups or phases of education: in these cases, our findings indicate no evidence of benefit for depressive symptoms and possible adverse effects among participants with greater intervention exposure, although the latter findings should be interpreted in light of the assumptions underlying CACE analyses and do not provide definitive evidence of harm. While broader indicators of functioning represent an important avenue for future work, the present findings speak specifically to intervention effects on depressive symptoms, consistent with the original aims of the INSPIRE trial. Relatedly, the magnitude of several standardised CACE estimates was larger than is typically observed in universal school-based mental health intervention trials. These estimates should be interpreted considering the CACE estimand, which reflects effects among estimated compliers rather than average effects of intervention allocation across the full trial sample.

Where positive effects are evident (RBPI primary and MBI secondary), a minimally effective dose is necessary. A key practical implication of the current study, therefore, relates to how schools can be encouraged and supported to ensure more regular, frequent delivery of sessions. Enhancing initial training (e.g., to incorporate theory and evidence relating to the importance of frequent and consistent session implementation), provision of ongoing implementation support (e.g., external coaches to monitor, model and supports' teacher delivery or provide booster training sessions), and careful consideration for how the underlying pathways can result in symptom deterioration, are three immediate

avenues for exploration in the next iteration of these interventions. Qualitative work is a valuable avenue by which detrimental effects can begin to be explored. For example, focus groups undertaken with a subset of schools as part of the Education for Wellbeing trial indicated feelings of boredom and frustration with MBI and RBPI materials (particularly their repetitive nature) and finding activities unhelpful were among reported negative perceptions about the two interventions reported by some pupils, although these analyses did not discriminate between primary and secondary schools [45].

Our CACE analyses are the key strength of this study, since they took implementation variability into account, and in doing so facilitated estimation of intervention effects in the context of compliance, while retaining the bias protective principle of randomisation [23]. Furthermore, the use of two compliance thresholds enabled us to consider whether treatment effects were amplified at higher levels of cumulative intervention intensity. However, these thresholds were arbitrary, and further work is needed to model the dose-response relationship of both interventions in a more nuanced way.

Although examining multiple dimensions of implementation using a multiverse approach would provide a more comprehensive picture of how implementation influences intervention effects [46], the present study included only a single measure of compliance, operationalised as intervention dosage. This reflects constraints inherent in CACE estimation and aligns with common practice in the literature [46]. While, in principle, separate CACE models could be estimated for different aspects of implementation (e.g., quality or fidelity), the analyses were restricted due to the already substantial number of models estimated (two interventions; primary and secondary school samples; moderate and high compliance thresholds). In addition, dosage (total minutes delivered) was derived from teacher self-report data. Implementation surveys were completed retrospectively by delivering staff, meaning that reports may have been subject to recall bias. Teacher self-report of implementation may also be influenced by impression management or variability in interpretations of implementation requirements [47]. The absence of observational or independent measures of implementation quality should therefore be considered when interpreting these findings, and future research should, where possible and feasible, incorporate multiple aspects of implementation.

We also note that, as is the case with all CACE analyses, only a single aspect of implementation was considered, though dosage is arguably the most important as it is the most variable in school-based interventions (compared to others, such as quality) [48]. Some of the CACE analyses reported here yielded no statistically significant predictors of compliance or entropy < 0.60 (i.e., below the recommended threshold for class separation) [49]. Although findings from these CACE models are still robust, these results should be interpreted with a degree of caution when compared to those with significant compliance predictors/entropy > 0.60. Finally, our analyses comprised secondary analysis of a larger RCT, which did not prospectively include measures specifically intended to predict compliance. Consequently, we identified plausible predictors post-hoc from the measures already available in the dataset, selecting variables that, based on theory and prior evidence, could reasonably be expected to relate to participants' likelihood of engaging with the intervention. Future trials intending to conduct CACE analyses should identify relevant predictors a priori and ensure that these are deliberately measured during data collection.

Conclusions

This Article highlights the importance of intervention implementation as a treatment effect modifier of universal school-based mindfulness and relaxation interventions. In doing so, it has provided important new insights into the effects that can be expected in the context of moderate and high compliance. Going forward, enhanced initial training (including an understanding of how increased dosage of MBI and RBPI may have detrimental effects in certain phases of education) and provision of ongoing implementation support may be required in order for positive effects to be achieved consistently.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00787-026-03152-x.

Acknowledgements We would like to thank all schools, pupils and parents who engaged with or participated in the programme. We would also like to thank the advisory group members, members of the data monitoring and ethics committee and members of the trial steering committee.

Author contributions JD was Principal Investigator and with support from NH and PS, conceptualised the programme of interventions known as Education for Wellbeing comprising the INSPIRE and AWARE trials. DH and AT were trial managers for Wave 1 and Wave 3, respectively. ET, supported by CK, JS and CM, led on decisions relating to data management. JS, EA, RM, AM, DH, AT, CK, ET and CM contributed to data acquisition, cleaning and related duties. ET, supported by CK and MP, led on data analysis. ET, CK and NH interpreted the results and drafted the manuscript. ET, CK, NH, DH, AT and JD revised the manuscript. All authors read and approved the final manuscript.

