

Complete Delivery Matters: Multilevel Complier Average Causal Effect Analysis of Universal Mental Health Curricula in two Randomised Controlled Trials in UK Schools

Christopher Knowles¹, Emma Thornton¹, Margarita Panayiotou¹, Daniel Hayes^{2,3}, Abigail Thompson², Jessica Deighton², Joao Santos¹, Carla Mason¹, Emma Ashworth⁴, Rosie Mansfield⁵, Anna March², Paul Stallard⁶, and Neil Humphrey¹

¹ Manchester Institute of Education, University of Manchester

² Evidence Based Practice Unit, University College London and Anna Freud Centre

³ Department of Behavioural Science and Health, University College London

⁴ School of Psychology, Liverpool John Moores University

⁵ Centre for Longitudinal Studies, University College London

⁶ Department of Health, University of Bath

Corresponding author: Christopher Knowles (christopher.knowles@manchester.ac.uk)

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.

Availability of data and materials

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

Competing interests

The authors declare that they have no competing interests.

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.

Authors' 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. CK, supported by ET and MP, led on data analysis. CK, ET and NH interpreted the results and drafted the manuscript. CK, ET, NH, DH, AT and JD revised the manuscript. All authors read and approved the final manuscript.

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.

Complete Delivery Matters: Multilevel Complier Average Causal Effect Analysis of Universal Mental Health Curricula in two Randomised Controlled Trials in UK Schools

Abstract

Objective: Complete implementation is assumed in intention-to-treat (ITT) analyses of universal, school-based mental health literacy curricula, but implementation data routinely reveal programme delivery is often incomplete (or in extreme cases, no implementation occurs). Complier average causal effect (CACE) estimation can rigorously determine whether this makes a difference to intervention outcomes. **Method:** We used multilevel CACE estimation to establish if intervention compliance modified treatment effects for universal, school-based mental health literacy interventions delivered in two large UK cluster randomised controlled trials (INSPIRE and AWARE): Strategies for Safety and Wellbeing (SSW, $k = 64$ primary schools, $N = 3397$ pupils, 91% compliers; $k = 29$ secondary schools, $N = 3692$ pupils, 87% compliers), The Guide ($k = 84$ secondary schools, $N = 4879$ pupils, 71% compliers) (complete vs incomplete delivery: effects on help-seeking intentions); and Youth Aware of Mental Health (YAM, $k = 87$ secondary schools, $N = 5408$ pupils, 51% compliers) (complete delivery vs no delivery: effects on emotional difficulties). **Results:** A compliance effect was observed for YAM ($d = -0.97$), The Guide ($d = 0.17$) and SSW when implemented in secondary schools ($d = 0.24$) but not primary schools. **Conclusion:** This study provides important evidence that universal school-based mental health literacy curricula can serve as potentially efficacious programmes to reduce emotional difficulties and enhance help-seeking behaviour, but that effects in secondary schools are largely contingent on, or amplified by, complete session delivery. To maximise impact, research must prioritise development of effective strategies and systems of support to overcome commonly identified barriers to implementation.

Key words: *Complier Average Causal Effect, Help-Seeking, Internalising Symptoms, Mental Health Literacy, Randomised Controlled Trials*

Public health significance statements:

1. School-based mental health literacy programmes can positively impact pupils' emotional wellbeing and help-seeking behaviours, but benefits in secondary schools may be contingent on full implementation.
2. Findings underscore the need for robust implementation strategies in secondary schools; without complete session delivery, the potential public health impact of mental health curricula is often diminished.
3. By applying compliance-based causal methods, this study highlights that efforts to ensure complete implementation may be as crucial as programme content in achieving mental health outcomes at scale.

Approximately one in five children and young people aged 8-16 have a probable mental health disorder (NHS Digital, 2023). Schools are important contexts for the prevention and treatment of such difficulties. Universal mental health literacy (MHL) interventions focus on improving knowledge and understanding of mental health, mental disorders, and reducing stigma, and are increasingly common in schools. However, the current evidence base is equivocal regarding their impact on MHL (e.g., intentions to seek help) and/or mental health (e.g., emotional difficulties) outcomes (Mills et al., 2023; MA et al., 2023; Seedaket et al., 2020). Relationships, sex and health education (including mental health education and related provision) became compulsory in English schools in 2021 (Department for Education, 2021) creating an urgent need to identify efficacious programmes.

This Article focuses on three universal MHL interventions: Youth Aware of Mental Health (YAM; Wasserman et al., 2015); The Mental Health and High School Curriculum Guide (The Guide; Milin et al., 2016); and Strategies for Safety and Wellbeing (SSW; Hayes et al., 2019a), which were selected by England's Department for Education (DfE) on the basis of their potential to improve mental health literacy and/or mental health outcomes in English schools. Specifically, we examine the extent to which complete delivery of these curricula (i.e., compliance) modifies treatment effects on primary outcomes that were observed in intention-to-treat (ITT) analyses in the INSPIRE and AWARE cluster randomised controlled trials (RCTs; Hayes et al., 2019a, 2019b). Despite the numerous advantages of ITT analyses (e.g., bias protective principle of randomisation, preservation of sample size and statistical power; Gupta, 2011), they assume participants behave as expected (i.e., they receive the intervention in full in the treatment arm). Implementation data routinely reveal that in practice this is not the case, with delivery often being incomplete (or in extreme cases, not occurring at all) in school-based

interventions. This is for myriad reasons including time constraints, competing priorities, and/or staffs' negative perceptions of the intervention in question (Durlak, 2015). In instances where an intervention being trialled is efficacious, but there is substantial non-compliance, ITT analyses may therefore underestimate the potential magnitude of the intervention effect (i.e., that which would occur in the context of compliance; McCoy, 2017).

For years, educational psychology theorists have emphasised that ensuring intervention integrity in the quantitative analysis of randomised controlled trials is essential for assessing the feasibility and full potential of school-based prevention efforts (Dane & Schneider, 1998). However, robust methods for doing so are rarely utilised in practice, highlighting a critical gap in understanding that undermines the validity of many prevention studies. This gap has significant implications for educational psychology theory, including misinformed cost-effectiveness estimates and the premature discontinuation of potentially effective programmes due to misattribution of null effects to intervention design rather than shortcomings in implementation. To this end, the current study furthers theoretical understanding of the efficacy and applicability of universal school-based interventions in England by employing complier average causal effect (CACE) estimation.

CACE estimation offers a useful complement to ITT analyses as it enables a robust (and, arguably, more accurate) assessment of intervention effects that accounts for variability in implementation (Peugh et al., 2017). In cluster RCTs (with multilevel components), CACE analysis is typically undertaken through mixture models (latent class analysis), where two latent classes are estimated to identify compliers and non-compliers. Compliance status for those in the intervention arm is usually known, but equivalent data for control participants is missing (for obvious reasons) and must be probabilistically estimated. Item responses on plausible predictors

of compliance are compared between those in the treatment and control arms to identify those who would likely have complied with the intervention (or not), had they been randomised to receive it (Kitmitto et al., 2018). Thereafter, the main outcome of interest is compared between compliers in the treatment arm against ‘would-be’ compliers in the control arm of the trial.

Despite its potential utility, the application of CACE in RCTs of school-based interventions (including those focusing on mental health promotion and/or literacy) is not yet widespread (Peugh et al., 2017), but examples of where it has been used provide vital insights. First, in a scenario in which both ITT and CACE estimates are null (as in the recent RCT of a school-based science intervention, *Thinking, Doing, Talking Science*; Kitmitto et al., 2018), CACE helps to rule out implementation failure as the reason for the lack of a treatment effect. Second, where a null ITT estimate contrasts with a compliance effect in CACE estimation (as in a recent RCT of the *Good Behaviour Game*; Humphrey et al., 2021), this can give a robust indication of the likely magnitude of intervention effect that would occur in the context of compliance. Third, a treatment effect observed in ITT contrasting with a null or negative CACE effect can imply a ceiling effect or, in some cases, that high levels of compliance may cause iatrogenic effects (as observed in another RCT of the *Good Behaviour Game*; Berg et al., 2017). Finally, where ITT analysis yields a treatment effect, and this is amplified in CACE (as in a trial of the *Promoting Alternative Thinking Strategies (PATHS)* curriculum; Panayiotou et al., 2020), this indicates that the intervention is efficacious but can be even more potent in the context of high compliance.

The current study

The current study is the first to report the effects of a series of universal school-based interventions (YAM, The Guide, and SSW) designed to reduce emotional difficulties and improve intentions to seek help through enhanced mental health literacy that takes implementation variability into account. The primary aim was to establish intervention efficacy through multilevel CACE analysis, a highly illuminative yet underutilised statistical technique. We hypothesized that for each intervention, significantly improved primary outcomes would be observed, on average, for children and young people in intervention arm schools that complied (by delivering said intervention in full), when compared to ‘would-be’ compliers in control arm schools (identified using maximum likelihood estimation techniques).

Compliance was defined as complete session delivery (i.e., five sessions of YAM, six sessions of The Guide; eight sessions of SSW), with non-compliance defined as anything less than complete session delivery (i.e., less than six sessions of The Guide; less than eight sessions of SSW; or, in the case of YAM, no delivery) for both theoretical and empirical reasons. First, each session targeted a distinct aspect of mental health literacy (knowledge, stigma reduction, help-seeking, coping), so missing one could mean omitting an entire conceptual domain. Second, alongside the modular design of the interventions, a baseline level of learning around key concepts was assumed to be carried forward from one session to the next for pupils to fully understand content in subsequent modules (e.g., what it means to be ‘emotionally safe’). Third, and perhaps most importantly, implementation data (Figure 1) showed that schools treated the curricula as complete packages: when implemented, delivery was high, indicating schools largely delivered the full curriculum or not at all. This distinction reflects real differences in implementation capacity. While partial delivery can be meaningful in interventions with defined core components, the curricula studied here were explicitly designed (and in practice,

implemented) as complete programmes. Our operationalisation therefore aligns with both the theoretical intent and the empirical reality of these interventions. For this reason, we henceforth refer to intervention *completeness* rather than *compliance* throughout, reserving the latter term only for technical discussion of CACE.

<<INSERT FIGURE 1 HERE>>

Methods

Trial design and participants

Full details of the INSPIRE and AWARE RCTs are provided in their respective protocols and main impact papers (Deighton et al., 2025a, 2025b; Hayes et al., 2019a, 2019b;). The trials were delivered throughout England (United Kingdom) in three waves spanning 2018-2024 (though COVID-19 disruption and school closure led to Wave 2 being abandoned). SSW was delivered in both primary (Year groups 4 and 5) and secondary (Year groups 7 and 8) schools. YAM and The Guide were delivered in secondary (Year group 9) schools only (see Figures 2 and 3 for total number of schools allocated to each trial arm). Demographic characteristics of each sample can be found in Table 1. Extended demographic information for each sample, including pupils' ethnicity and the region where schools were located is also provided as Supplementary Table S1.

