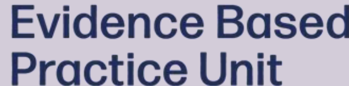
EVALUATION REPORT

More Good Days at School

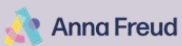
Trial report

Dr Jessica Stepanous, Dr Tanya Lereya, Dr Emily Stapley, Ayesha Sheikh, Arthur Pander Maat and Prof Julian Edbrooke-Childs

April 2026



A partnership of



Home Office

About the Youth Endowment Fund

The Youth Endowment Fund (YEF) is a charity with a mission that matters. We exist to prevent children and young people from becoming involved in violence. We do this by finding out what works and building a movement to put this knowledge into practice.

Children and young people at risk of becoming involved in violence deserve services that give them the best chance of a positive future. To make sure that happens, we'll fund promising projects and then use the very best evaluation to find out what works. Just as we benefit from robust trials in medicine, young people deserve support grounded in the evidence. We'll build that knowledge through our various grant rounds and funding activities.

And just as important, is understanding children and young people's lives. Through our Youth Advisory Board and national network of peer researchers, we'll ensure they influence our work and that we understand and are addressing their needs. But none of this will make a difference if all we do is produce reports that stay on a shelf.

Together, we need to look at the evidence and agree what works, then build a movement to make sure that young people get the very best support possible. Our strategy sets out how we'll do it. At its heart, it says that we will fund good work, find what works and work for change. You can read it here.

For more information about the YEF or this report, please contact:

Youth Endowment Fund
C/O Impetus
10 Queen Street Place
London
EC4R 1AG

www.youthendowmentfund.org.uk
hello@youthendowmentfund.org.uk
Registered Charity Number: 1185413

Contents

About the Youth Endowment Fund **2**

About the evaluator **5**

Acknowledgements **5**

Executive summary **6**

The project 6

Introduction **8**

Overview of evaluation 10

Intervention 10

Evaluation objectives 15

Ethics and trial registration 17

Data protection 17

Project team/stakeholders 18

Methods **21**

Trial design 21

Participant selection 23

Outcome measures 24

Sample size 30

Randomisation 31

Statistical analysis 31

Implementation and process evaluation 35

Impact evaluation results **39**

Participant flow, including losses and exclusions 39

Attrition 43

Participant characteristics 44

Outcomes and analysis 47

Implementation and process evaluation results **60**

Implementation 60

Quality and experience 64

Perceived need 68

Barriers and facilitators 69

Changes 76

Long-term impact and sustainability 84
Cost information 86

Conclusion **90**

Impact evaluation and implementation process evaluation integration 90

References **96**

Appendix A: Recruitment documents: Memorandum of Understanding, Pupil Information Sheet and Pupil Privacy Notice **100**

Appendix B: Effect size estimation **112**

Appendix C: Full model output **114**

Appendix D: Logic model **130**

Appendix E. YEF Security Rating **132**

About the evaluator

The Evidence Based Practice Unit (EBPU) is a partnership between Anna Freud and University College London's Faculty of Brain Sciences. Established in 2006, it combines cutting-edge research with innovative practice to improve the mental health and wellbeing of children and young people. The EBPU focuses on four core themes – risk, resilience, change and choice – conducting evaluations, developing tools, and shaping policy and practice in schools, health services and communities. Its work is underpinned by a commitment to equity, diversity and inclusion, ensuring that support is informed by real-world evidence and responsive to the needs of all young people. For more information, about the EBPU please visit this website or contact epbu@annafreud.org.

Acknowledgements

Many thanks to all the schools that participated in the More Good Days at School evaluation.

The executive summary is authored by the Youth Endowment Fund. The rest of the report is authored by the evaluator.

Executive summary

The project

More Good Days at School (MGDAS) aims to use whole-school trauma-informed training and support for staff to improve children's behaviour and wellbeing. Trauma happens when someone experiences something deeply distressing or harmful that has lasting effects on their wellbeing. Trauma-informed practice means shaping services to recognise this impact and avoid causing further harm. MGDAS aims to equip teachers to meet the needs of children who have experienced trauma and hopes that by doing so, children's behaviour will improve. Central to this approach is ensuring children develop at least one trusting relationship with a safe, reliable adult at school. MGDAS also encourages an approach to behaviour management that frames challenging behaviour as an inability to self-regulate and sanctions as an opportunity to learn from mistakes.

Delivered by Knowledge Change Action and Warren Larkin Associates, MGDAS offers two stages of training to all school staff (including teaching staff, leadership teams, and pastoral and inclusion (PI) leads) over 15 months. Stage one, collaborative enquiry, involves focus groups with staff and surveys with children to inform and co-design school training priorities. Stage two, training and reflective practice, delivered in person and online, provides school staff with three taught modules. The first is two hours of training on the brain, attachment and child development. The second includes two hours on behavioural manifestations of trauma and how to respond. The third offers three hours of emotion-coaching training. Senior leadership teams (SLTs), pastoral and inclusion (PI) leads also receive three 3-hour modules: supporting the mental health of school staff, therapeutic techniques and reflective practice. SLTs also attend four and a half hours of workshops on reflective practice and action planning, followed by three hours on amending school policies. PI leads are offered four and a half hours of additional reflective practice workshops on working with vulnerable children. In this project, the programme ran from November 2023 to July 2025.

The Youth Endowment Fund (YEF) and the Home Office funded an efficacy randomised controlled trial (RCT) of MGDAS. It aimed to establish whether MGDAS reduced children's externalising difficulties as measured by the Strengths and Difficulties Questionnaire (SDQ). It also aimed to establish whether the programme improved mental health, prosocial behaviours, social connections with teachers and peers, school inclusion, bullying, attendance and exclusion. It also assessed the impact on teacher-level outcomes, including knowledge and awareness of trauma, confidence in working with traumatised children, and staff wellbeing. The trial was run as a two-armed cluster RCT in 62 schools (13,900 children). All schools continued with existing wellbeing support practices (e.g. a worry box/drop-ins for advice, counselling for staff and mental health policies), and half of the schools (31) were randomised to additionally receive the training and support. The evaluation also included an implementation and process evaluation that explored whether MGDAS was implemented as intended and examined participant perceptions. This included interviews or focus groups with 21 children, 23 school staff members and nine trainers and a school survey. 60% of children in the trial were White; 25% Asian or Asian British; 6% Black, Black British, Caribbean or African; 5% from mixed or multiple ethnic groups; and 4% identified as from another ethnic group.

Table 1: Key conclusions
MGDAS had **no impact** on children's externalising behaviours. After the programme, children in MGDAS schools had similar levels of externalising behaviours compared to their counterparts in the control group. This result has a **very low** security rating.
MGDAS demonstrated no impact on prosocial behaviour, bullying perpetration, children's perceptions of the impact of their emotional or behavioural difficulties, their perceived relationships with teachers, their perception of peer support, and their perception of the relevance of schoolwork and their sense of agency over it. It had a small negative impact on children's internalising behaviour, school attendance, suspensions and likelihood of being a victim of bullying. It had a small positive impact on children's aspirations and goals. These are secondary outcomes and should be treated with more caution. There is also statistical uncertainty regarding these outcomes.
MGDAS demonstrated a moderate positive impact on teachers' understanding of the underlying causes of behaviour, their confidence in using trauma-informed practices, the extent to which responses to behaviour used trauma-informed practice, teachers' empathy and on reducing teacher burnout. Teachers in MGDAS schools were

also slightly more likely to report fulfilment from helping others and have positive feelings about their roles. There was no impact on teacher stress from exposure to trauma. These are secondary outcomes and should be treated with more caution, and all of these outcomes are statistically uncertain.
The trial suffered from very high attrition. 49.9% of children in the trial at randomisation were not included in the final analysis. All impact findings should therefore be treated with significant caution.
Teachers' attendance at training was low. On average, teachers attended only 29% of sessions. High workload, competing pressures, and long and frequent sessions impacted staff attendance. Some staff reported feeling better equipped to support children, while others described a lack of change in culture, policies and practices.

YEF security rating

These findings have a **very low** security rating. The trial was designed to be large enough to detect meaningful impacts on the primary outcome. However, 49.9% of children in the trial at randomisation were not included in the final analysis. We do not know if the effect found for MGDAS would be the same if the children missing from the final analysis were included.

Interpretation

MGDAS had **no impact** on children's externalising behaviours. This result has a **very low security rating** due to attrition. Attrition was caused by school dropout, pupils leaving schools and challenges in engaging schools to complete endline measures. School dropout was primarily due to staffing and capacity issues. MGDAS demonstrated small and highly uncertain impacts on a range of pupil-level outcomes. While there is some uncertainty in the estimate, the likelihood of children being victims of bullying was slightly higher in the intervention group. For staff outcomes, the direction of effects generally suggested greater alignment with trauma-informed practice in intervention schools. However, these differences were statistically uncertain and also consistent with there being no impact.

22 out of 25 MGDAS schools for which data was collected delivered all intended sessions. Staff members who attended generally enjoyed the sessions, and most were able to recall key concepts. However, staff attendance was limited (29% across all sessions). For the whole-school sessions, only one school achieved 80% session attendance. In interviews, staff frequently cited the pressurised school context and high workload as barriers to engagement, and they perceived sessions to be too long and frequent. Smaller group sessions, such as with SLTs and PI staff, were better attended. Low attendance may not explain the lack of impact on the primary outcome. In those schools where attendance was higher, MGDAS had a more negative impact on children's externalising behaviours (although this finding is underpowered, uncertain and should be treated with significant caution). The small number of children interviewed in MGDAS schools reported more frequent check-ins from staff about pupil wellbeing, new wellbeing initiatives and a shift away from detention as an immediate punishment. The small number of staff interviewed reported feeling better equipped to help pupils and having improved connections with pupils, and some reported general changes to ways of working (including reducing their reliance on traditional sanctions). Some staff, however, described a lack of change in culture, policies and practices following the training.

There has previously been very little research on the impact of trauma-informed training on children's outcomes. This evaluation finds that while it may lead to changes in teacher understanding of trauma informed practice, implementation challenges and a potential lack of pupil-level impact may question its use for the purpose of reducing children's involvement in violence. Given the relatively short period of evaluation, longer-term effects may not have been captured. This report and the primary and secondary outcome findings only present the findings of one study. When considering the implications, frontline professionals, policymakers and service commissioners should carefully consider the process evaluation, the wider evidence base and their own professional judgement.

Summary of impact

Outcome	Effect size	Impact	Evidence security	No. of children	P -value
SDQ externalising difficulties	0.014 [-0.047, 0.076]	No impact	1 of 5	6,959	0.652

Introduction

Background

Exposure to potentially traumatic events is a common experience among young people, with more than one in three encountering such events by the age of 18 (Lewis et al., 2019). These experiences can include experiencing abuse, being exposed to neglect, witnessing violence, or facing other adverse childhood experiences (ACEs), which have been shown to increase the risk of mental health issues, behavioural challenges, social problems and learning difficulties (Baglivio et al., 2020; Felitti et al., 1998; Neil et al., 2022). It has been shown that a greater number of ACEs leads to a higher likelihood of adverse behavioural, educational and social outcomes, including involvement in crime and violence (Gaffney et al., 2021).

Trauma can induce enduring changes in brain structure and function, including a heightened threat response system, abnormal hormonal regulation, altered reward processing and diminished executive functioning, all of which contribute to difficulties with emotional regulation, decision-making and behaviour (Blankenstein et al., 2022; McCrory et al., 2017; Van Voorhees & Scarpa, 2004). In schools, traumatised children often display behavioural and relational difficulties that are sometimes labelled as conduct disorders or antisocial behaviour. Punitive approaches, including zero-tolerance policies and exclusion, often exacerbate these problems and can contribute to cycles of school disengagement and later involvement with the criminal justice system (IPPR, 2017; Rainer et al., 2023).

Trauma-informed practice has emerged as an organisational-level, preventative approach aimed at addressing the impacts of trauma across settings such as schools, child welfare, healthcare and criminal justice (Gaffney et al., 2021). Unlike trauma-specific interventions, such as trauma-focused cognitive behavioural therapy, which directly support individual recovery, trauma-informed practice seeks to embed trauma awareness and responsive policies into organisational culture, practices and structures to prevent re-traumatisation (Gaffney et al., 2021). Key assumptions include realising the impact of trauma, recognising trauma symptoms, integrating trauma knowledge into practice and preventing re-traumatisation. Core principles include safety; trustworthiness and transparency; peer support; collaboration and mutuality; empowerment and voice; and attention to cultural, historical and gender issues (SAMHSA, 2014; Lang et al., 2015; Maynard et al., 2019).

Implementation of trauma-informed practice typically occurs across three domains: workforce development (training staff to respond effectively to trauma), trauma-informed services (screening for trauma, offering trauma-specific supports) and organisational change (modifying policies and practices to reduce re-traumatisation and improve coordination) (Gaffney et al., 2021). However, it is important to note that trauma-informed practice varies considerably in scope, ranging from one-off training sessions to comprehensive whole-school approaches (Gaffney et al., 2021). Effective implementation usually depends on senior leadership commitment, a whole-school approach, a culture of support, adequate staff training, listening to children and families, and interagency information sharing. Barriers to effective implementation of trauma-informed practice have been identified as staff capacity, turnover and limited ongoing support (Gaffney et al., 2021).

The last ten years have seen the growth of a bottom-up movement within schools in the UK, variously referred to as ”attachment-aware” or “trauma-informed”, with schools taking steps to recognise that a proportion of their pupils will be contending with a legacy of trauma or unmet attachment needs (Tivoli, 2024). This has been accompanied by growing recognition in areas of devolved national policy in Scotland, Wales and Northern Ireland and in the form of local initiatives in England (Gillett-Swan & Lundy, 2021).

Multiple local authorities (LAs) have established programmes to support schools in implementing trauma-informed approaches, including Derbyshire, Islington and Kirklees (Tivoli, 2024).

Despite their growing adoption, evidence on the effectiveness of trauma-informed interventions remains limited and mixed. The broad scope of interventions, variation in models and complexity of measuring organisational and pupil-level outcomes contribute to the uncertainty (Gaffney et al, 2021). Nevertheless, some studies indicate that whole-school interventions incorporating staff training, organisational change, practice adjustments and targeted pupil support can improve understanding of trauma, reduce re-traumatisation and contribute to positive educational and wellbeing outcomes (Avery et al., 2020). However, these findings come from a relatively small and methodologically mixed evidence base, and they should therefore be interpreted with appropriate caution. This is reflected in the Youth Endowment Fund's (YEF's) Toolkit rating for trauma-informed practice, which currently lists its impact as "unknown" with a confidence level of 0, underscoring the need for further high-quality evaluations in this area (Gaffney et al., 2021).

Given the high prevalence and severe consequences of trauma, the disproportionate impact on marginalised groups and the potential for schools to either exacerbate or mitigate trauma, there is a clear need for robust evaluation of trauma-informed interventions. Research demonstrates that traumatic experiences disproportionately affect minority youth, including historical and generational trauma, immigration and acculturation stressors, experiences of discrimination, family violence, and community violence, with minority youth less likely to access mental health care despite higher trauma exposure (Pumariage et al., 2022). In the UK, structural barriers manifest in stark educational disparities, such as White Gypsy and Roma pupils experiencing suspension rates three times the national average (3,371 per 10,000 pupils versus 1,106 the suspension rate for White British pupils), while Black Caribbean pupils are also excluded at disproportionately high levels (Department for Education [DfE], 2024a). Exclusion itself is a significant risk factor for later involvement in serious violence, with evidence showing that children excluded from school are more likely to be drawn into offending and criminal justice contact (YEF, 2025). At the same time, Black children are not only over-represented among those arrested and in custody but also disproportionately likely to be victims of violence, highlighting that minoritised groups face a dual disadvantage of being more likely to be excluded from education and more likely to be affected by crime and violence (YEF, 2025). Recent analysis of the Millennium Cohort Study found that Black children had higher assault perpetration rates (53.3%) than White children (41.7%). However, this difference disappeared after controlling for socioeconomic characteristics and exposure to ACEs and positive childhood experiences (PCEs), suggesting ethnic disparities in violence are substantially explained by differential exposure to structural inequalities rather than inherent group differences (Villadsen et al., 2025).

The relationship between ACEs, PCEs and youth violence underscores schools' potential as sites of intervention. Having six or more ACEs increased the risk of assault perpetration by 45%, weapon involvement by 150% and gang involvement by 154% (Villadsen et al., 2025). Crucially, PCEs can mitigate these risks. Among children with three or more ACEs, those with five or more PCEs showed a 22% reduction in the risk of assault perpetration, a 49% reduction in weapons involvement and a 39% reduction in gang involvement (Villadsen et al., 2025). However, despite widespread adoption of trauma-informed approaches, there is considerable variability in definition and implementation, with relatively little known about benefits, costs and effectiveness (Berliner & Kolko, 2016; Watson et al., 2025). The combination of substantial investment, rapid proliferation, implementation variability and limited rigorous evaluation creates an urgent imperative for high-quality trials that examine not only effectiveness but also if trauma-informed approaches reduce or inadvertently perpetuate existing inequities across ethnic and

socioeconomic groups. The current trial examines the More Good Days at School (MGDAS) intervention – a whole-school trauma-informed approach – to determine its impact on pupils' social, emotional, and behavioural outcomes.

Overview of evaluation

The overall design of the study was a cluster (school) randomised controlled efficacy trial, with a nested mixed methods convergence design and qualitative-driven implementation process evaluation (IPE). The trial involved two arms: MGDAS and business as usual (treatment arm) and business as usual only (control arm). To maximise the recruitment of schools, randomisation took place at two time points: one in November 2023 (cohort 1) and one in February 2024 (cohort 2) and occurred on a school-level basis. Randomisation was stratified by local authority (LA) or combined authority (CA), with a minimum of two schools needed for an LA/CA to be included in the randomisation. A CA was defined as any area that, at the time of trial set-up, had a formally constituted statutory CA under UK Government legislation (e.g. Greater Manchester, West Yorkshire). Schools located within these statutory CAs were grouped into a single CA stratum; all other schools were stratified at the individual LA level. The primary outcome measure was externalising difficulties measured by the sum score of the conduct problems and hyperactivity subscales of the Strengths and Difficulties Questionnaire (SDQ) (Goodman et al., 1998).

The IPE involved a support description survey (SDS) collected twice from all schools, interviews with pupils and interviews or focus groups with school staff members, as well as intervention implementers. The MGDAS implementers were members of the delivery team (e.g. programme managers, senior leadership, associates and trainers) who oversaw programme management and delivery and who also acted as the trainers for participating schools. Lastly, implementation and monitoring data collected by the MGDAS programme team was used to examine the number and types of sessions provided to schools and the levels of engagement of school staff in these sessions.

Intervention

The intervention, MGDAS: Building relationships to promote health, happiness and learning, was delivered between January 2024 and June 2025 and is detailed below.

Table 3: Intervention description.