Funding This research was commissioned and funded by the UK Department for Education (DfE). The commissioner and funder of the study had no role in data collection, data analysis, data interpretation

or writing of the report. DfE selected the interventions to be trialled and chaired an advisory group of researchers and educators regarding the progress and quality of the research. Views expressed herein are those of the authors and not necessarily the views of the DfE.

Data availability An anonymised dataset will be made publicly available in the future (planned late 2026). Please contact the Principal Investigator or Trials Manager for further information.

Declarations

Ethics approval and consent to participate Ethics approval for this study was granted from University College London (UCL) Research Ethics Committee (6735/009 and 6735/014). Storage and processing of all research data aligned with General Data Protection Regulation legislation. Eligible schools that expressed interest in taking part, were required to have a member of their senior leadership team sign a Memorandum of Understanding, and data-sharing agreement. Informed opt-in assent and opt-out consent was obtained in writing from participating children and their parent/legal guardian prior to baseline data collection.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

References

1. NHS Digital (2023) Mental health of children and young people in England, 2023 - wave 4 follow up to the 2017 survey
2. Bor W et al (2014) Are child and adolescent mental health problems increasing in the 21st century? A systematic review. Aust N Z J Psychiatry 48(7):606–616
3. Patalay P, Gage SH (2019) Changes in millennial adolescent mental health and health-related behaviours over 10 years: a population cohort comparison study. Int J Epidemiol 48(5):1650–1664
4. Humphrey N (2023) Universal, school-based social and emotional learning interventions and their potential to improve students' mental health. In: Evans SW et al (eds) Handbook of school mental health. Springer, New York, pp 9–19
5. NHS Digital (2023) Mental health of children and young people in England. London
6. Foxcroft DR (2014) Can Prevention Classification be Improved by Considering the Function of Prevention? Prev science: official J Soc Prev Res 15(6):818–822
7. Deighton J et al (2025) Effectiveness of school mental health and wellbeing promotion Universal approaches in English primary and secondary schools
8. Rose G (2001) Sick individuals and sick populations. Int J Epidemiol 30(3):427–432
9. Hayes D et al (2025) Universal, school-based, interventions to improve emotional outcomes in children and young people: a systematic review and meta-analysis. Front Child Adolesc Psychiatry 4:1526840
10. Werner-Seidler A et al (2017) School-based depression and anxiety prevention programs for young people: A systematic review and meta-analysis. Clin Psychol Rev 51:30–47
11. Zhang Q, Wang J, Neitzel A (2023) School-based Mental Health Interventions Targeting Depression or Anxiety: A Meta-analysis of Rigorous Randomized Controlled Trials for School-aged Children and Adolescents. J Youth Adolesc 52(1):195–217
12. NatCen and National Children's Bureau (2017) Supporting mental health in schools and colleges. : London
13. Department for Education (2021) Statutory guidance on relationships education, relationships and sex education (RSE) and health education: London
14. Nisar A et al (2024) Mainstream and special schools' use of wellbeing programmes: A regional survey. Br Edu Res J 50(1):6–31
15. Department for Education (2017) Analytical associate pool: summary of projects
16. Dunning D et al (2022) Do mindfulness-based programmes improve the cognitive skills, behaviour and mental health of children and adolescents? An updated meta-analysis of randomised controlled trials. (1468-960X (Electronic)).
17. Hamdani SU et al (2022) Effectiveness of relaxation techniques 'as an active ingredient of psychological interventions' to reduce distress, anxiety and depression in adolescents: a systematic review and meta-analysis. Int J Mental Health Syst 16:1752–4458 (Print))
18. Deighton J et al (n.d.) Promoting mental health and wellbeing in schools: examining mindfulness-based exercises, relaxation practices and Strategies for Safety and Wellbeing in English primary and secondary schools (INSPIRE): a multi-school, cluster randomised controlled trial [Preprint]
19. Durlak JA (2015) Studying program implementation is not easy but it is essential. 16: pp. 1123–1127
20. Humphrey N et al (2021) Treatment Effect Modifiers in a Randomized Trial of the Good Behavior Game During Middle Childhood. J Consult Clin Psychol 89(8):668–681
21. Cipriano C et al (2023) The state of evidence for social and emotional learning: A contemporary meta-analysis of universal school-based SEL interventions. Child Dev 94(5):1181–1204
22. Jo B et al (2008) Cluster randomized trials with treatment noncompliance. Psychol Methods 13(1):1–18
23. Peugh JL et al (2017) Beyond intent to treat (ITT): A complier average causal effect (CACE) estimation primer. J Sch Psychol 60:7–24
24. Hayes D et al (2019) Promoting mental health and wellbeing in schools: examining Mindfulness, Relaxation and Strategies for Safety and Wellbeing in English primary and secondary schools: study protocol for a multi-school, cluster randomised controlled trial (INSPIRE). Trials 20(1):640
25. Knowles C, Thornton E, Panayiotou M, Hayes D, Thompson A, Deighton J, Santos J, Mason C, Ashworth E, Mansfield R, March A, Stallard P, Humphrey N (2026) Complete delivery matters: Multilevel complier average causal effect analysis of universal mental health curricula in two randomized controlled trials in U.K. schools. J Consult Clin Psychol 94(5):267–280. https://doi.org/10.1037/ccp0000997
26. Education for Wellbeing: Technical Report. 2025 https://assets.publishing.service.gov.uk/media/67a3325c567402152f553c54/Education_for_wellbeing_-_technical_report.pdf
27. Panayiotou M, Humphrey N, Hennessey A (2020) Implementation matters: Using complier average causal effect estimation