<<INSERT TABLE 1 HERE>>

Randomisation and masking

The flow of schools and pupils through INSPIRE and AWARE (excluding those recruited during the abandoned Wave 2) is illustrated in Figures 2 and 3, respectively. After standard data cleaning processes, the number of pupils allocated to each trial arm with data available were as follows: 1) INSPIRE: SSW Primary $n = 2566$; SSW Secondary $n = 2850$; Usual Practice Primary $n = 2088$; Usual Practice Secondary $n = 2988$; and, 2) AWARE: The Guide $n = 3997$; YAM $n = 4028$; Usual Practice $n = 4141$. While analysts were blinded during main outcome (ITT) assessment, CACE analysis requires knowledge of treatment assignment thus, blinding is not possible for this type of model estimation.

<<INSERT FIGURES 2 & 3 HERE>>

Procedures

SSW, YAM and The Guide were compared with a control group in which the usual school curriculum was delivered. All interventions were implemented after baseline assessment and randomisation. Follow-up assessment occurred after the implementation period. In response to the call for more comprehensive reporting of implementation strategies in intervention research (Rudd et al., 2020), the strategies employed for each intervention are described in Supplementary Table S2, following Proctor et al.'s (2013) taxonomy.

YAM, developed by Mental Health in Mind International AB, is a five-session weekly MHL intervention (60 minutes/session) that aims to improve mental health awareness via discussion of risk and protective factors. It provides young people aged 14–16 years with a safe space to explore topics such as depression, anxiety and suicidal thoughts. YAM also encourages reflection on approaches to problem-solving in challenging situations. Six main themes are

covered: 1) What is mental health?; 2) Self-help advice; 3) Stress and crisis; 4) Depression and suicidal thoughts; 5) Helping a friend in need; and 6) Who can I ask for advice? Sessions are delivered in the classroom by trained external instructors and helpers, who are professionals with a relevant professional background (e.g., education, psychology, youth work). Supporting learning materials include posters and booklets. Role-play sessions (comprising three themes: awareness about choices; depression and suicidal thoughts and feelings; and, how to manage stress and crisis situations) are central to YAM, giving young people the opportunity to explore and act out relevant issues from their daily lives in a safe and confidential environment (Wasserman et al., 2015; Hayes et al., 2019b).

The Guide, originally developed by Dr Stanley Kutcher and the Canadian Mental Health Association, consists of six weekly 60-minute sessions, aiming to increase mental health literacy in both young people and staff. Modules are as follows: 1) The stigma of mental illness; 2) Understanding the relationship between mental health and mental illness; 3) Understanding specific mental illnesses; 4) Adolescents' experiences of mental illness; 5) Seeking help and finding support; and 6) The importance of positive mental health. The Education for Wellbeing intervention development team worked with original developers to modify content to include more resources from England (the original version having been developed in Canada), and less emphasis on PowerPoint presentations in favour of interactive discussion. All modifications were informed by a pilot study and approved by original developers. The six sessions, delivered by trained school staff, provide coverage of a range of mental disorders (e.g., bipolar disorder, panic disorder, eating disorders, depression), support and where to get help, and stress. Homework activities to extend and consolidate learning are also included in the curriculum

(Milin et al., 2016; Hayes et al., 2019b). Staff received a lesson plan for each session along with a half-day face-to-face training session in late 2018.

Based on a protective behaviours framework but adapted to focus more on wellbeing rather than safeguarding (Dale et al., 2016), SSW aims to increase children and young people's knowledge and skills relating to their safety, mental health and wellbeing, and how to access sources of support. Trained school staff deliver eight weekly 40-minute sessions. Sessions cover the following topics: 1) It is safe to talk about mental health; 2) You are never too young to talk mental health (primary schools)/We all have mental health (secondary schools); 3) What is safety?; 4) Early warning signs – noticing our bodies; 5) Early warning signs – noticing our feelings and thoughts; 6) Developing our safety networks; 7) Safe friendships; and, 8) Safe ways of managing emotions (Hayes et al., 2019a). Staff implementing SSW were provided lesson plans suitably adapted for implementation in primary and secondary schools and invited to attend a half-day face-to-face training session.

Schools in the control group continued their usual mental health provision without restrictions. To assess existing provision, school staff completed a baseline survey measuring the extent of mental health training, universal provision (e.g., structured lessons on mental health topics), and targeted provision (e.g., counselling). Each domain was assessed using a matrix format, where respondents could select multiple applicable service providers. For each matrix, the number of selected options (scored as 1) was summed and divided by the total possible options, generating proportional scores for staff training, universal provision, and targeted provision. This data reduction method aligns with Mansfield, Humphrey, and Patalay (2021), who demonstrated that greater school-wide training provision was associated with increased

teacher-reported capacity to support student mental health. All control schools were offering some degree of staff training, universal provision and targeted provision at baseline (Table 1).

Compliance data were derived from a combination of surveys administered to implementing school staff (or, in the case of YAM, external instructors), who were asked to indicate which of the sessions noted above had been delivered, and correspondence from schools (e.g., those who indicated, post-randomisation, that they would not be able to deliver their allocated intervention). Complete implementation was defined in terms of dosage, reflecting the suitability of this dimension for CACE estimation. In the current study, dosage data were disaggregated from the classroom-level to the pupil-level. Detailed theoretical justification for doing is provided in Discussion, and in previous studies adopting this approach (Humphrey et al., 2021; Panayiotou et al., 2020). Further, to contextualise the quality of implementation and support the rationale for focusing on completeness, descriptive data for other available implementation dimensions are provided (Table 2), summarised and discussed herein.

Additional dimensions for which data were available were: *fidelity* (the extent to which teachers adhered to intervention guidelines); *reach* (the proportion of classes where teachers reported 80% of pupils were present for session delivery); and *pupil responsiveness* (the extent to which pupils actively participated, engaged with, were interested in, enthusiastic about, and enjoyed the intervention).

<<INSERT TABLE 2 HERE>>

Outcomes

The primary outcome for both SSW and The Guide was intended help-seeking, as assessed by the self-reported personal-emotional problems section of the General Help-Seeking Questionnaire (GHSQ; Wilson et al., 2005). The GHSQ is widely used in young people, assessing intended help-seeking on a 7-point Likert-type scale (from one, extremely unlikely, to seven, extremely likely) for each of 11 different help source options (e.g., friend, parent, teacher) when experiencing personal-emotional problems. Participants mean score was used (range 1 to 7) with higher scores indicating greater intentions to seek help. The GHSQ has previously reported to have good construct, convergent and divergent validity as a measure of intended help-seeking in adolescence, high internal reliability (Cronbach's $\alpha = 0.85$) and test-retest reliability (0.92 over a three-week period; Wilson et al., 2005).

The primary outcome for YAM was emotional difficulties, as assessed by the self-reported Short Moods and Feeling Questionnaire (SMFQ; Angold et al., 1995); a free to download screening tool for depressive symptoms that has been translated into multiple languages. 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). Scores were summed (range 0 to 26) with higher scores indicating greater emotional difficulties. The SMFQ has reported high construct, concurrent and discriminant validity as a measure of depressive symptoms in children and adolescents (Sharp et al., 2006) with Cronbach's α previously reported as 0.91 (Thabrew et al., 2018).

Covariates

Covariates were selected from trial survey data and were included on the basis of theoretical plausibility (following expert discussions) and/or empirical precedent (e.g., baseline

levels of need predicted compliance in the PATHS curriculum; Panayiotou et al., 2020). Pupil-level covariates (baseline score for the primary outcome, sex, baseline emotional difficulties, baseline behavioural difficulties) were entered at the within-level, while school-level covariates (existing mental health provision, staff mental health training, trial wave, proportion of pupils eligible for free school meals) were entered at the between-level (Panayiotou et al., 2020).

Statistical analysis

There are several underpinning assumptions for CACE estimation: 1) the potential outcome of each participant is distinct from the treatment assignment of others; 2) random allocation to treatment and control arms; 3) there are no defiers (i.e., those who do the opposite of what is expected; monotonicity) or always-takers (i.e., those who receive treatment regardless of intervention arm); and, 4) there is no effect of the intervention for non-compliers (i.e., exclusion restriction; Peugh et al., 2017; Jo et al., 2008). As randomisation was at the school level, and control schools did not have access to intervention materials, we were fully confident the first three assumptions were met (Peugh et al., 2017; Berg et al., 2017). Conversely, while the exclusion restriction is often violated in school-based research, strong predictors of compliance can reduce bias in model estimation when this assumption is not met.

For each intervention (SSW primary, SSW secondary, The Guide, YAM), we estimated a two-level CACE model within a latent class mixture framework to delineate pupil- and school-level effects (TYPE = TWOLEVEL MIXTURE). To ensure clustering was appropriately handled, standard errors were adjusted to account for residual non-independence of pupils within schools using a sandwich estimator. Compliance was operationalised as a binary latent variable (U) derived from dosage information (0 = non-complier, 1 = complier), with values set to

missing for participants in the control arm. Treatment allocation was represented by an observed binary variable (IV ; 0 = control, 1 = treatment). Two latent classes were specified corresponding to compliers and non-compliers, defined by fixing the probability of $U = 1$ at 100% and 0%, respectively. Full information maximum likelihood (FIML) estimation was used to handle missing data on the outcome and compliance variables. As in analogous studies, individuals missing covariate data, or both outcome *and* compliance data were excluded due to computational burden and convergence problems associated with their inclusion (Humphrey et al., 2021).

The modelling strategy explicitly combined two elements: (i) estimation of the intervention (CACE) effect, obtained as the coefficient of IV on the primary outcome within the complier class; and (ii) estimation of predictors of compliance, obtained by regressing U on pupil- and school-level covariates within the same mixture framework. Fulfilling part (i), the primary outcome for each intervention (emotional difficulties or intended help-seeking behaviour at 3-6 month follow up) was linearly regressed on IV and all covariates simultaneously (with the exception of behavioural difficulties which was excluded to aid model convergence). The intervention-outcome association was fixed at zero in the non-complier class, consistent with the exclusion restriction assumption (Jo et al., 2008; Berg et al., 2017). Fulfilling part (ii), we conducted multinomial logistic regression of U on all covariates, inclusive of behavioural difficulties, in light of evidence of its relevance to implementation feasibility in comparable school-based interventions (Panayiotou et al., 2020).