Item	Description
Brief name	More Good Days At School (MGDAS): Building relationships to promote health, happiness and learning.
Aims	The aim of MGDAS was to interrupt the school-to-offending/prison pipeline and reduce the impact of adverse experiences – and the associated reduced social relationships and support – through increasing protective factors and school engagement. To achieve this, the programme promoted an environment of emotional and relational safety to reduce pupils' stress responses and associated avoidance or externalising behaviours and to enhance their capacity to learn and develop a prosocial identity. Central to this approach was the development of at least one trusting relationship with a safe, reliable adult within the school setting, which was key to fostering emotional and relational safety. For school staff to be able to create such an environment, it was essential that they also experienced emotional and relational safety themselves. MGDAS aims to support the wellbeing of school staff through learning and development activities,

	focusing on the impact of collective trauma, while also building the school's capacity to be reflective and supportive of its staff. Overall, by aiming to equip staff with trauma-informed knowledge, emotion-coaching skills and reflective practice skills, the programme aims to enable staff to support pupils' emotional regulation, engagement and learning in their day-to-day interactions.
Who	The intervention was delivered across the whole school by a specialist team with expertise in trauma-informed practice, organisational development and school-based mental health. The programme primarily targeted school staff, who were expected to implement changes at both the classroom and organisational levels. This included all staff members, senior leadership teams (SLTs), and pastoral and inclusion (PI) leads, reflecting the whole-school focus of the intervention. Although the pupils were not direct participants in the training, the intended beneficiaries were all children and young people attending the schools, as changes in school culture, relationships and behaviour approaches were designed to influence their educational and wellbeing outcomes.
Content	The intervention was a universal whole-school approach. The content focused on training and capacity building for all school staff in four areas: 1. Understanding the impact of individual and collective adverse experiences and trauma 2. Understanding the importance of the staff–pupil relationship as being foundational to helping pupils address barriers to learning and increase engagement 3. Utilising relational rather than punitive approaches to address challenging or dysregulated behaviour 4. Utilising tools, techniques and strategies to look after one's own mental health and that of others (e.g. pupils, colleagues) The programme was delivered during school time, and there were two main stages of the programme. Stage one was a collaborative enquiry. It involved focus groups with staff (1–2 hours with three attendees on average) and surveys with pupils to identify existing knowledge and practice, explore strengths and challenges, and inform and co-design priorities for enabling the school to become trauma-informed. This stage was also an opportunity for the delivery team to form relationships and model the trauma-informed approach with school staff by using, for example, collaborative, empowering and asset-based approaches. Stage two was training and reflective practice, delivered in person and online. Training was tailored to different audiences, and it was supported by a learning platform with resources (e.g. further e-learning courses, journals, workbooks). School staff received three taught modules: 1. Five to thrive: 2-hour training on the brain, attachment and child development 2. Mending hurts: 2-hour training on behavioural manifestations of trauma and stress and ways of responding to these manifestations to promote emotional and relational safety 3. Relationships, trauma and the brain: 3-hour emotion training for responding to dysregulated behaviour (e.g. recognising, validating and labelling feelings;

	setting limits on behaviour; problem solving to support emotional regulation skills) SLTs and PI leads received three taught modules: 1. Building and maintaining individual and community resilience: 3-hour training on being resilient, supporting one's own mental health and supporting the mental health of staff in their school 2. Therapeutic ideas for school leaders and specialist practitioners: 3-hour training on gaining fundamental knowledge and skills in therapeutic processes, using appropriate therapeutic techniques and understanding limitations of therapeutic practices for school staff 3. Developing resilient practitioners through reflective practice: 3-hour training on reflective practice techniques SLTs received two additional workshops: 1. Reflective practice consultation group: 4.5-hour training over three sessions to increase confidence and competence in reflective practice and to build an action plan on becoming trauma-informed 2. Review of policies, processes and procedures: 3-hour training over two sessions to first review, update and possibly re-write behaviour and attendance management policies; consider reward structures; and discuss emotional and mental wellbeing support processes. Then, in the follow-up session, senior leads review progress in implementing changes to policies, processes and procedures. PI leads received an additional workshop: Reflective practice: 4.5-hour training over three sessions to explore the impact of working with vulnerable pupils and families and how (if at all) this has changed over the course of the programme. SLTs and PI leads were able to use the online learning platform to curate their own directory of locally relevant information, resources and support. In particular, the directory aimed to provide information and signposting for wraparound support for vulnerable pupils and their families. This dynamic directory could then be accessed by all school staff.
When and how much (duration and dosage of the intervention)	The taught whole-school elements totalled seven hours, whereas the elements targeted at senior staff totalled an additional eight hours of taught material and between 7.5 and 12 hours of workshops. Sessions for schools in the first cohort were delivered between January 2024 and March 2025. Sessions for schools in the second cohort were delivered between April 2024 and June 2025. The whole-school and senior staff elements were delivered in parallel. Though there were some variations, in most schools, both the whole-school and senior staff elements were delivered in the order they are presented in above, although sequencing was, at times, amended due to the availability of schools.
Inclusivity	The programme was designed to be sensitive to different racial and ethnic groups by recognising how adversity and exclusion are shaped by social and cultural context. It fostered practice that was attentive to race, culture and inequity through reflective activities on identity, bias and systemic factors and by emphasising cultural atonement in trauma responses and behaviour. Delivery was adapted to reflect the

	ethnic and faith backgrounds of school communities, including tailored imagery and scheduling considerations, ensuring relevance and inclusivity.
Providers	The delivery team comprised staff from two organisations, Knowledge Change Action (KCA) and Warren Larkin Associates (WLA), from practitioner and research backgrounds, including clinical psychology, criminal justice, education, homelessness, nursing and social work. WLA staff delivered most sessions, including Therapeutic Ideas, Focus Group, and Reflective Practice Training sessions, as well as some Reflective Practice Group Consultation sessions. All other sessions were delivered by KCA.
Training and quality assurance of providers	A core group of trainers with an education or public health background were selected to deliver the programme. All trainers delivering the programme were experienced practitioners, educators or clinicians who underwent a structured preparation process prior to delivery. This included: • **Familiarisation with programme content and objectives:** Trainers completed a detailed walkthrough of the curriculum, including the theoretical model underpinning the work, session-by-session intentions and required facilitation approaches. • **Programme-specific briefing sessions:** Trainers from KCA and WLA attended joint orientation meetings to ensure a consistent understanding of aims, target audience, delivery methodology, safeguarding expectations and reporting requirements. • **Standardised delivery resources:** Trainers were issued session plans, slides, handouts and delivery notes to ensure a coherent and consistent learning experience across both cohorts and all schools. Quality assurance was ensured through multiple mechanisms: • **Pre-delivery checks:** All training materials and plans were reviewed by KCA's Director of Learning and Director of Strategy and Partnerships (overseeing the programme) and Dr Warren Larkin to ensure alignment with the programme design and fidelity to the intended learning outcomes. • **Ongoing supervision and support:** Trainers had access to clinical/reflective supervision (as appropriate to their discipline) and operational support throughout delivery to troubleshoot challenges and maintain quality. • **Observation and feedback loops:** A selection of recorded sessions was observed by programme leads, and trainers received structured feedback focused on fidelity, engagement and effectiveness. • **Participant feedback:** Post-session evaluations were collected and reviewed to monitor trainer performance and inform continuous improvement. • **Internal review meetings:** KCA and WLA held debriefs to discuss delivery quality, emerging themes and any adjustments required to maintain high standards. Together, these steps ensured that all trainers were well prepared and that delivery remained consistent, safe and aligned to the programme's intended outcomes.
Completion and Delivery	Dates of the delivery were between January 2024 and March 2025 for cohort 1 and between April 2024 and June 2025 for cohort 2. Overall, implementation coverage was high, with almost all intervention schools receiving the full set of planned sessions, aside from a reduced number of reflective practice sessions. Three schools missed one or two sessions due to limited engagement with the programme, and in some other cases, sessions were recorded when attendance was low to ensure staff still had access to the material. Delivery was primarily online, with a small number of in-person sessions across a few schools. Although implementation coverage was high, overall attendance was low.

	Schools showed some flexibility in how sessions were scheduled. Several schools split whole-staff sessions across multiple slots, and some requested slightly shorter sessions. Reflective consultation groups varied by school, with some holding separate groups for different staff roles and others running combined sessions. A small number of schools received additional catch-up or Key Concepts Consultancy sessions for senior staff who had joined their role partway through delivery. Minor adaptations to session ordering also occurred in a few cases. Further details on session completion, adaptations, attendance patterns and delivery mode are provided in the Results section.

Theory of change

Prior to the evaluation, a logic model was co-developed with the delivery team (see Appendix D). The MGDAS programme is based on the understanding that children and young people affected by adverse experiences and trauma are more likely to experience difficulties with emotional regulation, relationships and learning within school environments. To address these challenges, the programme aimed to improve educational inclusion by promoting an environment of emotional and relational safety within schools. Creating such an environment was expected to reduce pupils' stress responses, which could otherwise lead to avoidance or externalising behaviours, and to increase their capacity to engage in learning and develop a prosocial identity. The intention was to provide benefits to all staff and pupils, not limited to those children with a history of ACEs.

Central to the programme theory was the premise that developing at least one good relationship with a safe, reliable adult in the school setting is key to building emotional and relational safety for pupils. The intervention is intended to enhance staff capacity to recognise and respond to pupils' needs and to model positive relational practices. Through these strengthened skills and behaviours, staff members are more likely to engage consistently and supportively with pupils, creating the conditions in which at least one meaningful, trusting relationship can form. For school staff members to enable such relationships, they also needed to experience emotional and relational safety themselves. Accordingly, the MGDAS programme supported staff wellbeing through learning and development on the impact of trauma and adversity and through fostering a reflective, supportive school culture. This approach aimed to strengthen adults' capacity to respond to the needs of pupils and to model positive relational practices. By improving staff understanding of trauma and by embedding reflective and supportive structures within schools, MGDAS sought to build the capacity of school systems to nurture both staff and pupil wellbeing. In turn, this was expected to decrease pupils' externalising difficulties, mental health difficulties and involvement in bullying and improve friendships and school engagement (i.e. increase attendance and decrease exclusions). The programme theory recognised that changes would occur over different timeframes: improvements in staff awareness and practice were expected relatively quickly, while changes in pupil outcomes were likely to emerge more gradually. Given the relatively short period between baseline and follow-up, the evaluation may have been constrained in its ability to capture these longer-term effects. Moreover, the multi-component nature of trauma-informed practice programmes makes it challenging to define what constitutes meaningful change and to determine what effects are reasonable to expect within a given timeframe.

Evaluation objectives

The overarching research question was: How effective is the implementation of a whole-school approach to trauma-informed practice, including senior leadership support, frontline practitioner training and ongoing reflective practice (context), in improving safe social connections (mechanism) and reducing externalising difficulties (primary outcome), thereby reducing the likelihood of young people becoming involved in crime and violence in the future (long-term outcome)?

The primary research question was: Is there a difference in mean externalising difficulties at the end of programme follow-up between pupils in secondary schools receiving MGDAS and business as usual compared to pupils in secondary schools receiving business as usual only?

The secondary research questions for pupils were:

1. Is there a difference in mean mental health difficulties at the end of programme follow-up between pupils in secondary schools receiving MGDAS and business as usual compared to pupils in secondary schools receiving business as usual only?
2. Is there a difference in mean prosocial behaviours at the end of programme follow-up between pupils in secondary schools receiving MGDAS and business as usual compared to pupils in secondary schools receiving business as usual only?
3. Is there a difference in mean safe social connections with teachers at the end of programme follow-up between pupils in secondary schools receiving MGDAS and business as usual compared to pupils in secondary schools receiving business as usual only?
4. Is there a difference in mean safe social connections with peers at the end of programme follow-up between pupils in secondary schools receiving MGDAS and business as usual compared to pupils in secondary schools receiving business as usual only?
5. Is there a difference in mean school inclusion at the end of programme follow-up between pupils in secondary schools receiving MGDAS and business as usual compared to pupils in secondary schools receiving business as usual only?
6. Is there a difference in mean bullying (perpetration and victimisation) at the end of programme follow-up between pupils in secondary schools receiving MGDAS and business as usual compared to pupils in secondary schools receiving business as usual only?
7. Is there a difference in educational attendance at the end of programme follow-up between pupils in secondary schools receiving MGDAS and business as usual compared to pupils in secondary schools receiving business as usual only?
8. Is there a difference in exclusions (fixed-term suspensions and permanent) at the end of programme follow-up between pupils in secondary schools receiving MGDAS and business as usual compared to pupils in secondary schools receiving business as usual only?

The hypotheses state that pupils in schools that participate in the intervention will show reductions in mean levels of externalising difficulties, overall mental health difficulties, experiences of bullying victimisation, bullying perpetration, and the incidence of suspensions and permanent exclusions. At the same time, these pupils are expected to show increases in mean levels of prosocial behaviours, the development of safe and supportive social connections and overall school attendance. As the intervention is delivered at the staff level, it is anticipated that effects on pupil outcomes will be modest in magnitude. Stronger effects are expected in staff-related measures, which are more directly targeted by the programme.

The secondary research questions for school staff were:

1. Is there a difference in mean knowledge and awareness of trauma at the end of programme follow-up between staff in secondary schools receiving MGDAS and business as usual compared to staff in secondary schools receiving business as usual only?
2. Is there a difference in mean confidence in working with young people who have experienced trauma at the end of programme follow-up between staff in secondary schools receiving MGDAS and business as usual compared to staff in secondary schools receiving business as usual only?
3. Is there a difference in mean wellbeing at the end of programme follow-up between staff in secondary schools receiving MGDAS and business as usual compared to staff in secondary schools receiving business as usual only?
4. Is there a difference in mean burnout at the end of programme follow-up between staff in secondary schools receiving MGDAS and business as usual compared to staff in secondary schools receiving business as usual only?
5. Is there a difference in mean ability to create an emotionally safe environment at the end of programme follow-up between staff in secondary schools receiving MGDAS and business as usual compared to staff in secondary schools receiving business as usual only?
6. Is there a difference in mean understanding of vicarious trauma and the need for staff support at the end of programme follow-up between staff in secondary schools receiving MGDAS and business as usual compared to staff in secondary schools receiving business as usual only?
7. Is there a difference in mean empathy-focused behaviours for behaviour management at the end of programme follow-up between staff in secondary schools receiving MGDAS and business as usual compared to staff in secondary schools receiving business as usual only?

For staff in schools that participate in the intervention, outcomes are expected to include increases in mean scores relating to knowledge and awareness of trauma, confidence in working effectively with young people who have experienced trauma and overall wellbeing. Staff are also expected to demonstrate greater ability to create emotionally safe environments, enhanced understanding of vicarious trauma and the importance of staff support, and more frequent use of empathy-focused approaches to behaviour management. In addition, mean burnout scores are expected to decrease.

IPE research questions were:

1. To what extent is the MGDAS programme implemented as intended?
2. Are some elements of the MGDAS programme implemented more successfully than other elements?
3. What are the facilitators of and barriers to the implementation of the MGDAS programme?
4. What changes to policy and practice do school staff implement following the MGDAS programme?
5. What are the facilitators of and barriers to changing policy and practice for school staff following the MGDAS programme?
6. What is the perceived need for the intervention amongst MGDAS implementers and school staff?
7. What is the experience of the MGDAS programme for implementers and school staff?
8. What is the experience of pupils of support from schools following the MGDAS programme, and do different groups of pupils (i.e. ethnic groups, eligibility for free school meals [FSMs], special educational needs and/or disabilities [SEND] and/or an Educational, Health and Care Plan [EHCP]) have different experiences of support?
9. To what extent, and how, does the MGDAS programme impact school culture?

The link to the published protocol can be found here.

The Statistical Analysis Plan can be found here.

Ethics and trial registration

The project was reviewed and approved by the University College London's (UCL) Research Ethics Committee (Project ID Number: 14037/012). The trial was registered with ISRCTN (registration number: 12074782; DOI: https://doi.org/10.1186/ISRCTN12074782). The review confirmed that the study design, recruitment procedures, consent processes and data management arrangements met ethical standards for research with minors and school staff. Recruitment was led by the delivery team, and agreement for schools to participate in the trial was through the signing of the Memorandum of Understanding (see Appendix A). Informed consent to participate in the trial was obtained, on an individual level, for all participants. Schools distributed information sheets and privacy notices to staff, Year 8 and 9 pupils, and parents/carers of Year 8 and 9 pupils (see Appendix A). When schools were enrolled in the trial, this started a two-week opt-out period for parents/carers to opt their child out of participating in the trial. After this period, Anna Freud securely informed schools of opted-out pupils and emphasised that these pupils were not to take part in the survey data collection. Pupils and staff were invited to participate in the survey data collection and were presented with information sheets and privacy notices pertaining to the study. Consent (and assent for pupils under 16 years old) to participate was recorded prior to the start of the survey.

Prior to sharing school-held data on participating pupils (e.g. demographic information and attendance data), schools signed a Data Sharing Agreement formalising the arrangements for sharing, holding and processing the data. Pupils whose parents or carers opted them out of the trial, or who joined the school after the opt-out period, were not included in these submissions.

An adverse events monitoring group was established to capture safeguarding issues that arose during the study. The designated MGDAS contact in each school, across both intervention and control arms, completed a secure online form if any safeguarding concerns occurred. The form recorded the category of the event (for example, self-harm, suicide attempt or risk to others) and whether it was perceived to be related to participation in the study, along with a brief rationale.

Data protection

The lawful basis for data processing under the General Data Protection Regulation (GDPR) was legitimate interest. The processing of data was expected to deliver social benefits by improving understanding of how to conduct this type of research and by informing the development of more effective services for young people at risk of youth violence while presenting only a minimal privacy impact on individuals. Under GDPR Article 6(1)(f), processing is permitted where it is necessary for the purposes of the legitimate interests pursued by the data controller and under Article 9(2)(j) for statistical and research purposes.

Parents/carers, pupils and staff were informed of the scope of the research through information sheets and privacy notices. These documents outlined the purpose of data processing, the lawful basis for processing, the methods of data sharing and storage, and the duration of data retention (see Appendix A).

To ensure participant confidentiality, all personal data was collected and stored securely within UCL's Data Safe Haven, a secure server environment accessible only to authorised members of the evaluation team. During the evaluation, the research team processed identifiable data because personal identifiers were required to match pupil responses across time points (e.g. baseline to endline). Once the matching process

was complete, the dataset used for analysis was pseudonymised by removing direct identifiers and replacing them with an internal study ID. All statistical analyses were conducted on this pseudonymised dataset.

The participant information sheet also explained that, following the completion of the evaluation, identifiable pupil data (name, unique pupil number and date of birth) would be transferred to the DfE. At this stage, personal identifiers would be deleted and replaced with a unique pupil matching reference number. The Anna Freud Centre acted as data controller until the transfer to DfE. Subsequently, the DfE would provide these pseudonymised data to the Office for National Statistics Secure Research Service, where they would be stored securely to facilitate future research exploring the long-term outcomes of interventions such as MGDAS. The YEF would become the data controller once the data was archived.

Project team/stakeholders

Evaluation team – Evidence Based Practice Unit, Anna Freud and University College London

- Prof Julian Edbrooke-Childs (Professor of Evidence Based Child and Adolescent Mental Health; Head of Evaluation) served as Principal Investigator and provided overall leadership and management of the project, including oversight of budget, risks and timelines. He acted as the strategic point of contact for the YEF and KCA/WLA.
- Georgina Mutton (Membership and Development Officer) provided operational support for all communication activities, primarily with schools.
- Dr Emily Stapley (Senior Research Fellow) oversaw the design and delivery of the IPE. Dr Emily McDougal provided cover during maternity leave.
- Ayesha Sheikh (Research Officer) supported the operational conduct and delivery of the research and helped with qualitative analysis.
- Angelika Labno (Research Officer) supported the operational conduct and delivery of the research.
- Navya Malik (Research Officer) supported the operational conduct and delivery of the research.
- Dr Emily Goodacre (Research Officer) supported the operational conduct and delivery of the research.
- Dr Suzet Tanya Lereya (Senior Research Fellow) led the quantitative side of the trial, overseeing the design and delivery of the quantitative research.
- Dr Jessica Stepanous (Research Fellow) supported project management, prepared the statistical analysis plan and conducted quantitative data analysis.
- Ben Ritchie (Informatics Lead) led the collection and management of local school data.
- Arthur Pander Maat (Research Officer) worked on the collection and management of local school data and led the production of school reports.
- Holly Rowland (Research Officer) supported the operational conduct and delivery of research activities focused on local school data.
- Millie Evans (Participation Programme Assistant) provided lived experience guidance from a young person's perspective and contributed to reviewing study processes, documentation, analysis and reporting.
- Saya Majid (Researcher Intern) contributed to quantitative and qualitative analysis and contributed to writing the findings.
- Rachel Hart (Information Governance Manager) oversaw and supported adherence to best practice in information governance.
- Bernadette Martin (Head of Participation) oversaw and supported the peer research work, ensuring that principles of co-production were considered and addressed throughout the project.

Steering group – Anna Freud and University College London

The research steering group comprised key stakeholders who, together, guided the direction of the project, offering expertise, feedback and oversight to ensure that the research is relevant, rigorous and ethically sound. The research steering group met monthly during the study.

- Prof. Jessica Deighton (Professor of Child Mental Health and Wellbeing; Director of Applied Research and Evaluation) provided critical feedback to the overall design, analysis and interpretation of the evaluation through the project's steering group.
- Nick Tait (Programme Manager) oversaw the engagement strategy and ensured bidirectional communication between schools and the evaluation team.
- Charli Atkinson-Ryan (EDI Manager) supported the project to ensure that equity, diversity and inclusion were consistently considered and addressed.
- Dr Laura Talbot (Joint AMBIT Lead) oversaw the project's safeguarding policy and worked with researchers to prepare for and respond to safeguarding issues.
- Prof. Sam Norton (Professor of Medical Statistics & Applied Health Research) provided expert input to the statistical analysis plan.
- Prof. Peter Fonagy (Head of the Division of Psychology and the Language Sciences at UCL) provided senior scrutiny of the methodological and intellectual conduct of the research.

Additionally, the Data Monitoring and Ethics Committee and Trial Steering Committee were established, and both groups were consulted following baseline data collection as part of the study conduct phase.

Delivery team

- Anisha Gadhia (Chief Executive, Legal and Criminal Justice Senior Leader and Practitioner, KCA) provided programme oversight.
- Lou Mee (Programme Manager, KCA) coordinated the delivery of the programme.
- Catherine Gordon (Director of Learning, KCA) acted as the quality assurance lead, providing trainer and associate supervision, trainer-consultant support, and trainer and content leadership.
- Alex Elander-Phoenix (Expert, KCA) acted as trainer-consultant.
- Jo McAndrews (Associate, KCA) acted as a trainer.
- Kate Cairns (Social worker, KCA) acted as a trainer-consultant and content lead.
- Barry Golten (expert, KCA) acted as a trainer-consultant.
- Ann Berry (Registered General Nurse and Public Health Commissioner, KCA) acted as a trainer-consultant.
- Richard Holmes (Managing Director, KCA) provided strategic oversight and helped with the collaborative enquiry.
- Dean Reilly-Sharp (Associate, KCA) acted as a trainer.
- Rachael Pryor (Associate, KCA) acted as a trainer.
- Brian Roberts (Associate, KCA) acted as a trainer.
- Sally Poskett (Associate, KCA) acted as a trainer.
- Sheila Mulvenny (Associate, KCA) acted as a trainer.
- Suzi Moore (Associate, KCA) acted as trainer.
- Dr Warren Larkin (Consultant Clinical Psychologist, WLA) provided programme oversight, designed training content, led the collaborative enquiry and acted as an associate trainer.