to determine the impact of the Promoting Alternative Thinking Strategies (PATHS) curriculum on children's quality of life. J Educ Psychol 112(2):236–253
28. Deighton J et al (2013) The Development of a School-Based Measure of Child Mental Health. J Psychoeduc Assess 31(3):247–257
29. Warren SF, Fey ME, Yoder PJ (2007) Differential treatment intensity research: a missing link to creating optimally effective communication interventions. Ment Retard Dev Disabil Res Rev 13:313–320
30. Angold A et al (1995) Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. US, John Wiley & Sons, pp 237–249
31. Berg JK et al (2017) Using Complier Average Causal Effect Estimation to Determine the Impacts of the Good Behavior Game Preventive Intervention on Teacher Implementers. Adm Policy Mental Health Mental Health Serv Res 44(4):558–571
32. Muthén LK, Muthén BO Mplus User's Guide.. Eighth Edition. ed. 1998–2017, Los Angeles, CA: Muthén & Muthén
33. Hallquist MN, Wiley JF (2018) MplusAutomation: An R Package for Facilitating Large-Scale Latent Variable Analyses in Mplus. Struct Equ Model 25(4):621–638
34. Angrist JD, Imbens GW, Rubin DB (1996) Identification of Causal Effects Using Instrumental Variables. J Am Stat Assoc 91(434):444–444
35. Humphrey N, Panayiotou M (2022) Bounce Back: randomised trial of a brief, school-based group intervention for children with emergent mental health difficulties. Eur Child Adolesc Psychiatry 31(1):205–210
36. Tymms P (2004) Effect sizes in multilevel models. In: Schagen I, Elliot K (eds) But what does it mean? The use of effect sizes in educational research. National Foundation for Educational Research, pp 55–66 https://www.nfer.ac.uk/media/4tkbnw2w/but_what_does_it_mean_the_use_of_effect_sizes_in_educational_research.pdf
37. Kuyken WA-O et al (2022) Effectiveness and cost-effectiveness of universal school-based mindfulness training compared with normal school provision in reducing risk of mental health problems and promoting well-being in adolescence: the MYRIAD cluster randomised controlled trial. BMJ Evidence-Based Mental Health. (1468-960X (Electronic)).
38. Montero-Marin J et al (2022) School-based mindfulness training in early adolescence: what works, for whom and how in the MYRIAD trial? Evid Based Mental Health 25(3):117
39. Foulkes L, Stringaris A (2023) Do no harm: can school mental health interventions cause iatrogenic harm? BJPsych Bull 47(5):267–269
40. Barlow DH, Allen LB, Choate ML (2016) Toward a Unified Treatment for Emotional Disorders – Republished Article. Behav Ther 47(6):838–853
41. National Center for Complementary and Integrative Health (2021) Relaxation Techniques: What You Need To Know
42. Kim H, Newman MG (2019) The paradox of relaxation training: Relaxation induced anxiety and mediation effects of negative contrast sensitivity in generalized anxiety disorder and major depressive disorder. J Affect Disord 259:271–278
43. Andrews JL, Foulkes L (2024) Debate: Where to next for universal school-based mental health interventions? Time to move towards more effective alternatives. Child Adolesc Ment Health
44. Guzman Holst C et al (2024) Scoping review: Potential harm from school-based group mental health intervention.
45. Stapley E et al (2025) Pupil perspectives on Mindfulness and Relaxation: A qualitative study of English school children's experiences of two brief, universal, classroom-based mental health and wellbeing interventions: Mindfulness and Relaxation. [Preprint]
46. O'Brien A et al (2025) A systematic review informing recommendations for assessing implementation variability in universal, school-based social and emotional learning interventions. Social Emotional Learning: Res Pract Policy 5:100112
47. Humphrey N et al (2016) Implementation and process evaluation (IPE) for interventions in educational settings: a synthesis of the literature. : London
48. HennesseyA, Humphrey N (2020) Can social and emotional learning improve children's academic progress? Findings from a randomised controlled trial of the Promoting Alternative Thinking Strategies (PATHS) curriculum. Eur J Psychol Educ 35:751–774. https://doi.org/10.1007/s10212-019-00452-6
49. Bauer J et al (2022) A Primer to Latent Profile and Latent Class Analysis, in Methods for Researching Professional Learning and Development: Challenges, Applications and Empirical Illustrations, M. Goller, Editors. Springer International Publishing: Cham. pp. 243–268