Transparency and openness

Data collected and used in the present study are currently being analysed in three additional manuscripts all under review at the time of writing: the current study's companion paper which reports CACE analyses of additional interventions (mindfulness-based exercises, and relaxation practices) delivered in primary and secondary schools in England as part of INSPIRE but not covered herein (Thornton et al., 2025); and two further studies which report the main ITT analyses for INSPIRE and AWARE, respectively (Deighton et al., 2025a, 2025b). CACE analyses of INSPIRE and AWARE interventions were split across two papers to support completeness of reporting and interpretability findings. We report how we determined our sample size, all data exclusions, manipulations, and measures used in the study, and the study follows JARS (Applebaum, et al., 2018). An anonymised dataset will be made publicly available via an online repository in the future (planned 2026). Please contact the principal investigator or trials manager for further information. Study analysis code have been made openly available via the Open Science Framework. All statistical analyses were performed using Mplus version 8.8 and the MplusAutomation package in R version 4.4.2 (Hallquist & Wiley, 2018). Both INSPIRE (Trial registration: ISRCTN16386254) and AWARE (Trial registration: ISRCTN17631228) RCTs were pre-registered prospectively. While CACE estimation was planned and included in the pre-registered statistical analysis plan (Boehnke et al., 2023), the model designs were not pre-registered in detail, reflecting their exploratory secondary nature. Interventions were not accompanied by a standalone manual however, detailed documentation of their content, delivery procedures and the materials used are provided in the Education for Wellbeing technical report and associated documents on the DfE government website (DfE, 2025).

Results

Descriptive statistics: dimensions of implementation

Implementation data revealed consistently high fidelity across all interventions. When sessions were delivered, adherence to the protocol was strong, ranging from 89.31% (YAM) to 95.89% (SSW primary). Reach was similarly high, with 72.32% (YAM) to 96.51% (SSW primary) of classes achieving $\geq 80\%$ pupil attendance during session delivery. Pupil responsiveness scores indicated moderate to high engagement across interventions ($M = 3.57$ to 4.11 on a 5-point scale) while dosage data – expressed as the percentage of classes receiving all scheduled sessions – was impressive but posed greater challenges than other implementation dimensions with 72.16% (SSW secondary) to 84.88% (SSW primary) of classes implementing interventions in full. It should be noted that, in all instances where YAM was delivered, the full programme was implemented across all classes.

The proportion of data missing for each variable used in analysis is provided as Supplementary Table S3. For CACE estimation, dosage was disaggregated to the pupil-level, based on classroom membership. Consequently, $n = 1964$ pupils (81.5%) in The Guide schools were compliers, $n = 2255$ pupils (78%) in YAM schools were compliers, and $n = 1613$ (87.6%) pupils in SSW primary schools and $n = 1209$ (76.4%) of pupils in SSW secondary schools were compliers. The number/proportion of compliers reported here is descriptive and based on the reported dosage information for intervention schools. The proportion of compliers reported hereafter reflect model-estimated values and include ‘would-be’ compliers in control schools. An overview of the key findings from each model are provided below and summary of findings Tables 3 and 4.

Strategies for safety and wellbeing: primary schools (SSW primary)

The analytical sample comprised $n = 3397$ pupils from 64 schools ($n = 3105$ compliers; 91.4%). No CACE effect was observed ($\beta = 0.38$, 95%CI [-0.41, 1.18], $d = 0.05$). Females were less likely to be compliers than males: OR = 0.89, [0.81, 0.97].

Strategies for safety and wellbeing: secondary schools (SSW secondary)

The analytical sample comprised $n = 3692$ pupils from 29 schools ($n = 3215$ compliers; 87.1%). Complete delivery led to significant increases in intended help-seeking scores: $\beta = 1.91$, [1.40, 2.42], $d = 0.24$. None of our predictor variables significantly predicted compliance so results should be interpreted with caution.

<<INSERT TABLE 3 HERE>>

Youth aware of mental health (YAM)

The analytical sample comprised $n = 5408$ pupils from 87 schools ($n = 2745$ compliers; 50.8%). Complete delivery led to significant decreases in SMFQ scores: $\beta = -1.99$, [-2.28, -1.70], $d = -0.97$. Higher levels of baseline emotional difficulties (OR = 0.54, [0.30, 0.96]) and baseline behavioural difficulties (OR = 0.91, [0.84, 0.99]) made compliance less likely.

The mental health and high school curriculum guide (The Guide)

The analytical sample comprised $n = 4879$ pupils from 84 schools ($n = 3483$ compliers; 71.4%). Complete delivery led to a significant increase in GHSQ scores ($\beta = 0.76$, [0.13, 1.39], $d = 0.17$). Relative to those participating in Wave 1, pupils participating in Wave 3 of AWARE were more likely to be compliers ($\beta = 4.61$, [1.06, 8.17]).

<<INSERT TABLE 4 HERE>>

Discussion

Understanding intervention efficacy under ideal implementation conditions is essential for: (1) determining whether null ITT effects reflect programme ineffectiveness or implementation challenges, (2) informing accurate cost-effectiveness calculations, (3) identifying which programmes require complete versus partial delivery, and (4) prioritising implementation support investments. We report results of CACE analyses of interventions undertaken as part of the INSPIRE (SSW) and AWARE (YAM, The Guide) RCTs, representing the first robust estimates of the effects of these interventions that account for implementation variability. Our primary aim was to advance theory surrounding the extent to which universal school-based mental health interventions for improving pupils' emotional difficulties and intended help-seeking behaviour are dependent on complete delivery.

CACE estimation differs from earlier approaches such as last observation carried forward (LOCF), which impute missing or incomplete data under restrictive assumptions that can bias estimates, by instead explicitly modelling compliance as part of the randomised design. Where traditional ITT analyses estimate the average treatment effect regardless of variability in implementation, CACE leverages randomisation while accounting for variability in adherence, thereby producing robust and unbiased estimates of intervention effects among compliers, in this case pupils receiving interventions in full. This makes CACE more robust than earlier imputation-based methods and extends beyond ITT by providing insight into the impact of

receiving the intervention as intended, offering greater interpretive value for both theory and practice.

Statistically significant CACE effects were observed for YAM ($d = -0.97$). This contrasts with no main ITT effect (Deighton et al., 2025b) and indicates that YAM produces significant reductions in emotional difficulties when it is actually delivered. For The Guide, complete delivery amplified the main ITT effect (ITT standardised mean difference = 0.10, CACE $d = 0.17$; Deighton et al., 2025b) on intended help-seeking indicating The Guide is more impactful in the context of complete delivery. For SSW in secondary schools, statistically significant CACE effects contrast with no main ITT effect (Deighton et al., 2025a), indicating that SSW may need to be delivered in full to bring about intended intervention outcomes. Lastly, for SSW in primary schools, the main intervention effect estimate (ITT standardised mean difference = 0.09; Deighton et al., 2025a) was not amplified in the context of compliance, suggesting complete delivery does not offer additional benefit over and above partial delivery.

Although YAM has prior RCT-based evidence of efficacy (e.g., reductions in suicide attempts and suicidal ideation), no data on implementation were previously reported, with authors instead noting, “Process assessments and quality control were done in a standard manner to ensure that all preparatory procedures were executed correctly... [and that that YAM was] implemented in a standard way” (Wasserman et al., 2015). The current study therefore provides the first published estimates of YAM intervention completeness and subsequent treatment modification effects. Particularly noteworthy is the 23.29% classroom-level non-completion rate (i.e., approximately one in five schools randomly allocated to deliver YAM failed to do so). While implementation variability in school-based interventions is inevitable (Durlak, 2015), we expected this to be minimized in YAM, given that trials of teacher-led interventions are more

likely to encounter implementation problems than those delivered by external professionals (Durlak et al., 2011). However, descriptive data on implementation dimensions show that non-completion rates for the teacher-led interventions were analogous to or lower than for YAM (Table 2; Figure 1). This was the case even though non-completion in YAM equated to schools not implementing at all, whereas in SSW and The Guide, non-completion could comprise some or even most sessions.

School staff interviewed in the AWARE implementation and process evaluation noted two key factors that were perceived to be obstructive to take up of YAM. First, delivery by external staff, and the requirement for school staff not to be present during session delivery; to which staff raised concerns about safeguarding and behaviour management (in alignment with our finding that higher levels of baseline emotional and behavioural difficulties made YAM compliance significantly less likely). Second, difficulties coordinating delivery (e.g., finding dates/times that worked for both the school and instructor; displacement of existing timetabled lessons; finding an appropriately large space (e.g., drama studio) to allow multiple small group activities to take place simultaneously; Stapley et al., 2025).

For The Guide, complete implementation amplified the intervention effect on intended help-seeking behaviour identified in the ITT analysis. Like YAM, The Guide also has prior RCT-based evidence of efficacy (Milin et al., 2016), including a recent trial in Wales (Simkiss et al., 2023), demonstrating improvements across several MHL outcome domains (e.g., knowledge, stigma, intended help-seeking). However, in the original Canadian trial, the authors reported that all teachers delivered the curriculum in full (Milin et al., 2016); in the Welsh trial, there was substantial non-compliance, but the effects of this were not assessed (Simkiss et al., 2023). Thus, all three trials provide evidence of the efficacy of this intervention in improving MHL

outcomes, but the current study provides new, rigorous evidence of further benefits gained from complete implementation, over and above the observed main effect. Schools that participated in Wave 3 (baseline in 2022) of AWARE were more likely to deliver in full than schools participating in Wave 1 (baseline in 2018). This may be indicative of a renewed/intensified impetus for the promotion of MHL in schools in the wake of the COVID-19 pandemic, which demonstrably worsened adolescents' mental health (Mansfield et al., 2022).

In contrast to YAM and The Guide, SSW was intentionally developed for INSPIRE, meaning that although its development was informed by a protective behaviours framework (Dale et al., 2016), no prior evidence is available regarding intervention efficacy and subsequent treatment modification effects (i.e., complete versus incomplete implementation). For a teacher-led intervention, dosage levels were high; when disaggregated to the pupil-level over 85% of primary school pupils and over 70% of secondary school pupils received all 8 sessions (Table 2). The SSW curriculum therefore appears feasible for most schools however, in primary school settings, the distinction between complete and incomplete implementation was inconsequential, given CACE estimation indicated complete implementation did not modify the main ITT effect estimate (Deighton et al., 2025a). One could argue that the very high completion rate in primary schools made an ITT effect more likely and a CACE effect less likely, especially given that a substantial number of cases of incomplete implementation still delivered several SSW sessions (undermining confidence in the exclusion restriction in CACE, which assumes that the intervention effect is zero for non-compliers; Peugh et al., 2017). By contrast, complete implementation in secondary schools was substantially (>10%) lower, making the contrast in findings between our ITT and CACE estimates less surprising. Guidelines do not currently specify the level of non-compliance necessary for CACE analyses to provide insights beyond

those obtained through traditional ITT. Collectively, our findings and evidence of null CACE effects from other school-based interventions with very high compliance (e.g., 96%; Kitmitto et al., 2018) suggest that when compliance levels approach ~85% or above, CACE may be less likely to add value.