- Dr Colin Baker (Research Consultant, WLA) led the quantitative data analysis for the collaborative enquiry survey.
- Dr Andrew Parker (Research Consultant, WLA) led the quantitative data analysis for the collaborative enquiry focus groups.
- Rob Dickinson (Experienced Service Manager, WLA) facilitated collaborative enquiry focus groups and acted as an associate trainer. Also facilitated reflective practice groups.
- Dr Jolien Vos (Research Consultant, WLA) facilitated collaborative enquiry focus groups.
- Emily Barlow (Associate Consultant, WLA) delivered training and facilitated reflective practice groups.
- Siobhán Garrett (Associate Consultant, WLA) delivered training and facilitated reflective practice groups.
- Dr Rachel Lee (Consultant Clinical Psychologist) delivered training and facilitated reflective practice groups.
- Dr Jo Gorry (Consultant Clinical Psychologist) delivered training and facilitated reflective practice groups.
- Shelley England (Associate Consultant, WLA) facilitated reflective practice groups.

Methods

Trial design

The efficacy trial was run as a two-armed cluster randomised controlled trial (RCT) involving 62 schools. Half of the schools were randomised to receive the training and support, and half were to continue with existing wellbeing support practices (e.g. a worry box and drop-ins for advice, nurture groups, counselling for staff, mental health policies, and staff training to support pupil mental health). Schools were randomised in two cohorts: cohort 1 in November 2023 and cohort 2 in February 2024. Randomisation was stratified by LA or CA, with a minimum of two schools needed for the LA or CA to be included in the randomisation.

The primary outcome measure was externalising difficulties, as measured by the sum score of the conduct problems and hyperactivity subscales of the SDQ (Goodman et al., 1998). Secondary outcome measures were related to pupil and teacher wellbeing and school connections. All measures based on young people's self-reports were collected via online self-completion surveys at baseline, midline (approximately five months after randomisation) and endline (after training/support was complete: approximately 16 months after randomisation). Data collection periods were coordinated between cohorts so that they occurred broadly at the same stages for training/support inputs. The trial was designed and implemented using an intention-to-treat (ITT) framework, meaning that all randomised schools were retained in the trial and analysed according to their allocated condition. However, the outcome analysis necessarily relied on observed data. Schools that withdrew from the study did not provide endline outcome data, making it impossible to include them in the outcome analysis. As a result, the analysis is based on complete cases only (i.e. schools with both baseline and endline data). The model was adjusted for cohort, LA/CA, and (to adjust for baseline imbalances) baseline internalising symptoms, gender, and ethnicity.

The research team comprised individuals from diverse demographic backgrounds, which supported reflexivity throughout data collection, analysis and dissemination. Schools were recruited from networks serving varied communities to ensure representation across different contexts. Pupils and staff were asked to self-report their ethnicity, enabling us to capture identity data directly from participants. In the analysis, exploratory subgroup analyses were conducted by ethnicity to examine potential variation in outcomes. While these analyses were not powered to detect subgroup differences and should be interpreted cautiously, they reflect our commitment to transparency and inclusivity in reporting.

To support participation, all schools were offered a £1,000 operational reimbursement from the outset. This was released in stages aligned with data collection points (£250 following baseline data collection, £250 after midline and £500 at the end of the programme) in order to maintain engagement across the study period. During data collection periods, the research team was in close contact with schools. During each data collection window, the research team maintained regular contact with schools to address any difficulties they encountered. When schools reported challenges with online survey administration, a team member visited the school to facilitate data collection directly. In a small number of cases, paper copies of the surveys were provided, and completed responses were subsequently entered into the system by the evaluation team. Moreover, during the final data collection period, a small payment (£30) was provided to any staff member who supported the administration of a survey. This was offered in recognition of the extra tasks they undertook alongside their usual responsibilities. There was also reimbursement for school staff members who took part in the qualitative interviews (£25), acknowledging the time they contributed to the study. Lastly, to ensure the surveys provided direct value to participating schools, each school received a tailored survey report. These reports presented anonymised, aggregated results for their pupils and

compared them with the overall study sample. They also highlighted areas that could be prioritised for improvement.

Table 4: Trial design

<table>
<tr><td colspan="2">Trial design, including the number of arms</td><td>Two-arm cluster randomised controlled trial</td></tr>
<tr><td colspan="2">Unit of randomisation</td><td>Cluster (school)</td></tr>
<tr><td colspan="2">Stratification variables (if applicable)</td><td>Local authority or combined authority</td></tr>
<tr><td rowspan="2">Primary outcome</td><td>Variable</td><td>Externalising difficulties</td></tr>
<tr><td>Measure (instrument, scale, source)</td><td>Sum of the conduct problems and hyperactivity subscales of the Strengths and Difficulties Questionnaire (SDQ) (Goodman et al., 1998)</td></tr>
<tr><td rowspan="2">Secondary outcome(s)</td><td>Variable</td><td>Pupil-reported surveys
1) internalising difficulties, 2) impact score, 3) prosocial behaviour, 4) teacher–student relationships, 5) peer support at school, 6) control and relevance of school work, 7) future aspirations and goals, 8) bullying perpetration, 9) bullying victimisation
Local school data
10) attendance, 11) fixed-term exclusions, 12) permanent exclusions
Staff-reported surveys
13) underlying causes of problem behaviour and symptoms, 14) self-efficacy at work, 15) response to problem behaviour, 16) reactions to work, 17) on-the-job behaviour (empathy and control), 18) compassion satisfaction, 19) secondary traumatic stress, 20) burnout</td></tr>
<tr><td>Measure (instrument, scale, source)</td><td>Pupil-reported surveys
Measure A: SDQ (Goodman et al., 1998)
Outcomes: mental health difficulties, prosocial behaviour
Variables: 1) internalising difficulties, 2) impact score, 3) prosocial behaviour
Measure B: Student Engagement Instrument (Appleton et al., 2006)
Outcomes: safe social connection(s) with teachers, safe social connection(s) with peers, school inclusion
Variables: 4) teacher–student relationships, 5) peer support at school, 6) control and relevance of school work, 7) future aspirations and goals
Measure C: Illinois Bully Scale (Espelage & Holt, 2001)
Outcome: bullying
Variable: 8) bullying perpetration, 9) bullying victimisation
Local school data
Outcome: academic attendance and exclusions
Variables: 10) attendance, 11) fixed-term exclusions/suspensions, 12) permanent exclusions</td></tr>
</table>

		Staff-reported surveys *Measure A:* Attitudes Related to Trauma-Informed Care 35 (Baker et al., 2021) Outcomes: knowledge and awareness, confidence, emotionally safe environment, vicarious trauma, empathy-focused behaviours Variables: 13) underlying causes of problem behaviour and symptoms, 14) self-efficacy at work, 15) response to problem behaviour, 16) reactions to work, 17) on-the-job behaviour (empathy and control) *Measure B:* Professional Quality of Life Scale (Stamm, 2010) Outcomes: wellbeing, burnout Variables: 18) compassion satisfaction, 19) secondary traumatic stress, 20) burnout
Baseline for primary outcome	**Variable**	Externalising difficulties
	Measure (instrument, scale, source)	Sum of the conduct problems and hyperactivity subscales of the SDQ (Goodman et al., 1998)
Baseline for secondary outcome	**Variable**	All secondary outcomes were measured at baseline, apart from the SDQ impact score. For local school data, baseline data was defined as the first academic year of the trial (2023/24).
	Measure (instrument, scale, source)	Measures are the same as the Secondary outcome(s) section.

Participant selection

Recruitment to the trial was managed by the delivery team with support from the evaluation team and conducted at the school level. Randomisation was stratified by LA or CA; hence, a minimum of two schools were needed for an LA/CA to be included in the study. The LAs and CAs involved in the study were Kent, Wiltshire, Lancashire, Nottinghamshire, West Yorkshire, Greater Manchester, West Midlands, Cheshire West and Chester, Derbyshire, Herefordshire, Leicestershire, Slough, and Suffolk.

Initial engagement with nominated LA areas (where the project team already had strong relationships) began while the YEF funding application was still in progress, from late 2022 through spring/summer 2023. However, by summer 2023, it became clear that LAs alone were unlikely to secure the number of schools required within the original timeframe. To respond to this, an education marketing company was commissioned to widen outreach and broker introductions directly with schools, which led to an extension of the recruitment period.

The final recruitment timeline was as follows:

- Autumn 2022 to spring/summer 2023: Early discussions with nominated LAs to identify and engage suitable schools.
- September 2023: Recognition that recruitment numbers were unlikely to be met solely through LA routes; decision taken to adjust the approach and initiate a national call-out.

- October–December 2023: Extended recruitment phase supported by an education marketing partner to broaden reach, mediate introductions with multi-academy trusts (MATs), and secure additional school participation.
- January 2024: First cohort of participating schools commenced the initial programme inputs.

This extended timeline enabled the project team to meet the required cohort size and ensure schools were fully briefed ahead of programme delivery.

Alongside this, schools were recruited through direct outreach methods, such as emails, phone calls and headteacher meetings, as well as promotion via Anna Freud newsletters and social media channels. Online information sessions were held to explain the programme and encourage participation in the study. Once a school showed interest, targeted follow-ups were conducted within a local area to make sure an even number of schools were recruited to the study. An education marketing firm (PLMR) also supported recruitment by generating leads, which were followed up by KCA/WLA and Anna Freud to confirm participation. This process resulted in one MAT enrolling, with 18 schools across multiple LAs.

Schools were required to meet KCA/WLA's eligibility criteria before joining the programme. These included evidence of senior leadership and practitioner buy-in, stability within the school (for example, no recent serious case reviews or forthcoming organisational changes likely to disrupt engagement), the ability to work within the agreed programme timelines (September 2023–July 2025) and a commitment to releasing staff to attend the required sessions. Schools were also asked to consider the practical implications of participation, including the number of hours involved, staff availability and the need for appropriate technology to support online delivery (e.g. access to devices with Wi-Fi, camera and microphone). Importantly, schools were not eligible if more than 10% of their staff members had previously received two days or more of trauma-informed practice training or support from another provider to ensure the programme was implemented in settings where this approach was not already embedded.

All pupils in Years 8 and 9 (aged 12 to 14, Key Stage 3) in participating schools were eligible to take part in the evaluation. Although the programme was a whole-school intervention, two year groups were selected to avoid over-burdening schools with data collection. These year groups were selected because they would not have major school transition points during the duration of the study, avoiding the confounding effect and disruptions to data collection of transitions.

All school staff members, irrespective of their roles, were eligible to take part, reflecting the whole-school approach of the programme. A single point of contact was identified at each participating school and was asked to complete an SDS, which captured both business-as-usual provision and any new support implemented during the course of the trial.

Outcome measures

Pupil self-report surveys were collected at baseline, midline (approximately five months after randomisation) and endline (after training was complete, approximately 16 months after randomisation).

Primary outcome

The primary outcome was measured using the SDQ (Goodman et al., 1998). The SDQ is a brief behavioural screening questionnaire designed to assess patterns of emotional and behavioural functioning among children and adolescents aged 3–16 years. The YEF routinely uses the SDQ, where appropriate, across its projects to ensure consistency and enable comparisons between evaluations. The trial employed the self-report version of the measure, designed for 11–17-year-olds. It consists of five subscales, each containing

five items, which assess emotional symptoms, conduct problems, hyperactivity/inattention, peer problems and prosocial behaviour. Participants rate each item as not true, somewhat true or certainly true, producing subscale scores ranging from 0 to 10. Lower scores indicate more favourable outcomes for all subscales except prosocial behaviour.

The primary outcome was externalising difficulties, calculated as the sum of the conduct problems and hyperactivity subscales of the SDQ. The conduct problems subscale of the SDQ captures behaviours such as aggression, rule-breaking and difficulties with authority, while the hyperactivity subscale reflects impulsivity, restlessness and problems with concentration. Both theoretical and empirical evidence support combining the SDQ's conduct and hyperactivity subscales into an "externalising" subscale, as well as combining the emotional and peer subscales into an "internalising" subscale (Goodman et al., 2010). In the context of this study, externalising difficulties was chosen as the primary outcome, as it aligns with the ultimate goal of investigating the impact of whole-school trauma-informed practice on youth violence. Primary analyses used the baseline and endline data.

Secondary outcomes

Several pupil-level secondary outcomes were assessed to explore broader effects on wellbeing, engagement and peer relationships.

Beyond the SDQ externalising difficulties score (the primary outcome), further SDQ outcomes were included in the secondary analysis. Internalising difficulties were calculated as the sum of the "emotional symptoms" and "peer relationship problems" subscales of the SDQ (Goodman et al., 1998). The SDQ impact score was calculated by summing the five items assessing the distress caused by difficulties and the extent to which these difficulties interfered with home life, friendships, classroom learning and leisure activities. Prosocial behaviour was also measured using the SDQ. In line with standard scoring guidance, SDQ subscale scores were prorated where fewer than two items were missing, by averaging the available items and scaling to the full subscale range. If more than two items are missing, the subscale score was treated as missing.

Safe social connections and educational/school inclusion were assessed using subscales from the Student Engagement Instrument (SEI) (Appleton et al., 2006). The SEI is a brief 35-item self-report measure of cognitive and affective engagement, validated for students in Grades 6–12. It comprises six subscales: teacher–student relationships (9 items), peer support at school (6 items), family support for learning (4 items), control and relevance of schoolwork (9 items), future aspirations and goals (5 items) and intrinsic motivation (2 items). Items are rated on a four-point scale – strongly disagree, disagree, agree, strongly agree – with lower scores reflecting higher engagement. Safe social connections were captured using the "teacher–student relationships" and "peer support" subscales, while educational/school inclusion was assessed using the "control and relevance of schoolwork" and "future aspirations and goals" subscales. Bullying perpetration and victimisation were assessed using the "bully and victim" subscales of the Illinois Bully Scale (IBS) (Espelage & Holt, 2001). The scale contains 18 items divided into three subscales: the "bullying" subscale (9 items), which addresses how often a pupil engages in bullying behaviours; the "victimisation" subscale (4 items), which assesses both physical and verbal types of bullying victimisation; and the "physical fighting" subscale (5 items). Pupils were asked to indicate the extent to which, in the last 30 days, they were involved in each behaviour by answering items organised in a Likert scale: 0 = never, 1 = 1 or 2 times, 2 = 3 or 4 times, 3 = 5 or 6 times or 4 = 7 or more times. No cut-off scores are used; instead, higher scores in each scale indicate more involvement in bullying behaviours or more victimisation experiences. The IBS has shown good psychometric properties in previous contexts (Espelage & Holt, 2001).

Educational attendance and suspensions/exclusions were obtained from schools' administrative records on the individual level. Attendance was recorded as each pupil's percentage attendance for the 2023/24 and 2024/25 academic years. Suspensions and exclusions were coded as binary indicators (0 = none, 1 = at least one suspension/exclusion) for both the 2023/24 and 2024/25 academic years.

Several secondary staff outcomes were assessed to examine changes in knowledge, attitudes and wellbeing associated with the implementation of trauma-informed approaches.

The 35-item Attitudes Related to Trauma-Informed Care (ARTIC) scale (Baker et al., 2021) was used to assess five secondary outcomes: knowledge and awareness ("underlying causes of problem behaviour" subscale score), confidence ("self-efficacy at work" subscale score), emotionally safe environment ("response to problem behaviour and symptoms" subscale score), vicarious trauma ("reactions to work" subscale score) and empathy-focused behaviours ("on-the-job behaviour [empathy and control]" subscale score). The ARTIC scale assesses staff beliefs and attitudes about trauma-informed approaches and the extent to which they align with trauma-informed principles. It includes items covering domains such as beliefs about the causes of behaviour, responses to escalations, empathy and compassion, flexibility in practice and views about organisational support. Items are presented as bipolar statements anchored at each end of a 7-point Likert scale, with higher scores indicating more trauma-informed attitudes. The ARTIC scale has demonstrated good internal consistency and construct validity across education and human-service settings (Baker et al., 2021).

The 30-item Professional Quality of Life scale (ProQOL) (Stamm, 2010) was used to assess wellbeing ("compassion satisfaction" subscale score) and burnout ("secondary traumatic stress and burnout" subscale scores). The ProQOL scale is a 30-item self-report measure that captures the positive and negative effects of working in helping or caregiving roles. It comprises three 10-item subscales: "compassion satisfaction" (the pleasure derived from being able to do one's work well), "burnout" (feelings of exhaustion, frustration and disengagement) and "secondary traumatic stress" (stress resulting from exposure to others' trauma). Items are rated on a 5-point scale from never to very often, with subscale scores calculated separately. Higher scores indicate greater levels of the respective construct. The ProQOL scale is widely used across health, education and social care settings and has demonstrated good reliability and validity (Stamm, 2010).

Sociodemographic data

Sociodemographic information, including each pupil's date of birth, sex, ethnicity, year group, social care status, eligibility for FSMs and the presence of SEND or an EHCP, was obtained from the schools' administrative records. The pupil survey additionally captured gender, ethnicity and year group, while the staff survey collected information on staff gender, ethnicity and role within the school.

Baseline measures

For all primary and secondary outcomes, baseline data was collected prior to randomisation and the start of the intervention. Each outcome measure at endline was compared with the corresponding pre-intervention measure, which served as the baseline variable for that outcome.

Table 5: Mapping of logic models onto measures.

Outcome	Variable	Measure	Research question	Aligned logic model component
Pupils				
Externalising difficulties	Externalising difficulties	Strengths and Difficulties Questionnaire	Primary	Reduced externalising difficulties
Mental health difficulties	Internalising difficulties Impact score	Strengths and Difficulties Questionnaire	2.1	Reduced mental health difficulties
Prosocial behaviour	Prosocial behaviour	Strengths and Difficulties Questionnaire	2.2	Safe social connection(s)
Safe social connection(s) with teachers	Teacher–student relationships	Student Engagement Instrument	2.3	Safe social connection(s)
Safe social connection(s) with peers	Peer support for learning	Student Engagement Instrument	2.4	Safe social connection(s)
School inclusion	Control and relevance of school work Future aspirations and goals	Student Engagement Instrument	2.5	Educational inclusion, attendance, attainment
Bullying	Bullying perpetration Bullying victimisation	Illinois Bully Scale	2.6	Reduced bullying
Academic attendance	Attendance percentage in the 23/24 and 24/25 academic years	Local school data and National Pupil Database for long-term follow-up	2.7	Educational inclusion, attendance, attainment
Exclusions	Fixed-term exclusions/suspensions Permanent exclusions	Local school data and National Pupil Database for long-term follow-up	2.8	Educational inclusion, attendance, attainment

Outcome	Variable	Measure	Research question	Aligned logic model component
Staff				
Knowledge and awareness	Underlying causes of problem behaviour and symptoms	Attitudes Related to Trauma-Informed Care scale	3.1	Understanding the processes and consequences of trauma
				Increases in staff knowledge and awareness (on attachment, trauma, recovery, wellbeing)
Confidence	Self-efficacy at work	Attitudes Related to Trauma-Informed Care scale	3.2	Staff confidence Staff "resilience"
Wellbeing	Compassion satisfaction	Professional Quality of Life scale	3.3	Staff wellbeing and job satisfaction
Burnout	Secondary traumatic stress	Professional Quality of Life scale	3.4	Staff burnout (sick leave, retention)
	Burnout			
Emotionally safe environment	Responses to problem behaviour and symptoms	Attitudes Related to Trauma-Informed Care scale	3.5	Relationship-positive strategies for addressing behaviour
				Feeling safe*
				Reductions in sanctions
Vicarious trauma	Reactions to work	Attitudes Related to Trauma-Informed Care scale	3.6	Space to recognise and support staff's trauma
Empathy-focused behaviours	Empathy and control	Attitudes Related to Trauma-Informed Care scale	3.7	Relationship-positive strategies for addressing behaviour
				Understanding of young people's needs
				Systemic relational approaches

Note. Mechanisms and outcome of the logic model not included in the above are: self (psychological) awareness for pupils, which will be assessed using qualitative data; involvement in criminal activities for pupils, which is beyond the timeframe of this evaluation; organisational change, which will be assessed through the support description survey (SDS); and efficacy in de-escalating and resolving behavioural challenges, which will be assessed using qualitative data. * = Although feeling safe is included in the table above, it is a pupil mechanism in the logic model and being assessed through staff survey data. Therefore, it will also be assessed using pupil qualitative data.