Another suggested explanation of null effects on intended help-seeking in primary schools is that the SSW programme helped pupils recognise their symptoms were not at the level of needing support hence, no increase in intended help-seeking intentions was observed. Indeed, it is plausible that improvements in mental health literacy can help pupils to better differentiate between what does and does not constitute a mental health disorder, potentially resulting in more appropriate rather than simply more frequent help-seeking. Importantly, our outcome measure captured *intentions* to seek help rather than actual instances of help-seeking. While enhanced understanding might reduce intentions to seek help among pupils who now recognise their experiences as within the normal range, it is less plausible that such education would reduce help-seeking where support is genuinely warranted. Thus, in line with hypothesized intervention effects prospectively stated at the outset of the trial (Hayes et al., 2019a), the absence of an effect on intentions should be interpreted as an indication that the intervention did not achieve its primary outcome, rather than as evidence of a positive impact of increased discernment.

In summary, our analyses indicated that when YAM, The Guide and SSW are implemented, effects in secondary schools are largely contingent on, or amplified by, complete delivery. Accordingly, a key practical implication of the current study is the need for strategies and systems of support that make this more likely. Insights regarding staff perceptions of facilitators to implementation (e.g., recognising intervention benefits; fitting within the school context; ease of implementation; consistency and familiarity of intervention format; flexibility to

make adaptations) and barriers/challenges experienced (e.g., mixed views on a given intervention and its potential benefits; issues of accessibility, interest, and engagement for a given intervention; lack of staff knowledge and confidence; logistical problems, such as those noted above for YAM) derived from the implementation and process evaluations for INSPIRE and AWARE provide a helpful starting point for the likely foci of these (Stapley et al., 2025). Alongside this, a key research implication is the utility of CACE estimation as an important complement to ITT analyses in future RCTs of school-based MHL interventions. Researchers planning such studies would be well advised to prospectively identify (and then collect) theoretically relevant co-variate data to predict compliance; what implementation dimension to use as the compliance proxy; and, where to set the threshold for compliance/noncompliance (Peugh et al., 2017).

A key strength of this study is the use of CACE analysis, which accounts for implementation variability while retaining the bias-protective benefits of randomisation, thereby providing robust estimates of intervention effects under complete delivery. Since we assessed universal school-based interventions (i.e., delivered to all young people, regardless of need), implementation data were primarily collected at the level of the facilitating member of staff (classroom-level); a method standard in large-scale trials with both strengths and limitations that should be born in mind. Adhering to current guidelines for conducting CACE, these data either had to be aggregated to the school-level, or disaggregated to the pupil-level. Data were disaggregated to align with established CACE practices modelling implementation as an individual-level construct (Jo et al., 2008). The approach is particularly salient in light of the limited prior application of multilevel CACE models within empirical research. Aggregation to the school-level would result in data loss and obscure within-school variation, given the

possibility for completeness to differ across classrooms within schools. Moreover, from a conceptual standpoint, intervention exposure operates at the pupil level: while sessions are delivered by implementers, it is pupils who directly experience them, consistent with the broader educational context in which students typically have limited autonomy over participation. (Jo et al., 2008). That said, while staff-reports capture intended delivery, they may be subject to impression management or demand effects, (Humphrey et al., 2016) social desirability or recall bias, and cannot account for fluctuations in attendance at the pupil-level (e.g., a teacher may have delivered all 8 sessions of SSW, but a given pupil may have been absent for some). This is particularly important given the widely reported increases in persistent absenteeism since COVID-19 (Long & Roberts, 2024) and a notable limitation, as true exposure not only depends on delivery, but also pupil presence. Future work could explore ways in which complete delivery within universal interventions is assessed via *both* implementing members of staff and their pupils (e.g., via school attendance records or pupil self-reports) to better distinguish delivery from receipt. Independent observation could also enhance validity, though feasibility in large-scale studies is a challenge. It is also important to note that due to the restrictions of the CACE estimation process, only a single aspect of implementation was considered.

Implementation data indicated that when schools delivered sessions, adherence to protocols and pupil engagement were high, inferring intervention completeness rather than delivery quality represented the primary challenge for schools in this instance hence, completeness was the primary treatment effect modifier in present analyses. Selecting an alternative dimension of implementation as the main indicator of compliance would likely have been affected by ceiling effects and offered limited additional insights beyond ITT analysis. However, this may differ in future interventions. Researchers are advised to assess all available implementation dimensions,

identify that in which delivery was least optimal, and examine whether variation in that dimension contributes meaningfully to intervention outcomes to determine whether offering need-specific implementation support is likely to enhance intervention effectiveness in practice for participating schools. CACE analysis of SSW in secondary schools yielded no statistically significant predictors of compliance (i.e., none of our included variables made complete delivery more or less likely). Given this, the results should be interpreted with a degree of caution. Finally, due to the large number of interventions delivered within INSPIRE and AWARE (which together formed the programme of interventions known as Education for Wellbeing) it was not possible to condense findings for CACE analyses of all interventions into a single paper. CACE estimation conducted for the remaining interventions (mindfulness-based exercises, and relaxations practices) delivered in both primary and secondary schools in England as part of INSPIRE can be found in the current articles companion paper (Thornton et al., 2025).

Conclusions

In conclusion, this Article highlights the importance of complete delivery as a treatment effect modifier of school-based MHL interventions (YAM, The Guide and SSW) in secondary schools and provides important new insights into the treatment effects that can be expected when said interventions are delivered as intended (as opposed to the *offer* of intervention, as in ITT). Findings reinforce the importance of effective strategies and systems of support to overcome commonly identified barriers to implementation.

List of Abbreviations

AWARE *Approaches for Wellbeing and Mental Health literAcy: Research in Education*

CACE *Complier Average Causal Effect*

COVID-19 *Coronavirus Disease 2019*

DfE *Department for Education*

INSPIRE *INterventions in Schools for PromotIng Mental Wellbeing: Research in Education*

ITT *Intention-to-treat*

MHL *Mental Health Literacy*

PATHS *Promoting Alternative Thinking Strategies*

RCT *Randomised Controlled Trial*

SSW *Strategies for Safety and Wellbeing*

The Guide *Mental Health and High School Curriculum Guide*

UP *Usual Practice*

YAM *Youth Aware of Mental Health*

References

- Angold, A., Costello, E. J., Messer, S. C., & Pickles, A. (1995). Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *Journal of Child Psychology and Psychiatry*, 36(5), 237–249.
- Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, 73(1), 3. <https://doi.org/10.1037/amp0000191>
- Berg, J. K., Bradshaw, C. P., Jo, B., & Jalongo, N. S. (2017). Using complier average causal effect estimation to determine the impacts of the Good Behavior Game preventive intervention on teacher implementers. *Administration and Policy in Mental Health*, 44(4), 558–571. <https://doi.org/10.1007/s10488-016-0738-1>
- Boehnke, J. R., Deighton, J., Evans-Lacko, S., Hayes, D., Humphrey, N., Patalay, P., Santos, J., Stapley, E., Thornton, E., Zhang, K., & Thompson, A. (2023). Education for Wellbeing Statistical Analysis Plan. <https://doi.org/10.17605/OSF.IO/H5NBE>
- Dale, R., Shanley, D. C., Zimmer-Gembeck, M. J., Lines, K., Pickering, K., & White, C. (2016). Empowering and protecting children by enhancing knowledge, skills and well-being: A randomized trial of Learn to BE SAFE with Emmy™. *Child Abuse & Neglect*, 51, 368–378. <https://doi.org/10.1016/j.chiabu.2015.07.016>
- Deighton, J., Thompson, A., Hayes, D., March, A., Thornton, E., Santos, J., Patalay, P., Mansfield, R., Deniz, E., Knowles, C., Humphrey, N., Stallard, P., Ashworth, E., Moltrecht, B., Nisbet, K., Mason, C., Stepanous, J., & Boehnke, J. R. (2025a). 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.

<https://doi.org/10.17605/OSF.IO/KXUG7>

- Deighton, J., Hayes, D., Thompson, A., March, A., Thornton, E., Santos, J., Patalay, P., Mansfield, R., Deniz, E., Knowles, C., Humphrey, N., Stallard, P., Ashworth, E., Moltrecht, B., Nisbet, K., Mason, C., Stepanous, J., & Boehnke, J. R. (2025b). School-based intervention study examining universal approaches for well-being and mental health literacy of pupils in Year 9 in England (AWARE): A multi-school, parallel group, cluster-randomised controlled trial. <https://doi.org/10.17605/OSF.IO/KXUG7>
- Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clinical Psychology Review*, 18(1), 23-45. [https://doi.org/10.1016/S0272-7358\(97\)00043-3](https://doi.org/10.1016/S0272-7358(97)00043-3)
- Department for Education (2025). Education for wellbeing programme findings. *Department for Education*. <https://www.gov.uk/government/publications/education-for-wellbeing-programme-findings> [Accessed 05/12/2025]
- Department for Education. (2021). Statutory guidance on relationships education, relationships and sex education (RSE) and health education. *Department for Education*. <https://www.gov.uk/government/publications/relationships-education-relationships-and-sex-education-rse-and-health-education> [Accessed 05/12/2025]
- Durlak, J. A. (2015). Studying program implementation is not easy but it is essential. *Prevention Science*, 16, 1123–1127. <https://doi.org/10.1007/s11121-015-0606-3>

- Durlak, J. A., et al. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82(1), 405–432. <https://doi.org/10.1111/j.1467-8624.2010.01564.x>
- Gupta, S. K. (2011). Intention-to-treat concept: A review. *Perspectives in Clinical Research*, 2(3), 109–112. <https://doi.org/10.4103/2229-3485.83221>
- Hallquist, M. N. & Wiley, J. F. (2018). MplusAutomation: An R Package for Facilitating Large-Scale Latent Variable Analyses in Mplus. *Structural Equation Modeling*, 25, 621-638. <https://doi.org/10.1080/10705511.2017.1402334>
- Hayes, D., Moore, A., Stapley, E., Humphrey, N., Mansfield, R., Santos, J., Ashworth, E., Patalay, P., Bonin, E., Moltrecht, B., Boehnke, J., & Deighton, J. (2019a). 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, 4–8. <https://doi.org/10.1186/s13063-019-3762-0>
- Hayes, D., Moore, A., Stapley, E., Humphrey, N., Mansfield, R., Santos, J., Ashworth, E., Patalay, P., Bonin, E., Evans-Lacko, S., Moltrecht, B., Nisbet, K., Thornton, E., Lange, A., Stallard, P., Thompson, A., Boehnke, J. R., & Deighton, J. (2019b). School-based intervention study examining approaches for well-being and mental health literacy of pupils in Year 9 in England: Study protocol for a multi-school, parallel group cluster randomised controlled trial (AWARE). *BMJ Open*, 9(8), e029044. <https://doi.org/10.1136/bmjopen-2019-029044>
- Humphrey, N., Lendrum, A., Ashworth, E., Frearson, K., Buck, R., & Kerr, K. (2016). Implementation and process evaluation (IPE) for interventions in educational settings: A

- synthesis of the literature. *Education Endowment Foundation*.
https://d2tic4wvo1iusb.cloudfront.net/production/documents/evaluation/evaluation-design/IPE_Handbook.pdf [Accessed 05/12/2025]
- Humphrey, N., Panayiotou, M., Hennessey, A., & Ashworth, E. (2021). Treatment effect modifiers in a randomized trial of the Good Behavior Game during middle childhood. *Journal of Consulting and Clinical Psychology*, 89(8), 668–681.
<https://doi.org/10.1037/ccp0000673>
- Jo, B., Asparouhov, T., Muthén, B., Lalongo, N. S., & Brown C. H. (2008). Cluster randomized trials with treatment noncompliance. *Psychological Methods*, 13(1), 1–18.
<https://doi.org/10.1037/1082-989X.13.1.1>
- Kitmitto, S., González, R., Mezzanote, J., & Chen, Y. (2018). Thinking, doing, talking science: Evaluation report and executive summary. *Education Endowment Foundation*.
<https://dera.ioe.ac.uk/id/eprint/32754/> [Accessed 03/01/2025]
- Long, R., & Roberts, N. (2024). School attendance in England. *House of Commons Library*.
<https://commonslibrary.parliament.uk/research-briefings/cbp-9710/>
[Accessed: 05/12/2025]
- Ma, K. K. Y., Burn, A.-M., & Anderson, J. K. (2023). Review: School-based mental health literacy interventions to promote help-seeking – A systematic review. *Child and Adolescent Mental Health*, 28(3), 408–424. <https://doi.org/10.1111/camh.12609>
- Mansfield, R., Santos, J., Deighton, J., Hayes, D., Velikonja, T., Boehnke, J. R., & Patalay, P. (2022). The impact of the COVID-19 pandemic on adolescent mental health: A natural experiment. *Royal Society Open Science*, 9(4), 211114.
<https://doi.org/10.1098/rsos.211114>

- McCoy, C. E. (2017). Understanding the intention-to-treat principle in randomized controlled trials. *Western Journal of Emergency Medicine*, 18(6), 1075–1078.
<https://doi.org/10.5811/westjem.2017.8.35985>
- Milin, R., Kutcher, S., Lewis, S. P., Walker, S., Wei, Y., Ferrill, N., & Armstrong, M. A. (2016). Impact of a mental health curriculum on knowledge and stigma among high school students: A randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, 55(5), 383–391.e1. <https://doi.org/10.1016/j.jaac.2016.02.018>
- Mills, J., Cumbers, S., Williams, S., & Titley, W. (2023). The effectiveness of mental health literacy interventions in young people: A meta-analysis. *Mental Health Review Journal*, 28(4), 386–395. <https://doi.org/10.1108/MHRJ-07-2022-0052>
- NHS Digital. (2023). Mental health of children and young people in England, 2023. NHS Digital. <https://www.gov.uk/government/statistics/mental-health-of-children-and-young-people-in-england-2023-wave-4-follow-up-to-the-2017-survey> [Accessed 05/12/2025]
- 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. *Journal of Educational Psychology*, 112(3), 514–528. <https://doi.org/10.1037/edu0000360>
- Peugh, J. L., Strotman, D., McGrady, M., Rausch, J., & Kashikar-Zuck, S. (2017). Beyond intent to treat (ITT): A complier average causal effect (CACE) estimation primer. *Journal of School Psychology*, 60, 7–24. <https://doi.org/10.1016/j.jsp.2015.12.006>
- Proctor, E. K., Powell, B. J., and McMillen, J. C. (2013). Implementation strategies: recommendations for specifying and reporting. *Implementation Science*, 8(1), 139. <https://doi.org/10.1186/1748-5908-8-139>

- Rudd, B. N., Davis, M., and Beidas, R. S. (2020). Integrating implementation science in clinical research to maximize public health impact: a call for the reporting and alignment of implementation strategy use with implementation outcomes in clinical research. *Implementation Science*, 15(1), 103. <https://doi.org/10.1186/s13012-020-01060-5>
- Seedaket, S., Turnbull, N., Phajan, T., & Wanchai, A. (2020). Improving mental health literacy in adolescents: Systematic review of supporting intervention studies. *Tropical Medicine and International Health*, 25(9), 1055–1064. <https://doi.org/10.1111/tmi.13449>
- Sharp, C., Goodyer, I. M., & Croudace, T. J. (2006). The Short Mood and Feelings Questionnaire (SMFQ): a unidimensional item response theory and categorical data factor analysis of self-report ratings from a community sample of 7-through 11-year-old children. *Journal of Abnormal Child Psychology*, 34, 365-377. <https://doi.org/10.1007/s10802-006-9027-x>
- Simkiss, N. J., Gray, N. S., Kemp, A. H., Dunne, C., & Snowden, R. J. (2023). A randomised controlled trial evaluating the Guide Cymru mental health literacy intervention programme in year 9 (age 13–14) school pupils in Wales. *BMC Public Health*, 23(1), 1062. <https://doi.org/10.1186/s12889-023-15922-2>
- Thabrew, H., Stasiak, K., Bavin, L. M., Frampton, C., & Merry, S. (2018). Validation of the mood and feelings questionnaire (MFQ) and short mood and feelings questionnaire (SMFQ) in New Zealand help-seeking adolescents. *International Journal of Methods in Psychiatric Research*, 27(3), e1610. <https://doi.org/10.1002/mpr.1610>
- Stapley, E., March, A., Mansfield, R., Ashworth, E., Burrell, K., Thompson, A., Ravaccia, G., Deighton, J., & Hayes, D. (2025). A qualitative study of school staff experiences of

implementing five universal mental health interventions in England. (Pre-print).

<https://doi.org/10.17605/OSF.IO/KXUG7>

Thornton, E., Knowles, C., Panayiotou, M., Santos, J., Ashworth, E., Mason, C., Hayes, D., Thompson, A., Deighton, J., Mansfield, R., March, A., Stallard, P., & Humphrey, N. (2025). Implementation dosage as a modifier of treatment effects of universal mindfulness and relaxation interventions on emotional difficulties in primary and secondary schools: complier average causal effect estimation in the INSPIRE cluster randomised trial. (Pre-print). <https://doi.org/10.17605/OSF.IO/KXUG7>

Wasserman, D., Hoven, C. W., Wasserman, C., Wall, M., Eisenberg, R., Hadlaczky, G., Kelleher, I., Sarchiapone, M., Apter, A., Balazs, J., Bobes, J., Brunner, R., Corcoran, P., Cosman, D., Guillemin, F., Haring, C., Iosue, M., Kaess, M., Kahn, J., ..., Carli, V. (2015). School-based suicide prevention programmes: The SEYLE cluster-randomised, controlled trial. *Lancet*, 385(9977), 1536–1544.
[https://doi.org/10.1016/S0140-6736\(14\)61213-7](https://doi.org/10.1016/S0140-6736(14)61213-7)

Wilson, C. J., Deane, F. P., Ciarrochi, J., & Rickwood, D. (2005). Measuring help-seeking intentions: Properties of the General Help-Seeking Questionnaire. *Canadian Journal of Counselling*, 39(1), 15–28.

Table 1*Descriptive and baseline mental health provision statistics for each trial arm*

		INSPIRE				AWARE		
Variable		SSW (Primary)	Control	SSW (Secondary)	Control	The Guide	YAM	Control
<i>Pupil-level factors</i>								
Pupils	<i>N</i>	2566	2088	2850	2988	3997	4028	4141
Sex	Female <i>n</i> (%)	1220 (49.00)	996 (47.72)	1504 (52.77)	1600 (53.57)	1943 (49.92)	2107 (52.74)	1938 (46.85)
	Male <i>n</i> (%)	1270 (51.00)	1091 (52.28)	1346 (47.23)	1387 (46.43)	1949 (50.08)	1888 (47.26)	2199 (53.15)
Behavioural difficulties	T1 Mean (SD)	3.23 (2.29)	3.19 (2.30)	3.36 (2.37)	3.21 (2.22)	3.74 (2.55)	3.76 (2.55)	3.56 (2.51)
Emotional difficulties	T1 Mean (SD)	6.65 (5.03)	6.62 (4.93)	6.52 (5.72)	6.11 (5.33)	7.48 (6.36)	7.54 (6.30)	7.03 (6.15)
	T2 Mean (SD)	--	--	--	--	--	7.59 (6.57)	7.02 (6.40)
Intended help-seeking	T1 Mean (SD)	3.23 (1.26)	3.30 (1.32)	3.20 (1.15)	3.18 (1.12)	3.05 (1.14)	--	2.99 (1.12)
	T2 Mean (SD)	3.26 (1.25)	3.16 (1.29)	3.24 (1.15)	3.15 (1.12)	3.12 (1.22)	--	2.99 (1.15)
<i>School-level factors</i>								
Schools	<i>N</i>	36	36	18	18	50	51	52
Wave	1 <i>n</i> (%)	36 (100.00)	36 (100.00)	10 (55.55)	11 (61.11)	23 (46.00)	24 (47.06)	24 (46.15)
	3 <i>n</i> (%)	0 (0.00)	0 (0.00)	8 (44.44)	7 (38.89)	27 (54.00)	27 (52.94)	28 (53.85)
Free school meals *	<i>n</i> (%)	12 (33.33)	13 (36.1)	6 (33.3)	5 (27.8)	16 (32.0)	18 (35.3)	17 (32.7)
Staff training	% (SD)	7.6 (8.8)	10.0 (8.5)	8.8 (5.0)	17.7 (11.1)	14.3 (10.0)	12.3 (10.7)	15.5 (11.1)
Prior interventions	Yes <i>n</i> (%)	15 (41.67)	15 (41.67)	10 (55.56)	9 (50.00)	30 (60.00)	30 (58.82)	32 (61.54)