Data collection

The outcomes data was collected via an online self-completion survey of young people, hosted by the evaluation team. The baseline survey was completed prior to randomisation, with subsequent surveys conducted at midline (approximately five months after randomisation) and at endline (after training was complete, approximately 16 months after randomisation). Data collection periods were coordinated across cohorts to ensure they occurred at broadly the same stages relative to training inputs.

For the baseline survey, schools were sent links to the survey to distribute to pupils and staff. For subsequent surveys (midline and endline), where possible, young people were sent an email by the evaluation team containing a link to the online survey, or the school was sent pupil-specific access codes. Where it was not feasible to send pupil-specific access codes to schools, such as with school administrative constraints or preferences, the evaluation team generated a generic link, which school staff then passed on to the young person. Responses were subsequently matched to the previous survey using pupil demographic data, such as names, date of birth and year group. This was also cross-referenced with the school administrative data that was provided in instances where there were errors in the self-report data. Once matching was complete, identifying information was removed and responses were given a study-specific ID.

For the endline survey, evaluation team members visited schools that were experiencing difficulties with data collection (e.g. internet access issues or staff availability) and supported the administration of the survey. In some cases, paper copies of the survey were used and responses were later entered into the system by an evaluation team member.

The primary outcome was assessed using the endline survey. Members of the analysis team were blinded to allocation, and allocation information was not shared or accessible to them.

In addition to survey data, administrative data on study participants was collected. This included pupil demographic characteristics (date of birth, sex, ethnicity, year group, social care status), educational attendance and suspensions/exclusions, eligibility for FSMs, and SEND and/or EHCP status. The administrative data also included each pupil's Unique Pupil Number. This data was already held in schools' management information systems and, for all but one school, extracted through the use of standardised scripts, also known as reports. The data was shared with the evaluation team through a secure file transfer portal following the signing of a Data Sharing Agreement. Data covering the 2023/24 school year was requested, starting in March 2025, with an additional submission, including attendance and exclusion statistics from the 2024/25 school year, requested in July 2025. The reports used in data extraction were developed by School ICT, a consultancy specialising in management information systems data.

Similar to pupils, school staff completed surveys on understanding and practice of trauma-informed practice, via an online platform at baseline, midline, and endline.

A single point of contact at each school completed a business-as-usual SDS. Schools in the intervention arm additionally provided information on the trauma-informed practices they had implemented. The single contact also reported any costs incurred as a result of participating in the programme (e.g. staff time by role). These surveys were completed in both Year 1 and Year 2 of the trial, and in some cases, schools provided retrospective cost information for the entire trial in Year 2.

Sample size

An overview of the sample size calculations is presented in Table 6.

Table 6: Sample size calculations overview

		Protocol	Randomisation
Minimum Detectable Effect Size		Standardised mean difference = 0.15	Standardised mean difference = 0.17
Pre-test/post-test correlations	Level 1 (participant)	-	-
	Level 2 (cluster)	-	-
Intracluster correlations	Level 1 (participant)	0.03	0.041
	Level 3 (cluster)	-	-
Alpha		0.05	0.05
Power		0.8	0.8
One-sided or two-sided?		Two-sided	Two-sided
Average cluster size		186 pupils 159 staff	220 pupils 26 staff
Number of clusters	Intervention	29	31
	Control	29	31
	Total	58	62
Number of participants	Intervention	4,650 pupils 3,975 staff	6,429 pupils 724 staff
	Control	4,650 pupils 3,975 staff	7,194 pupils 896 staff
	Total	9,300 pupils 7,950 staff	13,623 pupils 1,620 staff

The cluster sample size was determined a priori using a Minimum Detectable Effect Size (MDES) of 0.15 (standardised mean difference [SMD]). The trial protocol considered a range of MDES values from 0.10 to 0.20. The calculation of an MDES of 0.15 assumed recruitment of 55 schools, inflated to 58 to account for 15% attrition. Estimates for the mean and standard deviation of externalising difficulties in the control arm were based on combined male and female scores from Elia et al. (2020), and the intraclass correlation coefficient (ICC) was assumed to be 0.03 at the participant level based on typical ICCs reported for outcomes in school-based cluster trials (Hedges & Hedberg, 2007). No adjustment was made for pre–post correlation due to limited evidence.

Power calculations were conducted in STATA using the following code:

```
power twomeans 5.21 4.76, sd(2.96) rho(0.03) m1(186) m2(186) power(0.8) cvcluster(0.5) cluster(25)
```

This corresponds to 4,650 pupils per arm, increased to 29 schools per arm to allow for 15% school attrition.

The primary population of interest was pupils, with staff also included. Based on figures from three LAs (Kent, Slough and Wiltshire), 186 pupils per school were assumed after accounting for 35% refusal and 25% attrition, giving a total of 9,300 pupils. For staff, 159 per school were assumed, giving a total of 7,950.

Following randomisation in February 2024, actual data updated the estimates. The average cluster size for pupils completing the primary outcome (SDQ externalising symptoms) was 219.73 (rounded to 220). The ICC was 0.041 and the coefficient of variation was 0.366, yielding power to detect an MDES of 0.17. STATA calculations estimated 24.3 schools and 5,333 pupils per arm, inflated to 29 schools and 6,372 pupils per arm to account for 15% attrition:

```
power twomeans 5.21 4.7, sd(2.96) rho(0.041) m1(219.73) m2(219.73)

power(0.8) cvcluster(0.366) nfracdisplay "schools per arm = " ceil(r(K1)/.85)

display "pupils per arm = " round(ceil(r(K1)/.85)*219.73)
```

Randomisation

Schools were randomised in two cohorts: cohort 1 in November 2023 (n = 12 schools) and cohort 2 in February 2024 (n = 50 schools). Randomisation was stratified by LA or CA, with a minimum of two schools required for an LA/CA to be included in the randomisation. The King's Clinical Trials Unit conducted the randomisation. Schools were enrolled by the programme delivery team (KCA and WLA), randomisation was carried out by the King's Clinical Trials Unit, and Anna Freud Research Officers informed schools of their allocation. The analysis team remained blinded to allocation, and allocation information was not shared or accessible.

Statistical analysis

Primary analysis

The primary outcome was externalising difficulties, calculated as the sum of the conduct problems and hyperactivity subscales of the SDQ. The outcome analysis followed an ITT approach, whereby all randomised schools were included in the groups to which they were originally allocated. This applied irrespective of the level of programme engagement, any withdrawal after randomisation or deviations in implementation. Analysing outcomes according to the initial assignment provides an unbiased estimate of the effect of offering the intervention and reflects real-world delivery conditions. It is also a conservative analytical strategy, as it captures the average effect across varied levels of adherence.

Group differences on the primary outcome were examined using a mixed-effects model consistent with the pre-specified statistical analysis plan (SAP). The primary analysis model estimated the effect of the intervention by including fixed effects for trial group (intervention vs control) while also controlling for baseline outcome scores. The model further adjusted for cohort, LA/CA, and (to adjust for baseline imbalances) baseline internalising symptoms, gender and ethnicity. A random intercept was included at the school level to account for the clustered trial design. Please note that the SAP specified random intercepts for pupils, but this was not possible given the wide structure of the data. As a result, there were no repeated observations at the pupil level, and a random intercept could not be estimated. For pupil (i) in school (j), we define the model as:

$$SDQextern_{ij} = \beta0 + \beta1(group_j) + \beta2(\text{baseline SDQ extern}_{ij}) + \beta3(cohort_j) + \beta4(\text{baseline SDQ intern}_{ij}) + \beta5(gender_{ij}) + \beta6(ethnicity_{ij}) + \beta7(LA/CA_{ij}) + u0_j + e_{ij}$$

The model was fitted using the lmer() function in R, using the following specification:

```
lmer(endline_SDQ_extern ~ group + baseline_SDQ_extern + cohort + baseline_SDQ_intern + gender + ethnicity + LA_CA + (1 | school), data = data)
```

P-values in the main body were taken from the lmerTest::summary() function, whereas the model outputs in the appendix were taken from the parameters::model_parameters(). The p-values in the appendix can differ very slightly from the ones in the main body, since these functions make slightly different calculations.

There were no instances where the differences were large enough to influence the conclusions of the analysis.

Secondary analysis

Secondary analyses were conducted on secondary pupil and staff outcomes. Secondary pupil outcome models followed the same structure as the primary analysis. For the SDQ impact score, baseline data was not collected to minimise participant burden, so the model was adapted to exclude this variable.

For the secondary staff outcomes, models included fixed effects for group, baseline outcome scores, cohort and LA/CA. A random intercept was specified at the school level. The model took the following form:

lmer(endline_outcome ~ group + baseline_outcome + cohort + LA _CA + (1 | school))

In line with the SAP, as the project contained multiple secondary and tertiary outcomes and analyses, false discovery rate (FDR) multiple comparison correction using the Benjamini-Hochberg procedure was applied to the significance level for tests for outcomes that are within the same measure, as these are likely to be correlated both theoretically and statistically.

Analysis in the presence of non-compliance

The compliance analysis was based on two data sources: the KCA's online environment (KCA Connected), where details of all delivered sessions were recorded, and an online attendance form that trainers prompted session attendees to fill out. For most attendance forms (3932/4258; 92%), a corresponding session was found on KCA Connected with the same session name, school and date. The remaining forms were matched, if possible, to a session being delivered at the same school on the same date.

Compliance for each school was calculated as the mean attendance rate across all sessions. Please note that this compliance metric was updated from the protocol, and this metric used is mentioned in the SAP, as this aligns with subsequent analyses. Mean session attendance accurately reflects the level of intervention exposure, and it provides a clearer and more appropriate measure of compliance. Attendance at sessions that were split into several parts was calculated as the average of the rates for the different parts. Sessions that form part of the programme but were not delivered were counted as having zero attendance. This includes the reflective practice consultation groups, three of which were intended to be delivered to senior leadership, along with three groups for PI leads. Where schools decided to merge their reflective practice consultation groups between the senior leadership team (SLT) and pastoral/inclusion (PI) leads, no zero attendance was added as long as three consultation groups were held for the combined audience. The numerator for the attendance rates was taken from attendance forms that trainers prompted session attendees to fill out during sessions. Two sessions (Reflective Practice Training and Therapeutic Ideas) very commonly had zero forms, and we were not able to confirm whether attendees were prompted to complete a form in each of these sessions, so these sessions were excluded from the compliance analysis.

The denominators for attendance rates at all-staff sessions were taken from whole-school staff counts requested at the start of the trial. The denominators for attendance rates at senior staff sessions were taken from invite counts on the delivery partner's monitoring system. As the delivery partner indicated that in some cases, these invites were forwarded by school staff without adding invites to the system, invite counts of three or less were discarded and replaced by the average of invite counts of four or more for the respective audience.

Complier average causal effect (CACE) analysis was used, adopting the same specification as the primary analysis model, to estimate the intervention effect on the primary outcome under a hypothetical scenario in which all schools in the intervention arm complied with the programme. In this framework, randomisation serves as an instrumental variable: it predicts the likelihood of compliance and allows estimation of the causal effect of compliance on pupil outcomes while preserving the benefits of random assignment. The average percentage of eligible staff attendance across training sessions for each intervention school was used as the compliance measure. As no school met the 80% threshold for compliance (consistent with common practices in CACE analysis; Ye et al., 2014), the upper quartile of the average percentage was used as the threshold.

The analysis followed a two-stage instrumental variable procedure. In the first stage, compliance status was modelled as a function of randomisation group and other covariates, producing predicted probabilities of compliance for each school. In the second stage, the primary outcome was regressed on these predicted compliance values, together with time and the covariates included in the primary model. This two-stage approach ensures that the estimated effect reflects the causal impact of compliance rather than being biased by non-compliance. A bootstrapping procedure which included both stages of the model was applied to obtain robust standard errors for the effect of compliance (Freedman, 1984).

Missing data analysis

Missing data analysis was conducted to clarify the missing data mechanism and support valid inferences. Schools that withdrew post-randomisation (six intervention schools and two control schools) were still asked to participate in endline data collection to minimise missing data; however, no schools agreed to participate in this. The extent of missing data in the primary outcome and covariates was examined. To assess whether data was missing completely at random (MCAR), Little's MCAR test (Little, 1988) was performed using the LittleMCAR() function in the mice package in R. Where data was MCAR, complete case analysis was considered to provide unbiased estimates. For variables with more than 5% missing data, the extent of systematic missingness was evaluated using multi-level logistic regression (modelling each variable as 0 = not missing and 1 = missing) to identify predictors of missingness for the primary outcome and covariates. If the primary outcome is missing conditional on covariates (MAR), then the primary analysis (complete cases analysis) will include covariates that predict missingness into the model. The results will be compared with the complete case analysis that does not include these covariates.

Where covariates were missing conditional on other covariates or outcomes (i.e. missing at random, MAR), multiple imputation (MI) was conducted, and results were compared with the complete case analysis. MI was performed for covariates used in the primary analysis using the mice package (van Buuren & Groothuis-Oudshoorn, 2011). When building the predictor matrix, variables strongly associated with missingness were included as predictors, and the temporal order of variables was maintained (e.g. midline scores were not used to predict baseline scores). School was also included as a predictor in the imputation model to account for clustering within schools. The number of imputations matched the percentage of missing data. Analyses were conducted on the imputed datasets, and pooled estimates were derived.

Non-response weighting was applied to account for biases arising from differential participation in completing the primary outcome. Weights were computed by first conducting a multi-level logistic regression to predict the probability of response based on covariates (gender, ethnicity, baseline internalising symptoms, baseline externalising symptoms, school characteristics) and then calculating the inverse of the predicted response probability (Valliant et al., 2013). The weights were normalised and incorporated into the primary analysis model.

Additional analyses and robustness checks

An additional analysis was undertaken in which the primary model was modified to include midline externalising difficulties scores. The mixed effects model included both midline and endline SDQ externalising difficulties as long-format data, and the predictors in the model included group, time (midline, endline), their interaction, baseline SDQ externalising symptoms and covariates, as in the primary model. This analysis was carried out using complete cases, as full information maximum likelihood (FIML) was not available in the lmer package, and an alternative single-level FIML approach (such as lavaan) was not considered appropriate given the clustered structure of the data. Results from this model were compared with those from the primary analysis to assess the consistency of findings.

Moderation analysis within the primary analysis

The primary analysis model included a cross-level interaction between group allocation and mean school-level scores on the staff-reported on-the-job behaviour (empathy and control) subscale of the ARTIC-35. This was to examine whether allocation to treatment vs control groups predicts different levels of externalising difficulties for pupils in schools where staff report higher levels of empathy-focused behaviours.

Longitudinal mediation analysis within the primary analysis

A longitudinal mediation analysis was conducted to examine whether safe social connections with teachers, measured by the SEI at midline, mediated the relationship between group allocation and changes in externalising difficulties from baseline to endline, while accounting for clustering by schools. The model was estimated using the lavaan package in R (version 0.6-17; Rosseel, 2012). Due to convergence issues, covariates were removed from the model. This analysis was exploratory and intended to inform potential future research.

Subgroup analyses

An exploratory analysis was undertaken to examine whether there was evidence of subgroup differences in the intervention effect across the primary outcome. Differential effects were tested through mixed effects models in which interaction terms were added between the randomisation group and each demographic variable (ethnicity, FSM eligibility and SEND/EHCP status). Ethnicity was entered using dummy-coded categories with White as the reference group, as it was the largest group, and FSM eligibility and SEND/EHCP status were each included as binary interaction terms. Estimates for each level were presented with 95% confidence intervals, with statistical inference based on omnibus tests across categories for multi-level variables.

Estimation of effect sizes

Effect sizes were calculated to quantify the impact of the intervention at endline. A multi-level model was fitted using the lmer() function in R, including fixed effects for group, baseline outcome scores and relevant covariates, with a random intercept for school to account for clustering. SMDs were calculated using the following formula:

$$ES = \frac{\Delta\text{group (adjusted)}}{\text{SD}_{pooled}}$$

where $\Delta\text{group (adjusted)}$ is the adjusted endline group coefficient from the mixed-effects model (representing the intervention–control difference at endline) and SD_{pooled} is the pooled unconditional standard deviation of the outcome across both groups at endline. This pooled variance was calculated using the endline analytical sample without adjustment for covariates or model structure, in accordance with YEF guidance. The resulting SMD was then converted to Hedges' g by applying the J-correction to adjust for small-sample bias. Effect sizes were reported with confidence intervals and p-values to reflect statistical uncertainty.

For binary outcomes, odds ratios from the mixed-effects models were transformed into an effect size comparable to Hedges' g using the Cox Index and are reported with 95% CIs and p-values to reflect statistical uncertainty.

Estimation of intraclass correlation coefficient

Clusters in this trial are schools. ICCs at the school-level were calculated separately at baseline and endline using the package rpt in R. Adjusted ICCs are reported with 95% confidence intervals.

Implementation and process evaluation

Research methods

The IPE explored the delivery of MGDAS within participating schools, the extent to which it was embedded in existing systems of support, and the experiences of pupils, staff and programme implementers. A mixed-methods approach was adopted to capture both quantitative and qualitative insights into implementation and perceived impact.

In each participating school, a nominated staff member (typically the MGDAS single point of contact) completed an SDS at the start and end of the trial: autumn 2023 and spring/summer 2025. The survey included both closed and open-ended questions and captured information on school policies, practices and provisions for mental health and wellbeing support. This data was used to describe the support typically available (business as usual) and to identify any changes in provision over the course of the study. The original plan was to use the support description to construct a composite score reflecting the number and type of policy and practice changes implemented following the MGDAS programme. However, complete information on the type of changes implemented was not available for all schools, which meant that constructing a reliable composite score was not feasible. Instead, we conducted a descriptive comparison of implementation activity using the data that was available. This involved comparing baseline levels of policy and practice indicators between intervention and control schools and, for schools that provided both baseline and endline data, examining changes over time descriptively across the two conditions. At baseline, all 31 intervention schools (100%) completed the SDS, as did 29 control schools (94%). In addition, there were 11 schools in the intervention arm (35%) and 11 in the control arm (38%) that also completed the survey at endline.

A multi-informant qualitative design was used to explore how MGDAS was experienced by pupils and staff and how it influenced school culture, policies and practices. Four schools were selected as qualitative case studies. Selection was based on a school's capacity to engage in the evaluation and their sociodemographic diversity. Sampling criteria were designed to ensure representation across key dimensions, including location, pupil ethnicity, socioeconomic context and levels of SEND. All case study schools were drawn from the intervention arm to allow an in-depth exploration of implementation processes and perceived outcomes. Schools from both cohort 1 (randomised in November 2023) and cohort 2 (randomised in

February 2024) were included to capture experiences from schools at different stages of programme delivery. Interviews were carried out by Anna Freud staff members who had some knowledge of the MGDAS programme but had not been involved in the facilitation or administration of the training sessions to minimise bias. The interview topic guide was directly informed by the IPE framework of research questions. Semi-structured interviews were conducted in person with four to six pupils per case study school during site visits from January to April 2025 (overall n = 21). Pupil interview lengths ranged from 12 to 51 minutes (M = 30.72, SD = 9.00).

Interviews were also conducted over Microsoft Teams from January to July 2025 with one to three school staff members across 13 schools (overall n = 23). Due to demands on school staff time and capacity and difficulties in securing participants, we decided not to restrict school staff interviewee recruitment to the four case study schools but to broaden our recruitment pool to all intervention schools. School staff interview lengths ranged from 16 to 46 minutes (M = 32.98, SD = 7.81).

To complement these perspectives, interviews were conducted with nine MGDAS implementers between May and July 2024 to capture insights into programme delivery and implementation across participating schools. Implementer interview lengths ranged from 34 to 59 minutes (M = 53.59, SD = 7.19).

Analysis

Qualitative data analysis was conducted by three members of the research team. The qualitative interview data was initially analysed drawing on the framework approach (Ritchie & Spencer, 1994). This involved the data being coded top-down or deductively to broad categories representing our research questions for each participant group. Following this, the data coded to each category for each participant group was then analysed, drawing on Braun and Clarke's (2020) reflexive thematic analysis approach. This involved the following stages: systematic recoding of the data coded to each category using a bottom-up or inductive approach, whereby transcript extracts were given labels or codes to describe their content; generating themes, whereby codes that represented similar perspectives, opinions or experiences were combined to form themes; developing and reviewing themes through discussion and doing additional coding as necessary; and ensuring that each theme had a name and definition that was clearly grounded within the data that it represented. This approach enabled us to triangulate findings across multiple qualitative data sources. Analyses were conducted by participant group (pupil, school staff and MGDAS implementers), with pupil qualitative data also explored alongside qualitative data from school staff at the case study schools.

A triangulation design convergence model (Creswell & Clark, 2018) was applied to address the overarching research question: How effective is the implementation of a whole-school approach to trauma-informed practice in improving safe social connections (mechanism) and reducing externalising difficulties (outcome), thereby decreasing the likelihood of young people becoming involved in crime and violence in the future (long-term outcome)? The qualitative and quantitative findings from the RCT and the IPE were interpreted together to provide an integrated understanding of programme impact and implementation.

Implementation and monitoring data collected by the MGDAS delivery team were used to examine programme reach and fidelity. This data captured the number and type of sessions delivered to schools, as well as levels of staff engagement. Descriptive statistics were used to summarise the extent of implementation and participation across schools.