Note: descriptive data reflect the sample proportions for pupils allocated to each trial arm prior to model estimation. See Tables 3 & 4 for model estimated proportions

** Number of schools in the top national tertile for proportion of pupils eligible for free school means (school-level measure)*

Table 2*Descriptives statistics of implementation dimensions reported by staff at the classroom-level*

	SSW ¹ primary	SSW ² secondary	The Guide ³	YAM ⁴
<i>Classroom-level</i>				
Allocated intervention (n)	100	115	152	156
Implemented at least partially (n; %)	85 (85.00)	81 (70.43)	124 (81.58)	112 (71.79)
Confirmed no implementation (ITT) (n; %)	1 (1.00)	16 (13.91)	6 (3.95)	34 (21.79)
Missing implementation data (n; %)	14 (14.00)	18 (15.65)	22 (14.47)	10 (6.41)
Fidelity ^a (%; SD)	95.89 (5.47)	92.17 (10.55)	89.95 (10.45)	89.31 (9.64)
Reach ^b (%)	96.51	95.06	87.10	72.32
Pupil responsiveness ^c (Mean; SD)	4.11 (0.70)	3.97 (0.77)	3.57 (0.76)	3.74 (0.86)
Dosage (Complete implementation) ^d (n; %)	73/86 (84.88)	70/97 (72.16)	100/130 (76.92)	112/146 (76.71)
<i>Pupil-level</i>				
Dosage (Complete implementation) ^e (n; %)	1935 (85.13)	1763 (70.95)	2642 (76.31)	2930 (76.68)

Note. Implementation statistics were derived from classes whose teacher returned an implementation survey; ITT schools did not provide survey data. Dosage was calculated using survey responses supplemented with administrative information indicating that ITT schools delivered 0 sessions, thereby capturing (at least) partial and non-implementation. In contrast, fidelity, reach, and pupil responsiveness were calculated only for classes with partial implementation, as survey data were unavailable for non-respondents.

¹ Implementation data were available for 85 classes allocated SSW in primary schools. There were 8 sessions of SSW in total.

² Implementation data were available for 81 classes allocated SSW in secondary schools. There were 8 sessions of SSW in total.

³ Implementation data were available for 124 classes allocated The Guide. There were 6 sessions of The Guide in total.

⁴ Implementation data were available for 112 classes allocated YAM. There were 5 sessions of YAM in total.

^a Extent to which teachers followed the guidance in the manual for each session of the intervention delivered.

^b Proportion of classes where teachers reported $\geq 80\%$ were present for session delivery.

^c Summed score: pupil participation, interest in, enthusiasm about, engagement with, and enjoyment of the intervention (scored 1-5).

^d Number and percentage of classes with implementation data delivering intervention in full (all scheduled sessions).

^e Dosage data disaggregated to the pupil-level. The number of compliers included in the analytical sample per trial arm.

Table 3*CACE Estimation for SSW in Primary and Secondary Schools (INSPIRE)*

School Type	Predictors of Help-Seeking for Pupils Allocated SSW β (SE)			
	Primary (Entropy = .85)		Secondary (Entropy = .87)	
	Non-compliers 419 (12.3%)	Compliers ¹ 2978 (87.7%)	Non-compliers 569 (15.4%)	Compliers ¹ 3123 (84.6%)
<i>Pupil-level predictors</i>				
Sex (0=Male; 1=Female)	.29 (.17)	.05 (.04)	.13 (.18)	.18 (.04)***
Emotional Difficulties (T1)	-.04 (.15)	.05 (.02)*	-.14 (.04)**	-.08 (.02)***
Intended help-Seeking (T1)	.12 (.07)	.28 (.02)***	-.06 (.27)	.39 (.04)***
<i>School-level predictors</i>				
SSW vs Control (CACE)	--	.38 (.40)	--	1.91 (.26)***
Wave (0=Wave 1; 1=Wave 3)	--	--	2.11 (.52)***	-.82 (.35)*
Free School Meals (0=No; 1=Yes)	.22 (.97)	.37 (.34)	-.56 (1.44)	.57 (.41)
Prior Interventions (0=No; 1=Yes)	.36 (1.16)	-.75 (.38)*	-1.31 (2.20)	.10 (.41)
Staff Training	-.57 (.33)	.18 (.13)	.40 (.72)	.50 (.22)*
<i>Compliance predictors ^a</i>				
	Sex: b = -.12 (.04)**; OR = .89 [.81, .97]			

¹ *Complier class comprises pupils whose dosage data shows they complied in the intervention arm alongside probabilistically estimated 'would-be' compliers in the control arm.*

^a *Statistically significant predictors of compliance for each model. CACE effects are marked in bold.*

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

OR, odds ratio

Table 4*CACE Estimation for YAM and The Guide (AWARE)*

School Type	Emotional Difficulties – YAM β (SE)		Help-Seeking – The Guide β (SE)	
	Secondary (Entropy = .84)		Secondary (Entropy = .75)	
	Non-compliers 2572 (47.6%)	Compliers ¹ 2836 (52.4%)	Non-compliers 1399 (28.7%)	Compliers ¹ 3480 (71.3%)
<i>Pupil-level predictors</i>				
Sex (0=Male; 1=Female)	.28 (.04)***	.22 (.05)***	.12 (.08)	.09 (.03)**
Emotional Difficulties (T1)	.59 (.02)***	.55 (.03)***	-.05 (.06)	-.02 (.03)
Intended help-Seeking (T1)	--	--	.48 (.05)***	.33 (.03)***
<i>School-level predictors</i>				
YAM/The Guide vs Control (CACE)	--	-1.99 (.15)***	--	.76 (.32)*
Wave (0=Wave 1; 1=Wave 3)	.23 (.43)	-.01 (.09)	-.83 (.98)	-1.35 (.19)***
Free School Meals (0=No; 1=Yes)	.44 (.53)	-.13 (.11)	.09 (.82)	-.10 (.39)
Prior Interventions (0=No; 1=Yes)	-.41 (.42)	-.26 (.11)*	.18 (.76)	.55 (.29)
Staff Training	-.03 (.23)	-.08 (.06)	.44 (.29)	.51 (.12)***
<i>Compliance predictors^a</i>	Emotional Difficulties (T1): b = -.62 (.29)*; OR = .54 [.30, .96] Behavioural Difficulties (T1): b = -.09 (.04)*; OR = .91 [.84, .97]		Wave: b = 11.13 (4.38)*	

¹ *Complier class comprises pupils whose dosage data shows they complied in the intervention arm alongside probabilistically estimated 'would-be' compliers in the control arm.*

^a *Statistically significant predictors of compliance for each model. CACE effects are marked in bold.*

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

OR, odds ratio with 95% confidence intervals

Figure 1. Bar charts illustrating the spread of sessions delivered across classes, disaggregated to the pupil-level