Data from the SDS completed by the single point of contact at each school was analysed to characterise existing business-as-usual support for pupil wellbeing and to explore how this evolved following participation in the MGDAS programme. Both quantitative and qualitative analyses were used to

descriptively summarise these findings, providing an understanding of how trauma-informed practice was interpreted and embedded within schools and the extent to which it differed from prior approaches.

The analysis of intervention fidelity was based on two data sources: KCA's online environment (KCA Connected), where all training sessions were logged, and a spreadsheet where trainers recorded variations on standard delivery.

Table 7: Implementation process evaluation (IPE) methods overview

Research methods	Data collection methods	Participants/ data sources	Data analysis methods	Research questions addressed	Implementation/ logic model relevance
Qualitative	Interviews	21 pupils	Thematic and framework analyses	IPE 1, IPE 8, IPE 9	Examine implementation; assess contexts, mechanisms and outcomes
Qualitative	Interviews Focus groups	23 school staff	Thematic and framework analysis	IPE 1, IPE 4, IPE 5, IPE 6, IPE 7, IPE 9	Examine implementation; assess contexts, mechanisms and outcomes
Qualitative	Interviews Focus groups	9 implementers	Thematic analysis	IPE 1, IPE 2, IPE 3, IPE 6, IPE 7, IPE 9	Examine implementation; assess contexts, mechanisms and outcomes
Quantitative	Implementation monitoring data	25 schools in the intervention arm	Descriptive statistics on numbers and types of sessions delivered and numbers of school staff engaged	IPE 1	Examine implementation
Mixed methods	School support description survey	Baseline n = 60 schools, Paired baseline and endline n = 22	Descriptive statistics and descriptive summaries of free text responses	IPE 1, IPE 4, IPE 5	Examine implementation

Timeline

The timeline of activities related to the evaluation and intervention delivery is presented in Table 8.

Table 8: Timeline

Cohort 1 dates	Cohort 2 dates	Activity	Staff responsible/ leading
May–Jul 23		Ethics approvals	Anna Freud
Jul–Dec 23		Operational set-up and recruitment	Anna Freud
Sep–Dec 23		Schools sign Memorandum of Understanding	Anna Freud/schools
Sep–Oct 23	Sep–Dec 23	Parents/carers of Year 8 and 9 pupils in participating schools sent study information	Anna Freud/schools
Oct 23	Jan 24	Parent/carer opt-out deadline	Anna Freud
Sep–Nov 23	Jan–Feb 24	Baseline data collected (pupil survey, school staff survey, SDS)	Anna Freud
Nov 23	Feb 24	Baseline data review and chasing	Anna Freud
Nov 23	Feb 24	Randomisation	Anna Freud

Nov 23–Jan 24	Feb–Apr 24	MGDAS collaborative enquiry with intervention schools	KCA/WLA
Jan 24–Mar 25	Apr 24–Jun 25	More Good Days At School (MGDAS) training delivery	KCA/WLA
Jan–Feb 24		Statistical analysis plan drafting (cohorts 1 and 2)	Anna Freud
Apr–Jul 24	Jun–Jul 24	Midline data collected (pupil survey, staff survey)	Anna Freud
May–Jul 24	Jul 24	Midline data review and chasing	Anna Freud
May–Jul 24		Implementer qualitative data collected	Anna Freud
May–Jul 24	Jul 24	MGDAS costs data from intervention schools collected	Anna Freud
May–Jul 24	Jul 24	Monitoring and implementation data collected and collated	Anna Freud
May–Jul 24	Jul 24	Local school data received	Anna Freud
Jan 25–Apr 25		School staff and pupil qualitative data collected (cohorts 1 and 2)	Anna Freud
Mar–Jul 25	May–Jul 25	Endline data collected (pupil survey, staff survey, SDS)	Anna Freud
May–Jul 25	Mar–Jul 25	Endline data review and chasing	Anna Freud
May–Jul 25	May–Jul 25	MGDAS costs data from intervention schools collected	Anna Freud
Mar–Jul 25	Mar–Jul 25	Monitoring and implementation data collected and collated	Anna Freud
Mar–Jul 25	Mar–Jul 25	Local school data received	Anna Freud
May–Nov 25		Data analysis and report write-up	Anna Freud

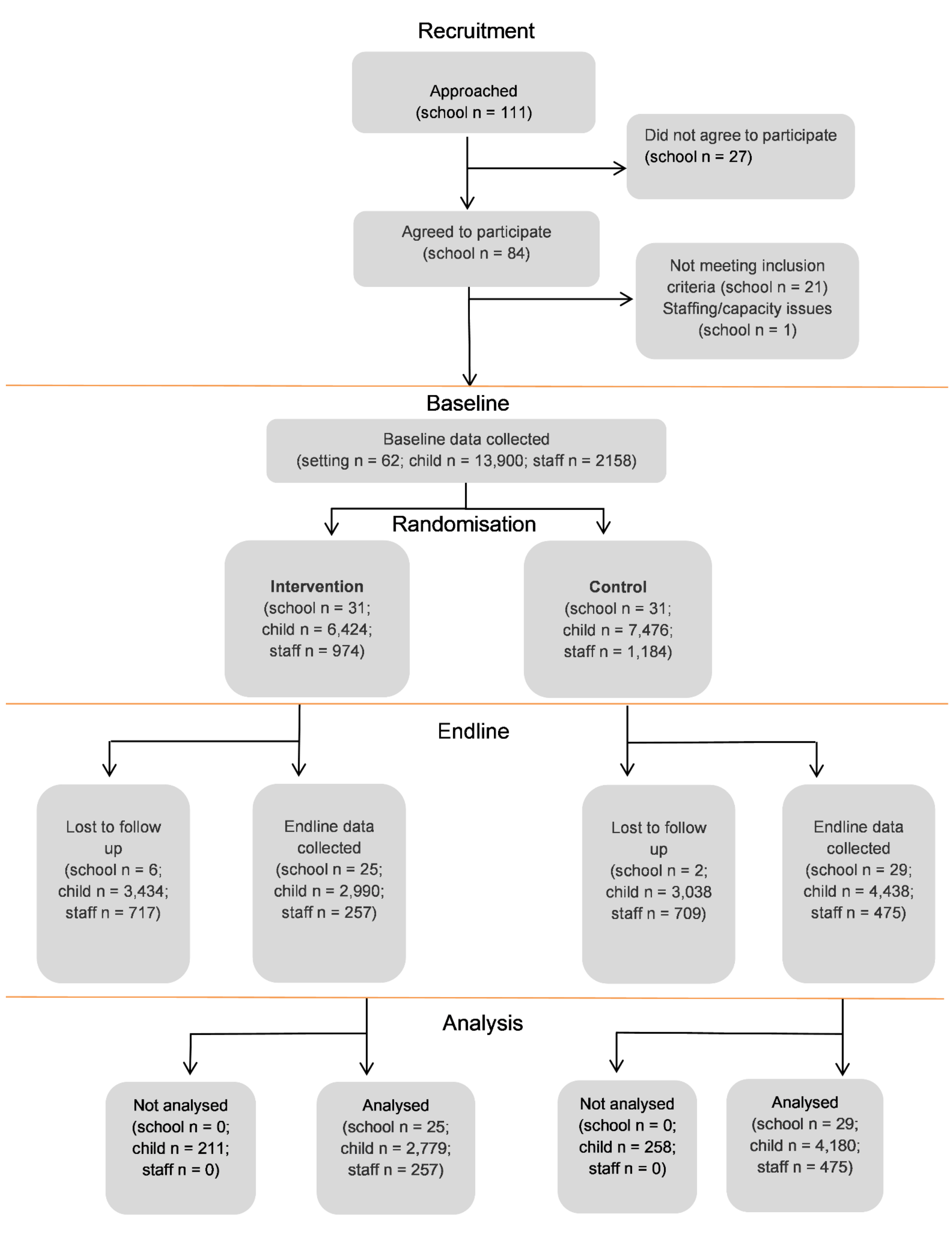
Note: SDS = support description survey.

Impact evaluation results

Participant flow, including losses and exclusions

The participant flow through the stages of the evaluation (recruitment, baseline data collection, randomisation, endline data collection and analysis) for both cohorts is presented in Figure 1. Baseline data was collected from 62 schools (31 per arm), including 13,900 pupils (intervention n = 6,424; control n = 7,476). Losses to follow-up comprised both school withdrawals and pupil-level attrition within retained schools. At analysis, 25 intervention schools (81%) remained, representing 2,779 pupils (43%), while 29 control schools (94%) remained, representing 4,180 pupils (56%). The sample sizes reported reflect the primary analysis. A detailed discussion of attrition is provided in the Attrition section.

Figure 1: Participant flow through the stages of the evaluation (recruitment, baseline data collection, randomisation, endline data collection and analysis) for both cohorts



Pupil demographic characteristics are presented in Table 9. At baseline, the majority of pupils identified as White (59% intervention; 56% control), with Asian or Asian British pupils forming the next largest group

(22% intervention; 25% control). Smaller proportions identified as Black (6% in both arms), Mixed (4–5%), or Other ethnic groups (around 4%), while approximately 5% of ethnicity data was missing. Gender distribution was broadly balanced: boys comprised 46% of the intervention group and 52% of the control group; girls 49% and 44% respectively; 2% identified as non-binary, questioning, or other; and around 2% of gender data was missing.

By endline and in the analytical sample, attrition reduced overall numbers, but the demographic profile remained similar. In the intervention arm, 2,779 pupils were retained for analysis, and 4,180 were retained in the control arm. The proportion of White participants decreased slightly (49% intervention; 55% control), while the proportion of Asian or Asian British pupils increased (33% intervention; 30% control). Gender distributions were stable, with boys representing about half of each arm, girls just under half and non-binary/other around 2%. Missing data was eliminated in the analytical sample, ensuring complete demographic information for the primary analysis.

Table 9: Pupil demographic characteristics of the sample at each stage of the evaluation.

	Baseline data n (%)	**Endline data n (%)**	**Analytical sample n (%)**
Ethnicity			
Asian or Asian British			
Intervention	1,423 (22.2%)	937 (31.3%)	929 (33.4%)
Control	1,862 (24.9%)	1,273 (28.7%)	1,252 (30.0%)
Black, Black British, Caribbean or African			
Intervention	398 (6.2%)	234 (7.8%)	231 (8.3%)
Control	455 (6.1%)	259 (5.8%)	256 (6.1%)
Mixed or multiple ethnic groups			
Intervention	261 (4.1%)	128 (4.8%)	124 (4.5%)
Control	368 (4.9%)	225 (5.1%)	223 (5.3%)
White			
Intervention	3,763 (58.6%)	1,383 (46.3%)	1,354 (48.7%)
Control	4162 (55.7%)	2,332 (52.5%)	2,296 (54.9%)
Other ethnic group			
Intervention	254 (4.0%)	145 (4.8%)	141 (5.1%)
Control	283 (3.8%)	154 (3.5%)	153 (3.7%)
Missing			
Intervention	325 (5.1%)	163 (5.5%)	0
Control	346 (4.6%)	195 (4.4%)	0
Gender			
Male			

Intervention	2,966 (46.2%)	1,468 (49.1%)	1,401 (50.4%)
Control	3,848 (51.5%)	2,288 (51.6%)	2,190 (52.4%)
Female			
Intervention	3,158 (49.2%)	1,402 (46.9%)	1,333 (48.0%)
Control	3,314 (44.3%)	1,985 (44.7%)	1,906 (45.6%)
Other (non-binary, questioning, other)			
Intervention	157 (2.4%)	52 (1.7%)	45 (1.6%)
Control	182 (2.4%)	93 (2.1%)	84 (2.0%)
Missing			
Intervention	143 (2.2%)	68 (2.3%)	0
Control	132 (1.8%)	72 (1.6%)	0

The staff demographic profile (see Table 10) was broadly similar across the intervention and control groups, with the majority identifying as White (around 72–75%), and females making up roughly 69–72% of the workforce. Asian or Asian British staff represented the next largest ethnic group (about 17–23%, depending on the sample), while other ethnic categories each accounted for only small proportions, with some suppressed due to low numbers. Teaching staff formed the largest role category, comprising just over half of participants in both groups, followed by support and leadership staff (15–20%) and smaller proportions of administrative staff and mental health leads. Overall, the analytical sample closely mirrored the baseline, maintaining a predominantly White, female and teaching-focused staff composition.

Table 10: Staff demographic characteristics at baseline and for the analytical sample.

	Baseline data n (%)	**Analytical sample n (%)**
Ethnicity		
Asian or Asian British		
Intervention	187 (19.2%)	43 (16.7%)
Control	219 (18.5%)	107 (22.5%)
Black, Black British, Caribbean or African		
Intervention	32 (3.3%)	#
Control	25 (2.1%)	#
Mixed or multiple ethnic groups		
Intervention	17 (1.7%)	#
Control	25 (2.1%)	#
White		
Intervention	705 (72.4%)	#
Control	882 (74.5%)	#
Other ethnic group		

Intervention	23 (2.4%)	#
Control	25 (2.1%)	#
Missing		
Intervention	10 (1.0%)	#
Control	15 (1.3%)	#
Gender		
Male		
Intervention	265 (27.2%)	70 (27.2%)
Control	364 (30.7%)	148 (31.2%)
Female		
Intervention	702 (72.1%)	185 (72.0%)
Control	813 (68.7%)	326 (68.6%)
Other (non-binary, questioning, other)		
Intervention	#	#
Control	#	#
Missing		
Intervention	#	#
Control	#	#
School role		
Administrative staff		
Intervention	44 (4.5%)	#
Control	63 (5.3%)	15 (3.2%)
Mental health lead		
Intervention	63 (6.5%)	15 (5.8%)
Control	74 (6.2%)	27 (5.7%)
School leadership		
Intervention	151 (15.5%)	41 (16.0%)
Control	183 (15.5%)	80 (16.8%)
Support staff		
Intervention	191 (19.6%)	47 (18.3%)
Control	194 (16.4%)	71 (14.9%)
Teaching staff		
Intervention	495 (50.8%)	143 (55.6%)
Control	610 (51.5%)	272 (57.3%)
Other		
Intervention	29 (3.0%)	#
Control	60 (5.0%)	#
Missing		
Intervention	#	#
Control	#	#

Note. # = number suppressed due to small n.

Attrition

At the pupil level, 13,900 participants were randomised across 62 schools (intervention n = 6,424; control n = 7,476). In the primary outcome model, which includes data from baseline and endline only, a complete case analysis was conducted on 6,959 pupils (intervention n = 2779; control n = 4180) across 25 intervention

schools and 29 control schools. This represents an overall attrition rate of 49.9% from randomisation to analysis. As shown in Table 11, attrition was higher in the intervention arm (56.7%) than the control arm (44.1%).

Losses to follow-up were attributable to school dropout, pupils leaving schools and challenges in engaging schools to complete endline measures. School dropout was primarily due to staffing and capacity issues. One intervention school withdrew due to a lack of perceived impact of the intervention, and another was withdrawn after not responding to the delivery team for several months at the start of the trial. These factors contributed to the differential attrition observed between arms. The analytical sample, therefore, reflects only those pupils with complete data available for the primary analysis.

Attrition differed significantly by ethnicity ($\chi^2(4) = 396.83$, $p < 0.001$). Post-hoc tests showed that Asian pupils were disproportionately retained in the analysed sample (standardised residual = 18.26, $p < 0.001$), whereas White pupils were disproportionately lost to attrition (standardised residual = −18.45, $p < 0.001$). These findings indicate that retention in the primary analysis was higher among Asian pupils and lower among White pupils compared with other ethnic groups.

The level of attrition substantially reduced the completeness of the dataset and introduced potential bias, as attrition was uneven across groups and demographics. The disproportionate withdrawal in the intervention arm, alongside differential retention by pupil ethnicity, may present challenges for the representativeness of the sample and the generalisability of findings.

Table 11: Pupil level attrition from the trial (primary model)

		Intervention	Control	Total
Number of participants	Randomised	6,424	7,476	13,900
	Analysed	2,779	4,180	6,959
Participant attrition (from randomisation to analysis)	Number	3,645	3,296	6,941
	Percentage	56.74%	44.09%	49.94%

We conducted an interview with one school after withdrawing from the study, and they described competing priorities as an important factor. It was mentioned that the requirements of the training were *"too onerous"*, especially when the school in question was *"undergoing a lot of transition"* in terms of leadership. This school found its time was too restricted to incorporate MGDAS alongside other required training: *"there's a lot of competition for, for example, staff inset time, and the amount of training that was required for the More Good Days at School just wasn't compatible with the amount of time that we had available"* (School staff).

One reflection from the withdrawn school was that there was *"a lot of communication"* from the programme team, which was *"quite overwhelming"*, and, to some extent, left them feeling even more aware that they would struggle to fit in the training sessions: *"that was quite overwhelming as well at the time because it was like, 'Oh crikey, how am I going to get that to work?' and, 'Where am I going to fit that in?'"* (School staff).

Participant characteristics

Baseline characteristics of the groups as randomised at the setting-level and participant-level are presented in Table 12. There was high missingness in pupil-level FSM eligibility and SEND data, primarily due to the non-submission of administrative records by some schools. Missingness was also higher for the SEI and IBS

subscales than for the SDQ. This likely reflects the fact that these subscales cannot be prorated, meaning that a single missing item results in a missing subscale score. In addition, these measures appeared later in the survey, so the higher missingness may also indicate non-response related to survey fatigue.

Table 12: Baseline characteristics of groups as randomised

Setting-level (categorical)	Intervention group		Control group	
	n/N (missing)	Count (%)	n/N (missing)	Count (%)
Cohort: 1		6 (19.4%)		6 (19.4%)
Cohort: 2		25 (80.6%)		25 (80.6%)
Setting-level (continuous)	**n/N (missing)**	**Mean (SD)**	**n/N (missing)**	**Mean (SD)**
Percentage of pupils receiving SEND support	31	14.12(6.02)	31	14.18 (3.79)
Percentage of pupils eligible for FSMs in the last six years	31	31.55 (11.89)	31	32.35 (14.02)
Total number of pupils on roll (all ages)	31	976.23 (376.25)	31	996.52 (371.12)
Participant-level (categorical)	**n/N (missing)**	**Count (%)**	**n/N (missing)**	**Count (%)**
Year group: 8		3,111 (48.43%)		3,601 (48.17%)
Year group: 9		2,860 (44.52%)		3,483 (46.59%)
Year group: missing		453 (7.05%)		392 (5.24%)
Pupil-level SEND: yes		551 (8.58%)		872 (11.66%)
Pupil-level SEND: no		3,218 (50.09%)		4,477 (59.88%)
Pupil-level SEND: missing		2,655 (41.33%)		2,127 (28.45%)
Pupil-level FSMs: yes		1,228 (19.12%)		1,493 (19.97%)
Pupil-level FSMs: no		2,541 (39.55%)		3546 (47.43%)
Pupil-level FSMs: missing		2,655 (41.33%)		2437 (32.60%)
Participant-level (continuous)	**n/N (missing)**	**Mean (SD)**	**n/N (missing)**	**Mean (SD)**
SDQ externalising symptoms	6,394 (30)	7.63 (4.24)	7,451 (25)	7.52 (4.23)
SDQ internalising symptoms	6,397 (27)	6.13 (3.75)	7,447 (29)	5.92 (3.76)
SDQ prosocial behaviour	6,413 (11)	6.61 (2.16)	7,458 (18)	6.61 (2.15)
SEI teacher-student relationships	5,791 (633)	23.05 (5.43)	6,936 (540)	23.10 (5.33)
SEI peer support at school	5,737 (687)	17.62 (3.7)	6,904 (572)	17.69 (3.77)
SEI control and relevance of schoolwork	5,818 (606)	23.75 (5.14)	7,026 (450)	23.67 (5.13)
SEI future aspirations and goals	5,814 (610)	15.85 (3.23)	7,017 (459)	15.91 (3.23)
IBS bullying perpetration	5,785 (639)	3.88 (5.73)	7,096 (380)	3.94 (5.69)

IBS bullying victimisation	5,638 (786)	3.90 (4.58)	6,895 (581)	3.79 (4.53)
ARTIC-35 underlying causes	863 (111)	4.86 (0.73)	1,092 (92)	4.66 (0.75)
ARTIC-35 responses	862 (112)	5.15 (0.81)	1,092 (92)	4.89 (0.84)
ARTIC-35 on-the-job behaviour (empathy and control)	864 (110)	5.29 (0.79)	1,095 (89)	5.12 (0.73)
ARTIC-35 self-efficacy	864 (110)	5.40 (0.89)	1,095 (89)	5.27 (0.93)
ARTIC-35 reactions	865 (109)	5.09 (0.79)	1,092 (92)	4.97 (0.83)
ProQoL compassion satisfaction	926 (48)	40.13 (5.41)	1,126 (58)	38.64 (5.84)
ProQoL burnout	935 (39)	23.29 (5.54)	1,139 (45)	23.77 (5.89)
ProQoL secondary traumatic stress	928 (46)	20.54 (5.09)	1,122 (62)	20.46 (5.11)

Note. The following are national-level means from DfE statistical releases (DfE, 2025a; DfE, 2025b): percentage of pupils receiving SEND support (12.9%), percentage of pupils eligible for FSMs in the last six years (27.3%), total number of pupils on roll (all ages; 948). SD = standard deviation; ARTIC = Attitudes Related to Trauma-Informed Care, FSM = free school meals, IBS = Illinois Bully Scale, ProQoL = Professional Quality of Life, SEND = special educational needs and/or disabilities; SDQ = Strengths and Difficulties Questionnaire; SEI = Student Engagement Instrument.