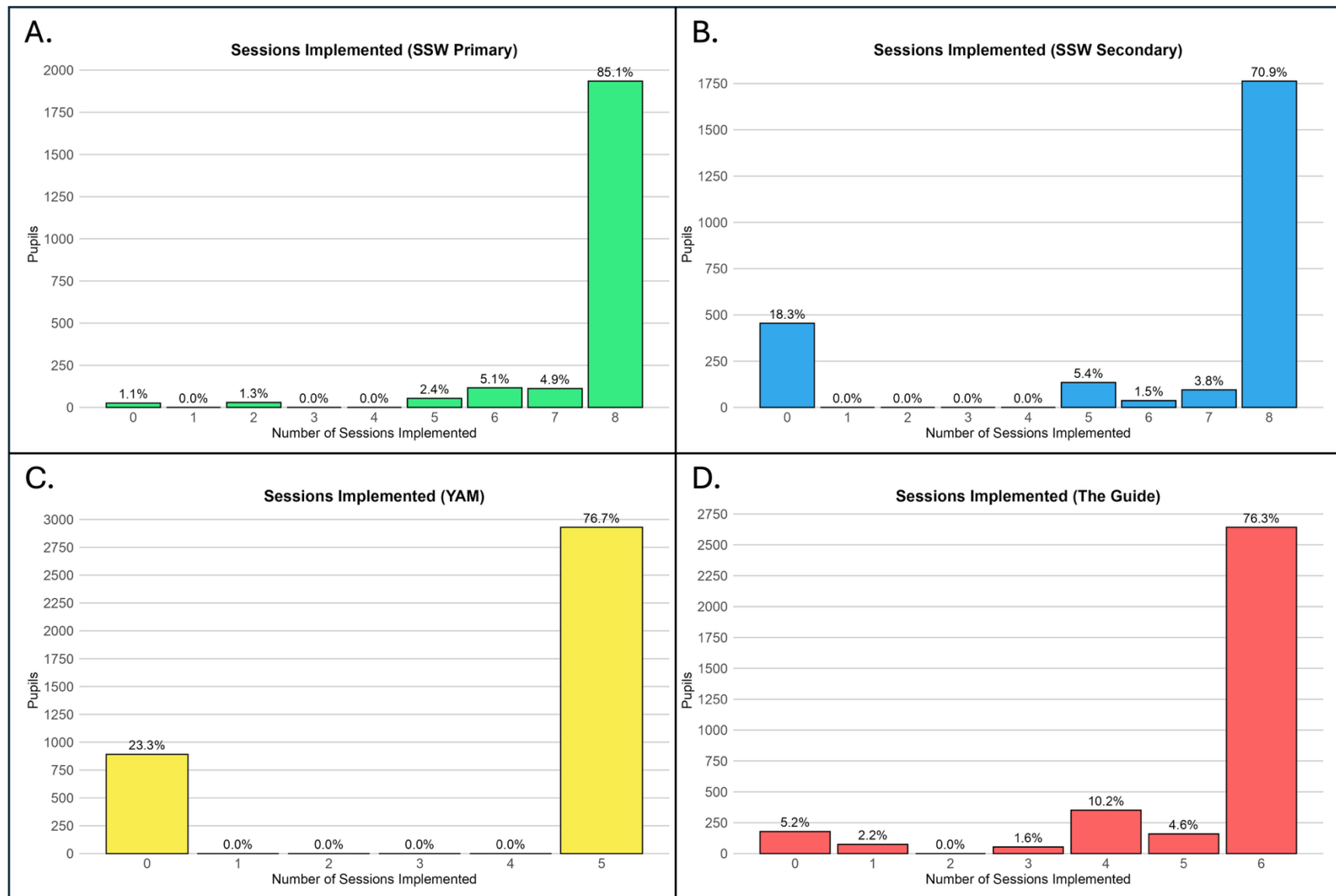


Figure 2. CONSORT Diagram Depicting Flow of Schools and Pupils through INSPIRE

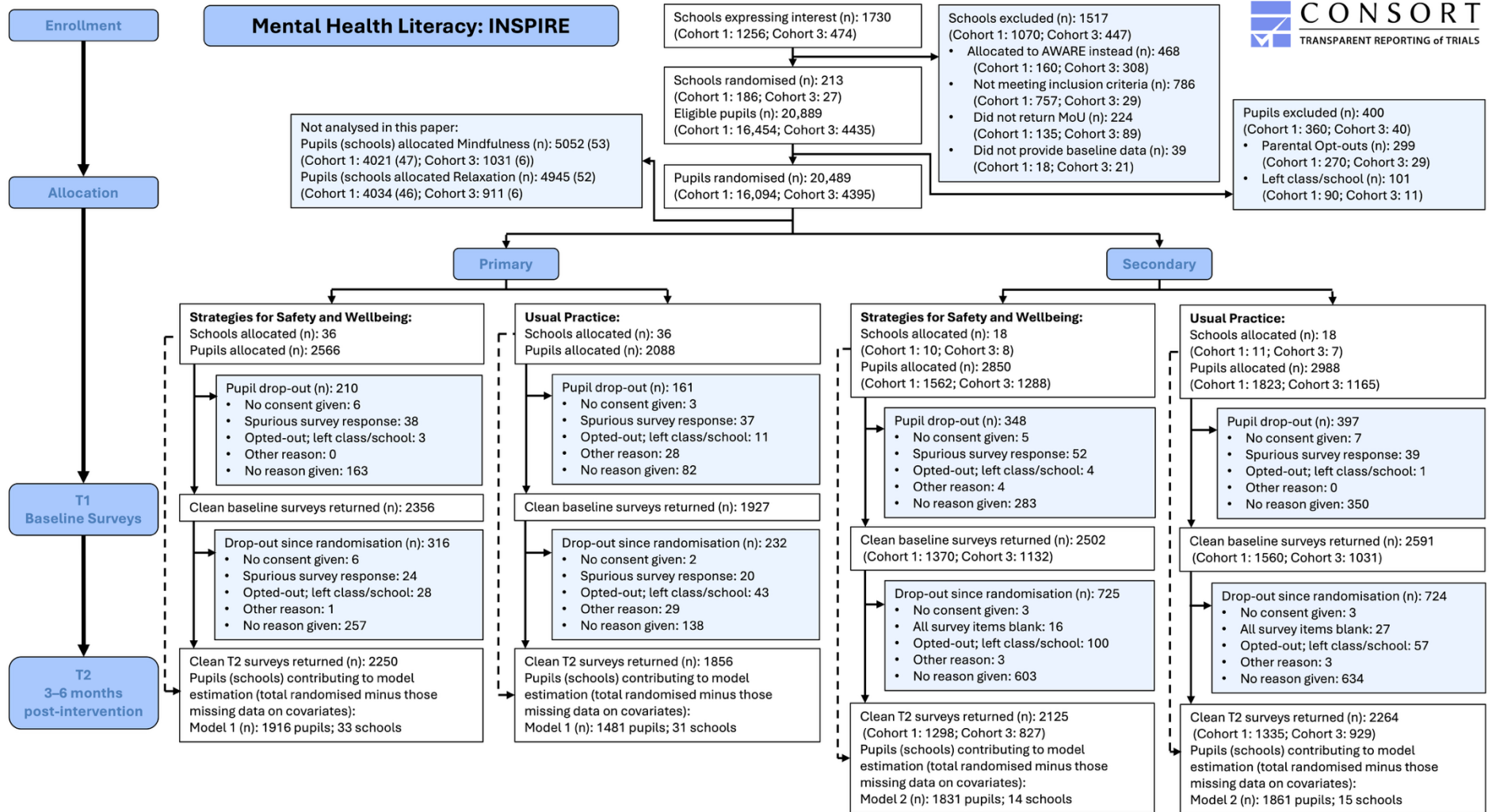
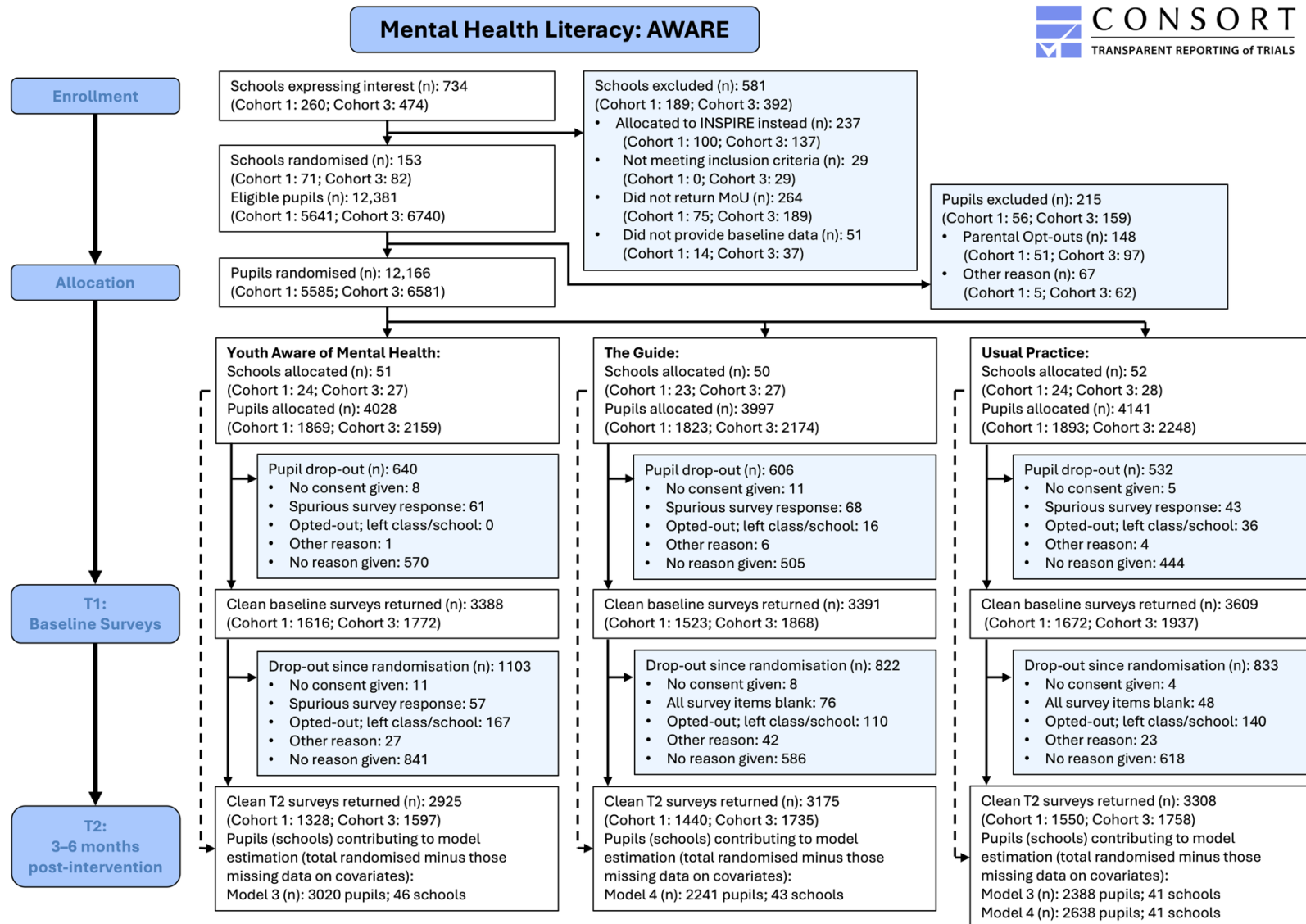