Baseline imbalance

Baseline imbalances in demographic and outcome measures were assessed following thorough data cleaning, including the removal of duplicates. Baseline comparisons revealed imbalances between intervention and control groups on several key variables:

- SDQ internalising difficulties scores were significantly higher in the intervention group at baseline ($t = −3.21$, $df = 13,540$, $p = 0.001$), with a mean score of 6.13 in the intervention group and 5.92 in the control group. This indicates that pupils in intervention schools were experiencing greater internalising difficulties than pupils in control schools, although the magnitude of the imbalance was small (Hedge’s $g = 0.055$)
- The gender distribution differed significantly between groups ($\chi^2 = 40.82$, $df = 3$, $p < 0.001$), with more females (standardised residual = 5.693, $p < 0.001$) and fewer males (standardised residual = −6.233, $p <.001$) in the intervention group; however, the imbalance was small (Cramer’s V = 0.052).
- Ethnicity varied across groups ($\chi^2 = 23.51$, $df = 5$, $p < 0.001$), with more White young people (standardised residual = 3.450, $p = 0.007$) and fewer Asian young people (standardised residual = −3.812, $p = 0.002$) in the intervention group. As with the other imbalances, this imbalance was small (Cramer’s V = 0.041).

There was no imbalance in terms of year group, school-level FSM, SEND percentages, school size, SDQ externalising difficulties (pupil- or school-level) and other outcome variables. Independent t-tests and chi-square tests were conducted for all variables, and these results were non-significant (all $p > 0.05$). Pupil-level FSM and SEND were not included in the imbalance checks, as this data was collected after randomisation and there was a high level of missingness due to school non-submission of administrative data.

To account for the imbalances and reduce potential confounding, all models with pupil data included baseline internalising difficulties, gender and ethnicity as covariates. This ensured that estimated intervention effects were adjusted for pre-existing group differences in baseline internalising difficulties, gender and ethnicity.

Outcomes and analysis

Primary analysis

The primary research question was: is there a difference in mean externalising difficulties at the end of programme follow-up between pupils in secondary schools receiving MGDAS and business as usual, compared to pupils in secondary schools receiving business as usual only? The primary outcome was externalising difficulties, measured using the SDQ externalising difficulties subscale.

The key finding is that unadjusted endline means were slightly lower in the intervention group, but the adjusted model – accounting for baseline differences and covariates – estimated a small positive group effect. This means that pupils showed marginally more behavioural problems in the intervention group compared to the control group. However, these differences were very small and indicated no evidence of a statistically significant intervention effect.

Descriptive statistics indicated similar mean scores across groups at endline: the intervention group had a mean of 6.53 (95% CI: 6.38 to 6.68), while the control group had a mean of 6.65 (95% CI: 6.53 to 6.77). The distribution of endline SDQ scores is shown in Figure 2 and indicates broadly similar scores across groups.

To estimate the adjusted effect of the intervention on endline SDQ externalising difficulties, the model included fixed effects for group, baseline SDQ externalising difficulties and covariates (cohort, baseline internalising symptoms, gender, ethnicity and LA/CA). A random intercept was specified for school to account for clustering. The following model was fitted using the lmer() function in R:

lmer(endline_SDQ_extern ~ group + baseline_SDQ_extern + cohort + baseline_SDQ_intern + gender + ethnicity + LA_CA + (1 | school), data = data)

The summary output of the adjusted and sensitivity (the latter of which excludes covariates of baseline SDQ internalising difficulties, gender and ethnicity) models is presented in Table 13. Effect size estimation inputs are presented in Appendix B.

Figure 2: Counts of Strengths and Difficulties Questionnaire endline externalising scores in the analytical sample by group

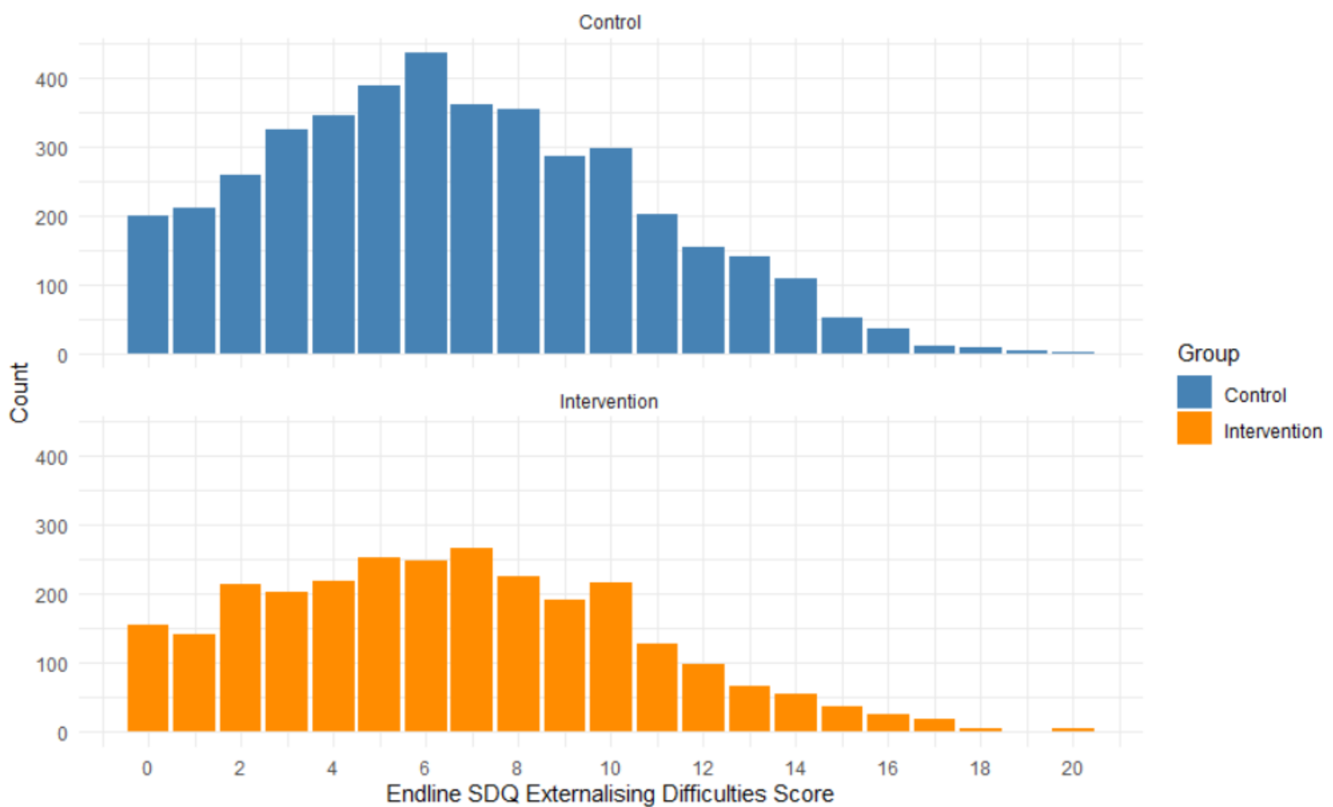


Table 13: Summary model results from the primary analysis.

	Unadjusted means				Effect size		
	Intervention group		Control group				
Outcome	**n (missing)**	**Mean (95% CI)**	**n (missing)**	**Mean (95% CI)**	**Total n (intervention; control)**	**Hedges g (95% CI)**	**p-value**
SDQ externalising difficulties (adjusted model)	2,779 (3,645)	6.53 [6.38, 6.68]	4,180 (3,296)	6.65 [6.53, 6.77]	6,959 (2,779; 4,180)	0.014 [-0.047, 0.076]	0.652
SDQ externalising difficulties (sensitivity model)	2,979 (3,445)	6.54 [6.39, 6.68]	4,420 (3,056)	6.66 [6.55, 6.78]	7,399 (2,979; 4,420)	0.005 [-0.056, 0.066]	0.874

Note. SDQ = Strengths and Difficulties Questionnaire; CI = confidence interval.

Model results indicated that the intervention group had slightly higher SDQ externalising difficulties scores than the control group (adjusted Hedge's g = 0.014, 95% CI [−0.047, 0.076], p = 0.652; sensitivity Hedge's g = 0.005, 95% CI [−0.056, 0.066], p = 0.874). This means that pupils in the intervention group reported marginally more behavioural difficulties. However, the effect sizes were very small, and the confidence intervals crossed zero, indicating no evidence of a statistically significant difference in SDQ externalising difficulties between the intervention and control groups.

The fixed effects represent the measured factors in the study (such as group, cohort, baseline scores and demographics). The random effect accounts for the fact that schools differ from one another in ways not captured by those measured factors. Fixed effects accounted for 40.2% of the variance (marginal R^2 = 0.402), while the full model, including the school random effect, explained 40.8% (conditional R^2 = 0.408). The random intercept for schools indicated low between-school variability (SD = 0.31), especially relative to the larger residual variability (SD = 3.06). The non-significant finding indicates no evidence of an impact of the intervention on pupil externalising difficulties. For full model output, please see Appendix C.

Secondary analysis

The secondary research questions for pupils are:

1. Is there a difference in mean mental health difficulties at the end of programme follow-up between pupils in secondary schools receiving MGDAS and business as usual, compared to pupils in secondary schools receiving business as usual only?
2. Is there a difference in mean prosocial behaviours at the end of programme follow-up between pupils in secondary schools receiving MGDAS and business as usual, compared to pupils in secondary schools receiving business as usual only?
3. Is there a difference in mean safe social connections with teachers at the end of programme follow-up between pupils in secondary schools receiving MGDAS and business as usual, compared to pupils in secondary schools receiving business as usual only?
4. Is there a difference in mean safe social connections with peers at the end of programme follow-up between pupils in secondary schools receiving MGDAS and business as usual, compared to pupils in secondary schools receiving business as usual only?
5. Is there a difference in mean school inclusion at the end of programme follow-up between pupils in secondary schools receiving MGDAS and business as usual, compared to pupils in secondary schools receiving business as usual only?
6. Is there a difference in mean bullying (perpetration and victimisation) at the end of programme follow-up between pupils in secondary schools receiving MGDAS and business as usual, compared to pupils in secondary schools receiving business as usual only?
7. Is there a difference in educational attendance at the end of programme follow-up between pupils in secondary schools receiving MGDAS and business as usual, compared to pupils in secondary schools receiving business as usual only?
8. Is there a difference in exclusions (fixed-term and permanent) at the end of programme follow-up between pupils in secondary schools receiving MGDAS and business as usual, compared to pupils in secondary schools receiving business as usual only?

Secondary outcome models were specified in line with the primary analysis. For the SDQ impact score, however, baseline data was intentionally not collected to reduce participant burden, so the model was adjusted to omit this variable.

Analysis of secondary outcomes showed broadly similar patterns across intervention and control groups, with all differences small and not statistically significant. For internalising difficulties, the intervention group had slightly higher scores than the control group (Hedges g = 0.042, 95% CI [−0.013, 0.098], p = 0.140), reflecting marginally greater emotional symptoms, such as anxiety or low mood. However, there is no statistically significant discernible effect and no evidence of a statistically significant difference in internalising difficulties between groups. For the SDQ impact score, intervention pupils reported a

marginally greater perceived impact of difficulties on daily life (Hedges g = 0.018, 95% CI [−0.036, 0.071], p = 0.523), but the difference was negligible and not statistically significant.

For prosocial behaviour, the intervention group showed slightly lower scores than the control group (Hedges g = −0.007, 95% CI [−0.090, 0.076], p = 0.867), indicating marginally fewer helpful or cooperative behaviours, though the difference was very small and not statistically significant. For teacher–student relationships, the intervention group reported marginally lower scores than the control group (Hedges g = −0.017, 95% CI [−0.110, 0.075], p = 0.712), suggesting slightly worse perceived relationships with teachers, but again, the difference was negligible and not statistically significant.

These findings indicate that the intervention did not produce measurable improvements in pupils' mental wellbeing, prosocial behaviours or relationships with staff compared with business as usual.

For the summary model output, please see Table 14. For full model output, please see Appendix C.

Table 14: Summary model results from the analysis of continuous secondary pupil outcomes.

	Unadjusted means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)	p-value
SDQ internalising difficulties (adjusted model)	2,778 (3,646)	5.62 [5.48, 5.76]	4,180 (3,296)	5.46 [5.35, 5.57]	6,958 (2,778;4,180)	0.042 [−0.013, 0.098]	0.140
SDQ impact score (adjusted model, no baseline outcome data or random intercept for pupils)	2,644 (3,780)	0.859 [0.796, 0.922]	4,007 (3,469)	0.800 [0.751, 0.849]	6,651 (2,644;4,007)	0.018 [−0.036, 0.071]	0.523
SDQ prosocial behaviour (adjusted model)	2,780 (3,644)	6.91 [6.83, 6.99]	4,180 (3,296)	6.83 [6.77, 6.89]	6,960 (2,780;4,180)	−0.007 [−0.090, 0.076]	0.867
SEI teacher–student relationships (adjusted model)	2,331 (4,093)	23.9 [23.6, 24.1]	3,618 (3,858)	23.8 [23.6, 23.9]	5,949 (2,331;3,618)	−0.017 [−0.110, 0.075]	0.712

Note. CI = confidence interval ; SDQ = Strengths and Difficulties Questionnaire; SEI = Student Engagement Instrument. Unless otherwise specified, the models included fixed effects for group, baseline outcome score, and covariates (cohort, baseline SDQ internalising difficulties score, gender, ethnicity and LA/CA). Random intercepts were specified for school to account for clustering.

Unadjusted proportions of pupils experiencing at least one suspension were very similar between groups (5.3% of pupils in the intervention group and 5.8% in the control group). After adjustment, the odds ratio was 1.187 and the corresponding standardised effect size, expressed as Hedges' g (Cox transformation), was very small (g = 0.095, 95% CI: −0.116 to 0.306, p = 0.379), indicating marginally higher odds of suspension in the intervention group. However, the result was not statistically significant, meaning there is no evidence of a statistically significant difference between groups.

Permanent school exclusions were rare in both groups. Due to the very low number of permanent exclusions and resulting convergence issues in multi-level modelling, this outcome was analysed using Fisher's exact test, which showed no evidence of a difference between groups. The intervention group had a slightly higher

exclusion rate (0.13%) compared to the control group (0.07%), corresponding to four exclusions in the intervention group and three in the control group. The odds ratio was 1.785, and the corresponding standardised effect size, expressed as Hedges' g (Cox transformation), was small (g = 0.319, 95% CI: −0.506 to 1.114, p = 0.470). This suggests higher odds of exclusion in the intervention group, but the wide confidence interval and non-significant result indicate no evidence of a statistically significant meaningful difference.

For the results for binary secondary pupil outcomes, please see Table 15 for the model output.

Table 15: Summary model results from the analysis of binary secondary pupil outcomes.

	Unadjusted means				**Effect size**		
	Intervention group		**Control group**				
Outcome	**n (missing)**	**Mean (95% CI)**	**n (missing)**	**Mean (95% CI)**	**Total n (intervention; control)**	**Hedge's g(95% CI)**	**p-value**
Suspensions (adjusted model)	3,158 (3,266)	0.053 [0.045, 0.060]	4,225 (3,251)	0.058 [0.051, 0.065]	7,383 (3,158;4,225)	0.095 [−0.116, 0.306]	0.379
Permanent school exclusion	3,158 (3,266)	0.0013 [0.00003, 0.0025]	4,225 (3,251)	0.0007 [0, 0.0015]	7,383 (3,158;4,225)	0.319 [−0.506, 1.145]	0.470

Note: CI = confidence interval

For tertiary outcomes, pupils in the intervention group reported marginally greater peer support at school and higher future aspirations and goals, suggesting somewhat stronger perceptions of peer social support and optimism about the future. Conversely, the pupils in the intervention group reported marginally lower scores on the control and relevance of schoolwork subscale compared to the control group. However, these differences were very small and not statistically significant, indicating no evidence of a statistically significant intervention effect.

In relation to bullying, both victimisation and perpetration scores were higher in the intervention group. For victimisation, the adjusted model showed a small positive effect size, suggesting greater pupil-reported experiences of being bullied in the intervention group. However, this effect was not statistically significant when covariates were removed from the model and after multiple comparison correction, meaning the finding should be interpreted with caution. This suggests that the observed difference may partly reflect adjustment for baseline differences, covariates or sample size rather than a consistent intervention effect. For perpetration, the intervention group also reported slightly higher scores than the control group. The effect size was negligible and not statistically significant, indicating no statistically significant evidence of a meaningful difference in bullying behaviours between groups.

For school attendance, pupils in the intervention group had slightly lower attendance rates compared to the control group. The effect size was small and not statistically significant, indicating no evidence of a statistically significant meaningful difference (see Table 16).

Table 16: Summary model results from the analysis of tertiary pupil outcomes.

	Unadjusted means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)	p-value
SEI peer support at school (adjusted model)	2,368 (4,056)	18.1 [18.0, 18.3]	3,660 (3,816)	18.1 [17.9, 18.2]	6,028 (2,368;3,660)	0.018 [−0.053, 0.089]	0.616
SEI control and relevance of schoolwork (adjusted model)	2,315 (4,109)	24.1 [23.9, 24.3]	3,602 (3,874)	23.9 [23.8, 24.1]	5,917 (2,315;3,602)	−0.013 [−0.103, 0.076]	0.776
SEI future aspirations and goals (adjusted model)	2,379 (4,045)	16.2 [16.1, 16.3]	3,688 (3,788)	16.0 [15.9, 16.1]	6,067 (2,379;3,688)	0.051 [−0.013, 0.114]	0.126
IBS bullying perpetration (adjusted model)	2,269 (4,155)	3.67 [3.44, 3.91]	3,558 (3,918)	3.67 [3.48, 3.86]	5,827 (2,269;3,558)	0.009 [−0.080, 0.098]	0.837
IBS bullying victimisation (adjusted model)	2,361 (4,063)	3.31 [3.13, 3.49]	3,702 (3,774)	2.96 [2.82, 3.09]	6,063 (2,361;3,702)	0.086 [0.017, 0.155]	0.019 (FDR-corrected = 0.057)
IBS bullying victimisation (sensitivity model)	2,525 (3,899)	3.31 [3.14, 3.48]	3,910 (3,566)	2.98 [2.85, 3.11]	6,435 (2,525, 3,910)	0.077 [0.008, 0.147]	0.035 (FDR-corrected = 0.057)
School attendance percentage (adjusted model)	3,155 (3,269)	93.5 [93.2, 93.9]	4,223 (3,253)	93.8 [93.5, 94.0]	7,378 (3,155;4,223)	−0.046 [−0.111, 0.019]	0.178

Note. CI = confidence interval; FDR = false discovery rate; IBS = Illinois Bully Scale; SEI = Student Engagement Instrument. Unless otherwise specified, the models included fixed effects for group, baseline SDQ outcome score and covariates (cohort, baseline SDQ internalising difficulties scores, gender, ethnicity and LA/CA). Random intercepts were specified for school to account for clustering.

The secondary research questions for school staff were:

1. Is there a difference in mean knowledge and awareness of trauma at the end of programme follow-up between staff in secondary schools receiving MGDAS and business as usual, compared to staff in secondary schools receiving business as usual only?
2. Is there a difference in mean confidence in working with young people who have experienced trauma at the end of programme follow-up between staff in secondary schools receiving MGDAS and business as usual, compared to staff in secondary schools receiving business as usual only?
3. Is there a difference in mean wellbeing at the end of programme follow-up between staff in secondary schools receiving MGDAS and business as usual, compared to staff in secondary schools receiving business as usual only?
4. Is there a difference in mean burnout at the end of programme follow-up between staff in secondary schools receiving MGDAS and business as usual, compared to staff in secondary schools receiving business as usual only?

5. Is there a difference in mean ability to create an emotionally safe environment at the end of programme follow-up between staff in secondary schools receiving MGDAS and business as usual, compared to staff in secondary schools receiving business as usual only?
6. Is there a difference in mean understanding of vicarious trauma and the need for staff support at the end of programme follow-up between staff in secondary schools receiving MGDAS and business as usual, compared to staff in secondary schools receiving business as usual only?
7. Is there a difference in mean empathy-focused behaviours for behaviour management at the end of programme follow-up between staff in secondary schools receiving MGDAS and business as usual, compared to staff in secondary schools receiving business as usual only?

For secondary staff outcome models, fixed effects were included for group, baseline scores, cohort and LA/CA. Random intercepts were included for school. Models followed the following specification:

lmer(endline_outcome ~ group + baseline_outcome + cohort + LA_CA + (1 | school))

For staff secondary outcomes, the majority of effects were small and not statistically significant. For summary output, see Table 17.

All ARTIC-35 subscales showed small positive effects, such as for beliefs about the underlying causes of behaviour (beliefs about why challenging behaviour occurs, including whether it is understood through a trauma-informed lens), self-efficacy (confidence in using trauma-informed practices), responses to problem behaviour (the extent to which responses to behavioural difficulties reflect trauma-informed principles), reactions to work (positive feelings about one's role), and empathy and control in on-the-job behaviour (balancing empathy with maintaining boundaries). These differences suggested slightly more positive attitudes among staff in intervention schools, but none were statistically significant.

On the ProQoL scale, staff in intervention schools reported slightly higher compassion satisfaction (greater fulfilment from helping others) and slightly lower burnout (less exhaustion) and secondary traumatic stress (stress from exposure to others' trauma), but these effects were not statistically significant.

Table 17: Summary model results from the analysis of secondary staff outcomes.

	Unadjusted means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)	p-value
ARTIC-35 underlying causes	223 (751)	4.88 [4.78, 4.98]	440 (744)	4.61 [4.54, 4.68]	663 (223;440)	0.123 [−0.019, 0.265]	0.098
ARTIC-35 self-efficacy	222 (752)	5.48 [5.36, 5.59]	439 (745)	5.34 [5.26, 5.43]	661 (222; 439)	0.132 [−0.037, 0.301]	0.132
ARTIC-35 response to problem behaviour	222 (752)	5.23 [5.10, 5.35]	439 (745)	4.87 [4.79, 4.95]	661 (222; 439)	0.124 [−0.036, 0.285]	0.139
ARTIC-35 reactions to work	223 (751)	5.17 [5.05, 5.29]	437 (747)	5.01 [4.93, 5.09]	660 (223; 437)	0.049 [−0.120, 0.219]	0.572

ARTIC-35 empathy and control (on-the-Job Behaviour)	223 (751)	5.46 [5.36, 5.56]	439 (745)	5.20 [5.13, 5.27]	662 (223; 439)	0.116 [−0.045, 0.278]	0.167
ProQoL compassion satisfaction	234 (740)	39.10 [38.4, 39.8]	437 (747)	38.0 [37.5, 38.6]	671 (234; 437)	0.048 [−0.083, 0.178]	0.474
ProQoL burnout	230 (744)	23.4 [22.7, 24.2]	436 (748)	24.0 [23.4, 24.5]	666 (230; 436)	−0.101 [−0.249, 0.046]	0.195
ProQoL secondary traumatic stress	229 (745)	20.6 [19.9, 21.2]	433 (751)	20.2 [19.7, 20.7]	662 (229; 433)	−0.011 [−0.171, 0.148]	0.893

Note. CI = confidence interval; ARTIC = Attitudes Related to Trauma-Informed Care; FDR = false discovery rate; ProQoL = Professional Quality of Life.

Analysis in the presence of non-compliance

The overall finding of this section is that a compliance analysis using a CACE framework showed that, as with the primary analysis, the intervention group had slightly higher SDQ externalising difficulties scores than the control group. However, there was no statistically significant difference, suggesting that non-compliance with the training programme may not have masked a potential impact.

Attendance data was available for 25 of the 31 schools originally recruited, as six schools withdrew from the intervention prior to the midline point. Average staff attendance across sessions varied considerably between schools, ranging from 5.7% to 65.7%. The mean attendance was 26.6%, with a median of 20.4%. This indicates that, while overall compliance was low, there was substantial variation across sites. Since no schools met the pre-specified 80% criterion, the upper quartile of staff attendance among intervention schools was used as the threshold for compliance in the CACE analysis (41.7%).

A linear regression model was used to predict compliance status based on randomisation group and other school-level variables (i.e. school size, school FSM percentage, school SEN percentage and mean school SDQ externalising difficulties at baseline), producing predicted probabilities of compliance for each school. In the second stage, we fitted a multi-level model, regressing endline SDQ externalising difficulties onto predicted compliance, baseline SDQ externalising difficulties and the covariates used in the primary model. The results of the CACE analysis are presented in Table 18, with the results from the primary analysis also provided for comparison.

For the CACE model, the estimated effect among pupils in schools that complied with the intervention was small (Hedges' g = 0.123), indicating that compliers showed a small increase in externalising difficulties relative to what they would have shown under control conditions. However, the 95% confidence interval crossed zero, and the result was not statistically significant. This indicates that there is no discernible statistically significant effect and no evidence of a statistically significant difference in externalising difficulties. Therefore, the CACE analysis shows that low compliance is unlikely to explain the lack of a significant main effect. However, it is acknowledged that no school reached the pre-specified 80% compliance threshold, so the analysis relied on a lower data-driven threshold (the upper quartile of staff attendance: 41.7%). As a result, the CACE estimate reflects differences between relatively low compliance groups rather than high compliers. This limits the extent to which CACE can meaningfully test a high-compliance effect. The full model output is presented in Appendix C.

Table 18: Summary model results from the CACE analysis and the primary adjusted analysis.

	Unadjusted means				Effect size		
	Intervention group		Control group				
Outcome	**n (missing)**	**Mean (95% CI)**	**n (missing)**	**Mean (95% CI)**	**Total n (intervention; control)**	**Hedges g (95% CI)**	**p-value**
SDQ externalising difficulties (CACE model)	2,779 (3,645)	6.53 [6.38, 6.68]	4,180 (3,296)	6.65 [6.53, 6.77]	6,959 (2,779;4,180)	0.123 [−0.152, 0.399]	0.380
SDQ externalising difficulties (adjusted model)	2,779 (3,645)	6.53 [6.38, 6.68]	4,180 (3,296)	6.65 [6.53, 6.77]	6,959 (2,779;4,180)	0.014 [−0.047, 0.076]	0.652

Note. CI = confidence interval; SDQ = Strengths and Difficulties Questionnaire.

Missing data analysis

Missing data analysis for the primary analysis model was conducted to clarify the missing data mechanism and support valid inferences. Overall, it was found that missing data analyses (including MCAR tests, such as predictors of missingness in the model, MI and non-response weighting) produced estimates aligned with the complete case analysis, suggesting that the non-significant effect is unlikely to be explained by biases arising from missingness.

Little's (1998) MCAR test suggested that the data was not MCAR ($\chi^2(117) = 1{,}259$, $p < 0.001$). Therefore, an additional analysis of the missingness mechanism was undertaken. The amount of missing data in the primary outcome (endline SDQ externalising difficulties) and covariates for pupils who had any baseline data is presented in Table 19.

Table 19: Count and percentage of missing data in the primary outcome and covariates for pupils who had any baseline data.

Variable	N Missing	% Missing
Endline SDQ externalising difficulties	6,486	46.66%
Ethnicity	671	4.83%
Gender	275	1.98%
Baseline SDQ externalising difficulties	55	0.40%
Baseline SDQ internalising difficulties	56	0.40%

Note. SDQ = Strengths and Difficulties Questionnaire.

Data was more than 5% missing for the primary outcome of endline SDQ externalising difficulties. Therefore, multi-level logistic regression (including school as a random intercept) was conducted to identify variables that predict missingness.

Missing primary outcome data

Missingness in the primary outcome was associated with pupil year group, pupil-level SEND status, baseline SDQ externalising difficulties, attendance, suspensions, school size and LA/CA (see Appendix C for the full model output).

Pupils in Year 9 at the start of the trial were more likely to have missing endline SDQ externalising symptoms than those in Year 8. Pupil-level SEND was predictive of missingness, and higher baseline SDQ externalising difficulties scores were positively associated with missingness. Lower attendance during the 2023/24 academic year was linked to increased missingness, and pupils who had experienced suspensions were also more likely to be missing. Smaller school size was associated with higher levels of missingness, and LA/CA differences were evident.

Based on these findings, year group, pupil-level SEND status, attendance, suspensions and school size were included as additional covariates in the primary analysis model to account for differential missingness. Please note that the remaining predictors of missingness were already present in the model.

The output from the model is presented in Table 20. The estimated effect size in the primary model was very similar to that in the model adjusted for covariates associated with missingness. Both estimates were small and non-significant, suggesting that the inclusion of these covariates did not meaningfully alter the primary findings.

Table 20: Summary output from the primary model with covariates that predict missingness included, alongside output from the adjusted primary model.

	Unadjusted means				Effect size		
	Intervention group		Control group				
Outcome	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)	p-value
SDQ externalising difficulties (adjusted missingness model)	2,460 (3,964)	6.41 [6.26, 6.57]	345,2 (4,024)	6.60 [6.47, 6.73]	5,912 (2,460;3,452)	0.013 [−0.057, 0.084]	0.711
SDQ externalising difficulties (adjusted primary model)	2,779 (3,645)	6.53 [6.38, 6.68]	4,180 (3,296)	6.65 [6.53, 6.77]	6,959 (2,779;4,180)	0.014 [−0.047, 0.076]	0.652

Note. CI = confidence interval; SDQ = Strengths and Difficulties Questionnaire.

Missing covariate data

As year group (6.08%), pupil-level SEND status (34.40%), suspensions (34.40%) and attendance (35.25%) were predictive of missingness in the primary outcome and had more than 5% missing data, predictors of missingness in these variables were also assessed in separate models to provide evidence that this data was MAR and was suitable for MI.

Missingness in pupil-level SEND status, attendance and suspensions was largely due to missing school data submissions (indeed, these variables were 100% missing in schools that did not provide a submission). Therefore, to help inform the imputation model, a logistic regression to predict school data submission was performed using school-level characteristics.

Average school baseline externalising difficulties negatively predicted school data submission (estimate = −1.11, SE = 0.475, p = 0.019), and school FSM percentage positively predicted school data submission (estimate = 0.063, SE = 0.032, p = 0.049). Therefore, these variables were added as predictors into the imputation predictor matrix, together with pupil characteristics (gender, ethnicity, year group), school ID and a binary indicator of whether the school had provided a school data submission.

Continuous covariates (e.g. attendance) were imputed using two-level normal models (2l.norm), while binary covariates (e.g. SEND status, suspensions, year group) were imputed using two-level logistic models (2l.bin). Analyses were conducted across the imputed datasets, with pooled estimates reported. The effect size resulting from the imputation model was similar to the primary model and not statistically significant (Hedge's g = 0.018, 95% CI [-0.044, 0.079], p = 0.571).

Non-response weighting

Non-response weighting was applied to account for biases resulting from differential participation of groups in completing the primary outcome. Weights were computed by first conducting a multi-level logistic regression to predict the probability of response of the primary outcome based on covariates (gender, ethnicity, baseline internalising symptoms, baseline externalising symptoms and school characteristics) and then calculating the inverse of the response probability (Valliant et al., 2013). Weights were then normalised and included in the primary analysis model. Effect sizes were similar to results from the primary model (Hedge's g = 0.010, 95% CI[−0.038, 0.059], p = 0.678).

Subgroup analyses

Exploratory subgroup analyses showed that most comparisons did not indicate differential intervention effects. Although there was some indication of an interaction suggesting a potentially different effect for Black pupils, this pattern was not statistically robust once adjustments for multiple comparisons were applied. It is important to note that this analysis was exploratory, rather than confirmatory or causal. Summary output is presented in Table 21, and full model output is presented in Appendix C.

For the ethnicity subgroup analysis, effects were generally small and non-significant across most groups. The only statistically significant interaction was between intervention group and Black ethnicity relative to the reference categories (control group and White ethnicity). This interaction indicated higher externalising scores for Black pupils in intervention schools (Hedge's g = 0.159, 95% CI: 0.011 to 0.308, p = 0.036). However, the number of Black pupils was relatively small (n = 256 in the control group; n = 231 in the intervention group), the confidence interval was wide and the p-value did not survive multiple comparison correction, suggesting considerable uncertainty around this estimate. This finding should therefore be interpreted with caution. For pupils who identified as White, Asian or mixed ethnicity, SDQ externalising difficulties scores were higher in the intervention group compared to the control group, although effect sizes were small and not statistically significant. The direction of the effect was reversed for young people who described their ethnic group as other – there were slightly lower externalising difficulties in the intervention group compared to the control group – but this was not a meaningful difference, as it was not statistically significant.

For young people who were eligible for FSMs or who had SEND, the effect size was negative, which indicates that SDQ externalising difficulties scores were lower in the intervention group compared to the control group, but this was not statistically significant. For those who were not eligible for FSMs or who did not have SEND, the effect size was reversed, but again, this was not a statistically significant difference.

Table 21: Summary output from the exploratory subgroup analyses by ethnicity, FSM eligibility and SEND status.

	Unadjusted means				Effect size		
	Intervention group		Control group				
Subgroup	n (missing)	Mean (95% CI)	n (missing)	Mean (95% CI)	Total n (intervention; control)	Hedges g (95% CI)	p-value
Ethnicity							
White	1,354 (2,409)	7.08 [6.87, 7.30]	2,296 (1,866)	7.21 [7.05, 7.37]	3,650 (1,354; 2,296)	−0.005 [−0.083, 0.074]	0.909
Asian	929 (494)	5.71 [5.46, 5.97]	1,252 (610)	5.69 [5.48, 5.90]	2,181 (929; 1,252)	0.021 [−0.066, 0.108]	0.637
Black	231 (167)	6.27 [5.77, 6.77]	256 (199)	5.88 [5.41, 6.34]	487 (231; 256)	0.159 [0.011, 0.308]	0.036 (FDR-corrected = 0.180)
Mixed	124 (137)	7.05 [6.33, 7.77]	223 (145)	6.80 [6.29, 7.31]	347 (124; 223)	0.044 [−0.133, 0.222]	0.625
Other	141 (113)	6.57 [5.98, 7.16]	153 (130)	7.16 [6.51, 7.82]	294 (141; 153)	−0.128 [−0.314, 0.058]	0.176
FSM eligibility							
Yes	764 (464)	6.82 [6.53, 7.10]	986 (507)	7.26 [7.01, 7.51]	1,750 (764; 986)	−0.035 [−0.127, 0.057]	0.460
No	1696 (845)	6.23 [6.04, 6.42]	2,510 (1036)	6.35 [6.20, 6.51]	4,206 (1,696; 2,510)	0.015 [−0.060, 0.090]	0.685
SEND status							
Yes	308 (243)	7.43 [7.00, 7.86]	505 (367)	8.03 [7.68, 8.39]	813 (308; 505)	−0.059 [−0.183, 0.064]	0.346
No	2,152 (1,066)	6.27 [6.10, 6.44]	3,036 (1,441)	6.36 [6.22, 6.50]	5,188 (2,152; 3,036)	0.018 [−0.053, 0.089]	0.628

Note. The outcome for all subgroup analysis models was SDQ externalising symptoms. CI = confidence interval; FSM = free school meals, SEND = special educational needs and/or disabilities.

Additional analyses and robustness checks

Primary model with midline data included

The primary model was modified to factor in midline externalising difficulties scores. This was a complete case analysis, as full information maximum likelihood was not available in the package lmer, and an alternative single-level FIML model (e.g. using lavaan) was not deemed appropriate, given the clustered

nature of the data. The interaction term between group (intervention) and time (endline vs midline) was not statistically significant (estimate = 0.140, $p = 0.117$). Effect sizes were non-significant (midline Hedge's g = −0.023, 95% CI: −0.071 to 0.024, $p = 0.334$; endline Hedge's g = 0.010, 95% CI: −0.038 to 0.059, $p = 0.672$). Full output is shown in Appendix C.

Moderation analysis within the primary analysis

In the primary analysis model, a cross-level interaction between group allocation and mean school-level scores on the staff-reported on-the-job behaviour (empathy and control) subscale of the ARTIC-35 was included to examine whether allocation to treatment vs control groups predicts different levels of externalising difficulties for pupils in schools where staff report higher levels of empathy-focused behaviours. There was no significant moderation effect found (intervention group x mean ARTIC on-the-job behaviour (estimate = 0.155, SE = 0.361, $p = 0.668$; Hedge's g = 0.034, 95% CI: −0.031 to 0.0999, $p = 0.302$). Full output is shown in Appendix C.

Longitudinal mediation analysis within the primary analysis

A longitudinal mediation analysis was conducted to test whether SEI teacher–student relationships at midline mediate the relationship between group allocation and changes in externalising difficulties from baseline to endline, accounting for clustering by schools. For the mediation model, the package *lavaan* (version 0.6-17) in R was used (Rosseel, 2012). Due to convergence issues, covariates were removed from the model. There was no statistically significant indirect effect of SEI teacher–student relationships (estimate = −0.015, SE = 0.110, $p = 0.891$) or the total effect (estimate = −0.171, SE = 0.295, $p = 0.561$). However, it is noted that the path from teacher–student relationships at midline to externalising difficulties at endline was statistically significant (estimate = −0.24, $p < 0.001$). This indicates that pupils who reported more positive relationships with teachers tended to show lower externalising difficulties over time. This pattern is consistent with the programme's logic model, which assumes that stronger relational climates are associated with improved behavioural outcomes. However, because there was no evidence that intervention produced measurable changes in teacher–student relationships, this pathway did not translate into a significant indirect effect. Full output is shown in Appendix C.

Estimation of intraclass correlation coefficient

ICCs for SDQ externalising difficulties scores were calculated separately at baseline and endline using the package rpt in R. Only school-level ICCs were estimated due to each pupil contributing a single observation. Adjusted ICCs are reported with 95% confidence intervals.

- ICC baseline school: 0.011, 95% CI [0.005, 0.019]
- ICC endline school: 0.010, 95% CI [0.003, 0.019]

School-level ICCs were very low at both baseline (0.011) and endline (0.010), indicating minimal outcome variation attributable to school differences.

Implementation and process evaluation results

Below we present the findings in the following sections: Implementation, Quality and experience, Perceived need, Barriers and facilitators, Changes, and Long-term impact and sustainability.

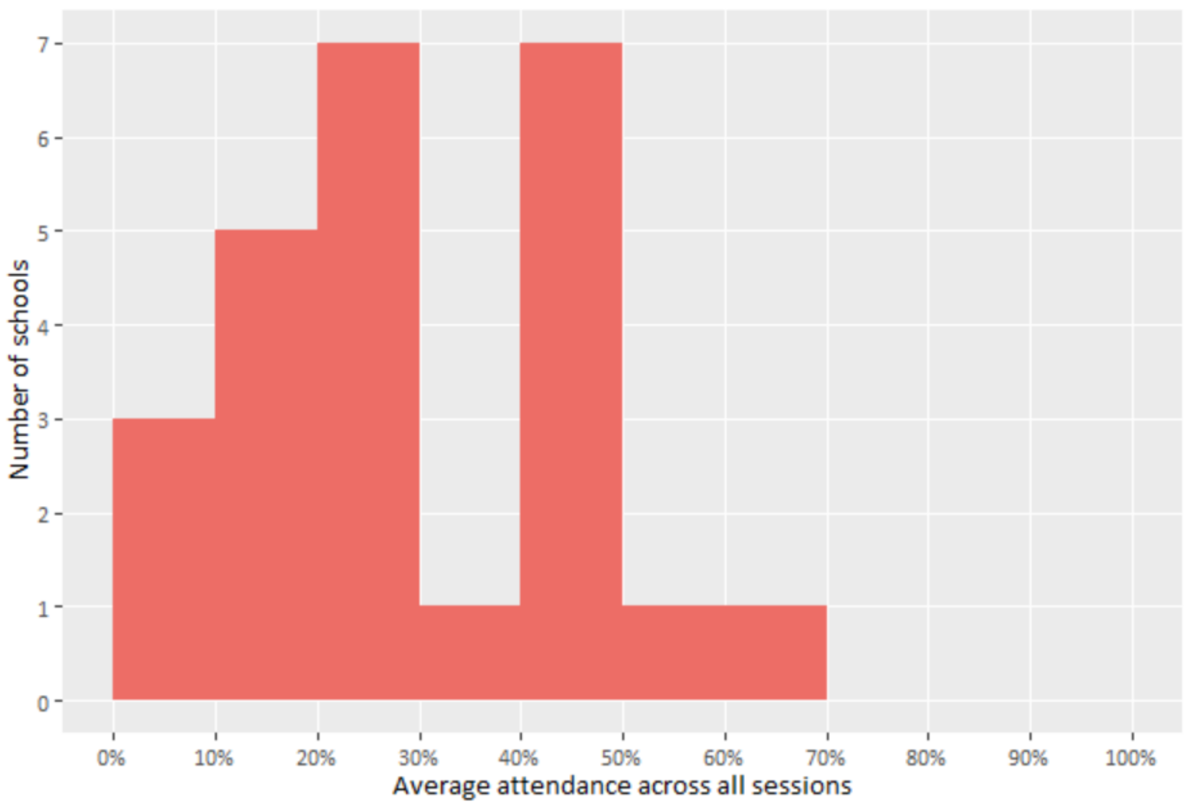
Implementation

Summary of key points in this section:

- The IPE research questions addressed in this section are: 1. To what extent is the MGDAS programme implemented as intended? and 2. Are some elements of the MGDAS programme implemented more successfully than other elements?
- Attendance data indicates that overall compliance with the programme was modest; however, this varied across sites.
- Whole-school sessions were consistently less well attended than sessions for specific staff (such as those on SLTs and PI leads).
- Most schools were offered every session, as intended. Many schools asked for variations in the way sessions were delivered, such as reduced session length or split sessions.
- Sessions were generally delivered online, with a small number being delivered in person to a few schools.
- Training was adapted to school schedules, leading to changes in sequencing and, at times, long gaps between sessions.