Figure 3. CONSORT Diagram Depicting Flow of Schools and Pupils Through AWARE



Supplementary Table S1: Additional Demographic Information for Each Trial Arm

		INSPIRE				AWARE		
		SSW Primary	Control	SSW Secondary	Control	The Guide	YAM	Control
<i>Pupil-level factors</i>								
Pupils	<i>N</i>	2566	2088	2850	2988	3997	4028	4141
Sex	Female <i>n</i> (%)	1220 (49.0)	996 (47.7)	1504 (52.8)	1600 (53.6)	1943 (49.9)	2107 (52.7)	1938 (46.8)
	Male <i>n</i> (%)	1270 (51.0)	1091 (52.3)	1346 (47.2)	1387 (46.4)	1949 (50.1)	1888 (47.3)	2199 (53.2)
Year Group	4 <i>n</i> (%)	1234 (48.1)	1049 (50.3)	n/a	n/a	n/a	n/a	n/a
	5 <i>n</i> (%)	1330 (51.9)	1037 (49.7)	n/a	n/a	n/a	n/a	n/a
	7 <i>n</i> (%)	n/a	n/a	1440 (50.5)	1515 (50.7)	n/a	n/a	n/a
	8 <i>n</i> (%)	n/a	n/a	1410 (49.5)	1473 (49.3)	n/a	n/a	n/a
	9 <i>n</i> (%)	n/a	n/a	n/a	n/a	3997 (100.0)	4028 (100.0)	4140 (100.0)
Ethnicity	White <i>n</i> (%)	1904 (74.2)	1696 (81.2)	2405 (84.4)	2220 (74.3)	2919 (73.0)	2724 (67.6)	2706 (65.4)
	Ethnic Minority <i>n</i> (%)	587 (22.9)	307 (14.7)	401 (14.1)	571 (19.1)	904 (22.4)	3081 (25.3)	972 (24.3)
Special Ed. Needs	Yes <i>n</i> (%)	395 (15.4)	347 (16.6)	403 (14.1)	345 (11.6)	613 (15.3)	550 (13.7)	528 (12.8)
Behavioural Difficulties	T1 Mean (SD)	3.23 (2.29)	3.19 (2.30)	3.36 (2.37)	3.21 (2.22)	3.74 (2.55)	3.76 (2.55)	3.56 (2.51)
Emotional Difficulties	T1 Mean (SD)	6.65 (5.03)	6.62 (4.93)	6.52 (5.72)	6.11 (5.33)	7.48 (6.36)	7.54 (6.30)	7.03 (6.15)
Emotional Difficulties	T2 Mean (SD)	n/a	n/a	n/a	n/a	n/a	7.59 (6.57)	7.02 (6.40)
Intended Help-Seeking	T1 Mean (SD)	3.23 (1.26)	3.30 (1.32)	3.20 (1.15)	3.18 (1.12)	3.05 (1.14)	n/a	2.99 (1.12)
Intended Help-Seeking	T2 Mean (SD)	3.26 (1.25)	3.16 (1.29)	3.24 (1.15)	3.15 (1.12)	3.12 (1.22)	n/a	2.99 (1.15)
<i>School-level factors</i>								
Schools	<i>N</i>	36	36	18	18	50	51	52
Wave	1 <i>n</i> (%)	36 (100.0)	36 (100.0)	10 (55.6)	11 (61.1)	23 (46.0)	24 (47.1)	24 (46.2)
	3 <i>n</i> (%)	0 (0)	0 (0)	8 (44.4)	7 (38.9)	27 (54.0)	27 (52.9)	28 (53.8)
Region	North-East <i>n</i> (%)	8 (22.2)	8 (22.2)	2 (11.1)	1 (5.6)	7 (14)	7 (13.7)	6 (11.5)
	North-West <i>n</i> (%)	10 (27.8)	10 (27.8)	4 (22.2)	4 (22.2)	5 (10)	10 (19.6)	8 (15.4)
	South-East <i>n</i> (%)	9 (25.0)	10 (27.8)	7 (38.9)	9 (50.0)	25 (50)	24 (47.1)	25 (48.1)
	South-West <i>n</i> (%)	9 (25.0)	8 (22.2)	5 (27.8)	4 (22.2)	13 (26)	10 (19.6)	13 (25)
Urban	<i>n</i> (%)	10 (27.8)	11 (30.6)	3 (16.7)	5 (27.8)	23 (46)	23 (45.1)	25 (48.1)
Rural	<i>n</i> (%)	26 (72.2)	25 (69.4)	15 (83.3)	13 (72.2)	27 (54.0)	28 (54.9)	27 (51.9)
Free School Meals	<i>n</i> (%)	12 (33.3)	13 (36.1)	6 (33.3)	5 (27.8)	16 (32.0)	18 (35.3)	17 (32.7)
Staff Training	% (SD)	7.6 (8.8)	10.0 (8.5)	8.8 (5.0)	17.7 (11.1)	14.3 (10.0)	12.3 (10.7)	15.5 (11.1)
Universal Provision	% (SD)	11.3 (4.7)	11.9 (5.6)	14.9 (8.0)	21.5 (12.2)	17.9 (8.8)	16.8 (7.9)	20.0 (10.3)
Targeted Provision	% (SD)	7.1 (4.9)	9.5 (6.8)	12.6 (7.5)	18 (13.3)	15.3 (7.8)	15.7 (9.5)	17.7 (8.9)
Prior Interventions	Yes <i>n</i> (%)	15 (41.7)	15 (41.7)	10 (55.6)	9 (50.0)	30 (60.0)	30 (58.8)	32 (61.5)

Supplementary Table S2: Specifying the Implementation and Reporting Strategies for Each Intervention ([Proctor et al., 2013](#))

Intervention & Discrete Intervention Strategies*	Actor	The action	Action targets	Temporality	Dose	Outcomes assessed	Theoretical justification
SSW - Primary Schools: Conduct educational meetings; Develop a formal implementation blueprint; Develop educational materials; Develop resource sharing agreements; Distribute educational materials; Facilitation; Identify and prepare champions; Obtain formal commitments; Use other payment schemes	School staff	Equip pupils with safety and wellbeing skills through interactive lessons rooted in the protective behaviours framework. Trained staff guide pupils through sequenced activities such as identifying bodily and emotional warning signs, mapping personal safety networks, rehearsing safe friendship behaviours, and practising emotional regulation strategies. Staff use lesson plans to structure classroom dialogue, model protective strategies, and prompt pupils to reflect on their own coping resources, thereby embedding skills into pupils' everyday contexts. Main themes covered in each session: 1) It is safe to talk about mental health; 2) You are never too young to talk mental health; 3) What is safety?; 4) Early warning signs – noticing our bodies; 5) Early warning signs – noticing our feelings and thoughts; 6) Developing our safety networks; 7) Safe friendships; and, 8) Safe ways of managing emotions.	Educate pupils on mental health. Increase help-seeking behaviour in school pupils in need of mental health support	1. Evaluation team train school staff in a half-day workshop prior to implementation 2. Evaluation team provide lesson plans adapted for primary schools 3. School staff deliver eight weekly 40-minute sessions	8 weekly sessions 40-minutes (Total dose: 5 hours 20 minutes)	Intended help-seeking behaviour; measured using the 18-item (self-report) General Help-Seeking Questionnaire (GHSQ)	Based on a protective behaviours framework (Dale et al., 2016 ; but adapted to focus more on mental health and wellbeing rather than safeguarding), which emphasises that recognising early warning signs, developing safety networks, and practising safe coping strategies can protect children and young people from harm and promote wellbeing. Sequenced lessons are intended to build knowledge, interpersonal safety skills, and emotional regulation capacity, thereby reducing vulnerability and enhancing resilience.

SSW - Secondary Schools:	School staff	Equip pupils with safety and wellbeing skills through interactive lessons rooted in the protective behaviours framework. Trained staff guide pupils through sequenced activities such as identifying bodily and emotional warning signs, mapping personal safety networks, rehearsing safe friendship behaviours, and practising emotional regulation strategies. Staff use lesson plans to structure classroom dialogue, model protective strategies, and prompt pupils to reflect on their own coping resources, thereby embedding skills into pupils' everyday contexts. Main themes covered in each session: 1) It is safe to talk about mental health; 2) We all have mental health; 3) What is safety?; 4) Early warning signs – noticing our bodies; 5) Early warning signs – noticing our feelings and thoughts; 6) Developing our safety networks; 7) Safe friendships; and, 8) Safe ways of managing emotions	Educate pupils on mental health. Increase help-seeking behaviour in school pupils in need of mental health support	1. Evaluation team train school staff in a half-day workshop prior to implementation 2. Evaluation team provide lesson plans adapted for secondary schools 3. School staff deliver eight weekly 40-minute sessions	8 weekly sessions; 40-minutes (Total dose: 5 hours 20 minutes)	Intended help-seeking behaviour; measured using the 18-item (self-report) General Help-Seeking Questionnaire (GHSQ)	Based on a protective behaviours framework (Dale et al., 2016 ; but adapted to focus more on mental health and wellbeing rather than safeguarding), which emphasises that recognising early warning signs, developing safety networks, and practising safe coping strategies can protect children and young people from harm and promote wellbeing. Sequenced lessons are intended to build knowledge, interpersonal safety skills, and emotional regulation capacity, thereby reducing vulnerability and enhancing resilience.
The Guide:	School staff	Deliver structured lessons that combine interactive discussions, stigma-reduction exercises, and factual teaching about mental disorder led by trained school staff. Teachers engage pupils in dialogue, present case examples of mental disorders, and assign	Educate pupils on mental health. Increase help-seeking behaviour in school pupils in need of	1. Evaluation team train school staff in a half-day workshop before delivery 2. Evaluation team provide structured	6 weekly sessions; 60-minutes (Total dose: 6 hours)	Intended help-seeking behaviour; measured using the 18-item (self-report) General Help-Seeking	Informed by mental health literacy theory, which proposes that increasing knowledge, reducing stigma, and clarifying when and how to seek help can positively influence mental health outcomes.

Develop educational materials; Develop resource sharing agreements; Distribute educational materials; Facilitation; Identify and prepare champions; Obtain formal commitments; Use other payment schemes		homework activities to consolidate knowledge. The approach emphasises education of mental health and illness, challenging misconceptions, and reinforces help-seeking behaviours through repeated staff–pupil interactions. Main themes covered in each session: 1) The stigma of mental illness; 2) Understanding the relationship between mental health and mental illness; 3) Understanding specific mental illnesses; 4) Adolescents’ experiences of mental illness; 5) Seeking help and finding support; 6) The importance of positive mental health.	mental health support	3.	lesson plans and homework tasks School staff deliver six weekly 60-minute sessions	Questionnaire (GHSQ)	By enhancing understanding of mental illness and challenging misconceptions, adolescents are expected to improve attitudes toward mental health and be more willing and able to access support. Training staff also strengthens the school environment’s capacity to deliver accurate, sustainable mental health education.	
YAM: Conduct educational meetings; Develop a formal implementation blueprint; Develop educational materials; Develop resource sharing agreements; Distribute educational materials; Facilitation; Identify and prepare champions;	Professional facilitators external to the school trained in YAM	Engage pupils in experiential learning through structured role-play and guided group discussions led by trained external instructors. Sessions create opportunities for pupils to practise recognising, expressing, and managing emotional difficulties, while also rehearsing adaptive coping and help-seeking strategies in response to challenging scenarios. Instructors facilitate dialogue, encourage peer support, and provide reflective feedback to strengthen emotional regulation and resilience. Main themes covered in each session: 1) What is mental health?; 2) Self-help advice;	Educate pupils on mental health. Reduce pupils’ risk of experiencing emotional difficulties	1.	Evaluation team recruit professional facilitator trained in delivery of YAM	5 weekly sessions; 60-minutes (Total dose: 5 hours)	Emotional difficulties; measured using the 13-item (self-report) Short Mood and Feelings Questionnaire (SMFQ)	Grounded in experiential learning and problem-solving theory, positing that active engagement in role-play and peer dialogue enables adolescents to practise adaptive coping responses. Providing a safe space for discussion is theorised to reduce emotional difficulties by enhancing self-awareness, normalising emotions, and strengthening adaptive regulation strategies.
				2.	External facilitators deliver sessions			

Obtain formal commitments; Use other payment schemes	3) Stress and crisis; 4) Depression and suicidal thoughts; 5) Helping a friend in need; 6) Who can I ask for advice?
---	--

SSW: primary, Strategies for Safety and Wellbeing: Primary Schools; *SSW: secondary*, Strategies for Safety and Wellbeing: Secondary Schools; *The Guide*, The Mental Health and High School Curriculum Guide; *YAM*, Youth Aware of Mental Health

* Discrete implementation strategies adopted to support intervention implementation (Powell et al., [2015](#)).

Supplementary Table S3: Proportion of Missing Data for Each Trial Arm

[illegible]

Note: Missing data is higher for the AWARE control group for some variables as usual practice pupils were presented either help-seeking or emotional difficulties items depending on the trial arm they were assigned to. Only those who returned data on the primary outcome for either trial were included in respective models