Attendance data was available for 25 of the 31 schools originally recruited, as six schools withdrew from the intervention prior to the midline point. Average staff attendance across all sessions varied considerably between schools, ranging from 5.7% to 65.7%. The mean attendance was 29.0%, with a median of 26.6%. This indicates that while overall compliance was modest, there was substantial variation across sites.

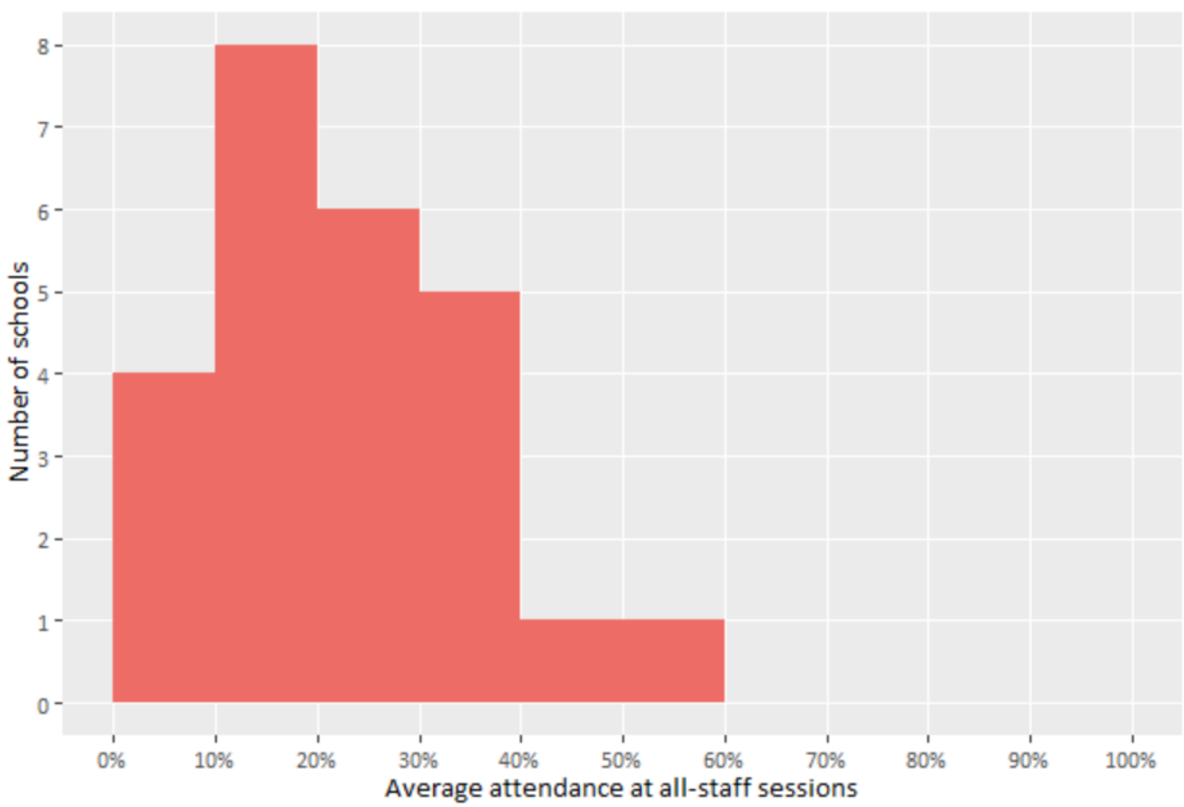
Figure 3: Histogram of average session attendance rates across all sessions



The attendance data indicates that some elements of the MGDAS programme were implemented more successfully than others. Whole-school sessions were consistently less well attended, reflecting lower levels of whole-school engagement for these sessions. By contrast, sessions for specific staff (SLTs and PI leads) generally had higher levels of attendance, although by nature of these sessions, fewer staff were needed in order to achieve full attendance.

Across the three whole-school elements, average attendance was consistently low. The original threshold for high compliance was 80% attendance; however, only one school met the threshold of above 80% compliance in at least one of the whole-school sessions (maximum = 80.5%). Mean attendance across schools on these elements was in the range 17–30%, with median values of 15–27%. In total, eight schools achieved attendance above 50% in at least one of these elements. This shows that although a small number of schools managed to engage more than half of their staff members in these sessions, the majority of schools struggled to secure wider staff participation. Note that 15 whole-school sessions across 12 schools were recorded in response to low or zero attendance, and there was no record available of the number of staff who watched the recordings. Therefore, the numbers above may underestimate true attendance.

Figure 4: Histogram of average session attendance rates across all-staff sessions



The attendance data indicated that the sessions targeted at SLT and PI staff appeared more successful than the whole-school elements. Mean attendance across schools at these sessions was in the range 24–49%, with median values of 17–50%. Variation within schools was high: 18 schools achieved attendance above 50% in at least one SLT/PI element, but all 25 schools had 0 recorded attendees for at least one SLT/PI session. Several components within this strand achieved strong engagement, with many schools reporting attendance levels well above 50% of eligible staff: Community resilience sessions that brought SLT and PI together, as well as reflective practice consultation groups for PI staff, were particularly well attended. Policy and practice review sessions for SLTs also had consistently high participation across multiple sites. Later or repeated iterations of the reflective practice consultation groups, especially those combining SLTs and PIs, were less successful, often drawing low numbers or not being scheduled at all.

Figure 5: Histogram of average session attendance rates across senior staff sessions

The training was largely offered to schools as intended, with 22 out of 25 intervention schools being offered every session and the remaining three schools missing one or two sessions each. These missed sessions occurred due to schools not engaging with the programme. Variations in the sessions that were offered to schools were agreed with schools and included reductions in session length by up to a third, splitting of a single session's material over several sessions and sessions being recorded in case of low or zero attendance. The analysis below only includes intervention schools in the analytical sample (i.e. those that completed all rounds of data collection).

Aspects of the programme that were not delivered as intended

All but three out of 25 intervention schools had all sessions delivered. Two schools had all but two of the sessions delivered, and one school had all but one session delivered. Note that this does not include sessions that were delivered to very few or no attendees. It also does not count senior staff consultation groups that were held fewer than the intended number of three times: this happened in 12 out of 25 intervention schools.

In 12 out of 25 intervention schools, at least one session was delivered in less time than originally planned at the request of the school. In these instances, time for discussion or reflection was cut, and delivery of the core knowledge base was prioritised.

In 15 sessions across 12 out of 25 intervention schools, a session was recorded for staff to watch at a later date; often, this was done in response to there being few or no attendees. As we do not have a record of the number of staff who watched these sessions, the attendance counts for these sessions are based purely on the number of completed attendance forms that were matched to the session. It is possible that this means we have underestimated compliance. It should also be noted that we do not know if attending a session and watching a video resulted in different levels of engagement with the material.

Aspects of the programme delivery that were enhanced or customised for schools

In 20 out of 25 intervention schools, at least one of the all-staff sessions was split into two or three shorter sessions. Most sessions were delivered virtually, but five sessions across four schools were delivered in person.

In 14 out of 25 intervention schools, two separate series of reflective practice consultation groups were held for SLTs and PI leads; in 10 schools, these were pooled between these two groups. In one school, only the SLT attended the consultation groups.

In seven out of 25 intervention schools, bespoke catch-up or Key Concepts Consultancy sessions were delivered for senior staff who had moved into their roles midway through the programme or had missed previous sessions.

In three out of 25 intervention schools, there were small modifications to the order of the sessions for senior staff.

A further implementation issue related to the sequencing and timing of training delivery. The programme was originally designed for whole-school core training to be delivered at the outset, followed by targeted sessions for SLTs and PI leads. However, delivery had to be adapted to fit school calendars and available training days. This resulted in changes to the intended order of sessions and long gaps between whole-school training sessions. These extended intervals, combined with the need in some schools to shorten sessions or deliver them as recordings, may have reduced continuity and limited opportunities for staff to embed learning and practise new approaches.

Quality and experience

Summary of key points in this section:

- IPE research questions addressed in this section are: 7. What is the experience of the MGDAS programme for implementers and school staff?
- Data presented here is mainly from qualitative insights, with some quantitative data from the SDSs.
- In interviews, school staff described that they generally enjoyed the MGDAS training and found it useful.
- Some staff had not experienced trauma-informed practice training before, but those who had felt that MGDAS was still useful for consolidating their existing knowledge and building a shared understanding between colleagues.
- Staff could recall many areas of the content of the training, such as emotion coaching and resilience sessions, and felt these were important for their work.
- Most staff described that they enjoyed the interactive nature of the training. They appreciated being able to discuss and reflect on their views and experiences, although some staff felt the breakout rooms could be uncomfortable.
- Both school staff and MGDAS implementers recognised that sessions could feel long, although they acknowledged that this was necessary for the volume of content that needed to be covered.
- Implementers made suggestions for future iterations of the programme, such as reducing the ask from schools and increasing the lead-in time for better relationship-building with senior staff.
- Content in this section was coded under the themes of feedback on the training and suggestions for improvement. Subsidiary themes are indicated in bold.

School staff gave their feedback on session content. They described that they found the MGDAS training enjoyable and that they felt like they **could use what they had learnt to help them make a difference for their pupils**: *"I think I left that meeting feeling… like I said, I was overwhelmed with the amount of detail there, but I was also taken back with, 'There is so much more that I can do for the kids'"* (School staff).

Some staff described **learning about *"the overall idea of trauma-informed practice"*** for the first time and how to start *"recognising kids [who] have gone through the trauma"*. They saw this as useful, as often they had not considered these principles before: *"We've never really been trained as much on how to spot it and how we deal with it. And that's one thing that really interested me, so that's what stuck out to me"* (School staff). For other staff, the training appeared to give them a **shared language**, as they realised that they had been using trauma-informed principles without necessarily knowing what the practices were called: *"I feel that it's what we do in our roles. And what I mean by that is I didn't know that's what it was called, but when it's discussed, you think, 'Oh, that's what we do'"* (School staff).

Relevant insights from quantitative data

School staff members' feedback about having or not having received training before aligns with data from the SDS. At baseline, 74% of intervention schools identified having training in 18 months prior to the intervention on understanding how past experiences (e.g. attachment, trauma, adversity) can affect pupil mental health and wellbeing. At endline, all of the schools reported having this training, which shows at the descriptive level that the intervention had the intended effect of providing training to staff to understand the processes and consequences of trauma.

In interviews, staff who had undertaken similar training prior to the MGDAS programme felt that they already had some understanding of the principles of trauma-informed practice, varying from beginner level to more advanced. Staff with existing knowledge perceived that the sessions were useful for **building understanding within the wider staff body** so that all staff members had a shared knowledge and knew *"why"* there was a need for the training. Staff mentioned that this was especially important in gaining the *"support of the whole school"* for making changes at the "*policy level".* The sessions were also considered to be useful for learning strategies for addressing trauma in young people and *"ways of managing it and planning for it"*. One staff member, although they saw the value of trauma-informed training, did voice concerns that sometimes the practice went *"too far",* and pupils were being afforded *"protection"* from their behaviour in school, which they wouldn't receive *"outside in the community"*. Staff shared how, due to the volume of mandatory training that they are required to do, they sometimes had **difficulty remembering the content** they had learnt, which included that from the MGDAS programme sessions: *"The amount of stuff I learnt this year, I've forgotten more than I remember, if I'm perfectly honest"* (School staff). School staff also mentioned that the gaps between sessions could sometimes mean that they forgot the content from one session to the next: *"There's so much going on all the time, it's difficult to retain that information from one [session] to the other"* (School staff). A minority of staff members struggled to remember specific details about the training sessions, such as what they were called. Some staff alluded to a lack of clarity over what was MGDAS training and what was not, with one staff member stating: *"I don't actually know whether the training I did was with this or not because it was a colleague that was signed up for that. I think I did a session. I can't remember, sorry"* (School staff).

Most staff, however, were able to **recall the content covered in training sessions**, such as *"brain science"* and *"the hand brain model"*. This included the way in which trauma *"affects the brain"* and can therefore impact behaviour in young people, which staff saw as being imperative to learn.

Sessions on *"emotional coaching"* and "*resilience"* were similarly well-received. Staff remembered the importance of emotion coaching in *"co-regulation and validating the emotions"* of young people, and how to support *"dysregulated"* pupils. These sessions were described as important in learning about feelings *"from a child's perspective"*, and were relevant to what many staff do in their roles every day. Staff also referenced that for many of their pupils, particularly those with *"additional needs"*, emotional regulation and resilience were not often practised, so these concepts felt especially relevant. Some staff talked about how they found the training on emotion coaching useful, as it gave them a new outlook on how to help pupils manage their emotions:

> *"I think it's just a different way of thinking about your interpretation of how you can actually just listen, without having to speak. It's like that silence, and it's about filling the silence. You don't need to do that. I think, sometimes, that can be more powerful, and I think that's what I took away from it"* (School staff).

School staff recalled learning about the impact of *"adverse childhood experiences"* on their pupils' emotions and behaviour. Staff shared that they did not previously realise that many different experiences could create trauma for young people and the extent to which it could affect them.

> *"I think it was I didn't realise there were so many of them [ACEs] and it was the impact that they could have. So, like now, like looking back and reflecting and knowing more about the children, because last year was like my first year in my role, and this year's my second year. And now I've got to know the children a bit more, it's understanding why they are the way that they are, because of past experiences that they might have been through or witnessed or seen"* (School staff).

School staff described learning about the concept of *"reflective practice"* and how to use this in their workplace to support their colleagues in the context of trauma in the form of a *"support session"*. Fewer staff mentioned learning about concepts of relational practice, including *"Five to Thrive"*, although one staff member said that the content on *"relationships and forming attachments"* stood out to them the most, especially regarding how staff can model positive relationships for pupils: *"That generated an interesting conversation about how, actually, it's not always the interactions that we're doing directly with the children, but just by being in an environment where they're seeing positive relationships and seeing us talk to each other"* (School staff).

Some staff had experienced sessions on **policies and procedures**, which were received positively by one senior staff member, but as a source of frustration for a teacher who was not a member of their school SLT, as it highlighted areas in their school policy that were lacking, but they were personally not in a position to impact change: *"We talked about the policy, but I wasn't at a level where I can change policy. I can only take [it] back to the team, but I'm not leadership. So, maybe that one wasn't helpful to me because it just frustrated me"* (School staff).

School staff gave feedback on the format and structure of the programme. They appreciated the **interactive nature of the training** and being able to have group discussions during the sessions. Some staff felt these interactive elements were more interesting than other parts of the learning due to the way that they were delivered:

> *"I remember some of them were quite didactic, and they just talked at [us] a lot. And actually, that feedback from the staff there was, so that was less effective. It would almost be that it was better when you might talk, present, right in your teams now discuss, actually in their feedback from there.*

So, where it became interactive, there was far more engagement and actually better feedback from that as well" (School staff).

Staff described that the **reflection sessions** within the training gave them the opportunity to receive *"supervision"*. They shared that this helped them to feel acknowledged and listened to, especially when reflecting on a real-life example of something that had occurred in school: *"I remember the session on reflective practice because I brought a case study to that, and that was really quite a cathartic process for me because I felt like it was the first time I'd had supervision, proper supervision"* (School staff).

Similarly, staff recalled that **breakout rooms and reflective spaces** in small groups were often a key part of training sessions. Common feedback from staff was that this was valuable in helping them consolidate some of the information they had covered and consider how to apply this to their own work practices during certain situations.

> *"I think one of the things that I found useful in the sessions is when we had the little breakouts. So, we did a bit of theory and [talked about it, and then when we had a breakout, we would chat to colleagues about specific incidents of things that had occurred and how we might have tried to use those strategies or how we might try to use those strategies next time something similar happens. So, I think that being able to put… taking it from the theory and putting it into the practice"* (School staff).

Staff also felt this time for reflection and discussion *"resonated more"* with them and allowed everyone the opportunity to share their thoughts: *"It was nice to hear other people's problems, conundrums, debates, things that they were thinking of for next steps"* (School staff). Although some staff found the breakout rooms less useful, they recognised that their colleagues may not share their sentiment: *"I don't think that's as effective, but you know, you may speak to 50 other people and they might say, 'Hey, that was fantastic,' but for me, I wasn't particularly a fan of that"* (School staff).

In terms of the structure of the programme, many staff members mentioned that they found some of the sessions *"quite long"* or ***"too lengthy",*** which sometimes made it more difficult to stay engaged, especially when the training was delivered at the end of the school day. They suggested that the sessions could have been *"more digestible"* if delivered in more frequent, shorter sittings.

> *"I think what I'd say is it was a lot to take in. So I think, maybe, it could have been chunked or broken down a little bit more for us to process it and then move on to another piece of the topic at a later date. So maybe, if it chunked a little bit better rather than it all being there at once"* (School staff).

However, staff also recognised that, to a certain extent, the programme needed to be delivered in long blocks due to the number of sessions and volume of content that was on the schedule: *"Maybe lengthy, maybe it was too lengthy, but I know that it had to be, then otherwise we'd have been having one every week for an hour"* (School staff). They also appreciated the option to review content following the sessions: *"The resources are amazing. I'm going to admit, yes, I've not had time to sit down with all the resources, but I do intend to read through all of them"* (School staff).

Implementers echoed this and suggested **reducing the ask for schools** participating in the programme in future iterations to make the programme *"slightly less onerous and demanding"* for schools. They suggested streamlining the programme content, such as by reducing the repetition of content and reducing the

number of training sessions. They also suggested implementing the programme over a longer period of time to reduce the demand on schools:

> *"We might go, 'Year one, it's just the leaders. Year two, it's just whole-school training,' and then, maybe, work alongside over that two- or three-year period, working with schools on their policies and their practices and their structures"* (Implementer).

Perceived need

Summary of key points in this section:

- IPE research questions addressed in this section are: 6. What is the perceived need for the intervention amongst MGDAS implementers and school staff?
- In interviews, implementers described MGDAS as being needed to change school staff members' knowledge and practices around managing pupils' behaviours.
- School staff similarly described a need to develop a shared understanding of trauma-informed practice. Some staff described less of a need for the programme if trauma-informed practice was already embedded in the school's culture.
- Staff highlighted the importance of the systematic and evidence-based approach of MGDAS, and the provision of knowledge in helping staff to support pupils in diverse and vulnerable communities.
- Both implementers and staff talked about the need to support staff wellbeing.
- Content in this section was coded under themes based on what implementers and school staff thought was the need for the training. Subsidiary themes are indicated in bold.

Implementers described the programme as advocating for a shift to *"trauma-informed practice and relational approaches to teaching"* to ensure that **school is a safe place for pupils, parents and carers, and staff**.

From implementers' perspectives, the programme aimed to change school staff members' knowledge and practices around managing pupils' behaviours. This involved supporting school staff in **developing a non-judgmental, compassionate and curious approach to managing pupils' behaviours** and understanding what might be behind them.

> *"If I understand trauma-informed practice, I understand that that behaviour isn't just bad behaviour. It's an expression of need, and it's an expression of something else that's going on, and I need to stop and reflect on that"* (Implementer).

Implementers indicated that **supporting staff wellbeing** was a key part of the programme's ethos, as to be able to support pupils' wellbeing, staff needed to feel supported in terms of their own wellbeing too: *"The idea here is to help staff to understand that a spotlight on their own resilience is really important"* (Implementer). Staff also saw the promotion of **support for staff** as a key reason that MGDAS was needed.

Other aspects of the MGDAS training session content that school staff implied were particularly needed included its **systematic approach and scientific basis**, provision of a **shared understanding** among staff of trauma-informed practice, and provision of knowledge about how to **help pupils in vulnerable communities** – such as those at risk of violence – with their mental health and behaviour.

> *"What I'm seeing is an increase of stuff that's happening at home that's being brought into school that we're having to deal with. And in my area, absolutely, because it's all relatable, and now you kind of have an idea how to deal with certain situations. I had an idea beforehand, but with this extra training and understanding, it makes my job a lot easier"* (School staff).

However, **not all staff members saw the training sessions as needed**, for example, if their school culture already aligned with the principles of the training, although the training sessions were still seen as beneficial for substantiating the trauma-informed approaches that schools were already taking:

> *"I don't know if I would use the word 'needed', because I'm going to say that I really do feel that it validated a lot of stuff that we were already doing and then really refined and polished it. But, I would say that we absolutely have benefited from all the training"* (School staff).

Barriers and facilitators

Summary of key points in this section:

- IPE research questions addressed in this section are: 3. What are the facilitators and barriers to implementation of the MGDAS programme? and 5. What are the facilitators and barriers to changing policy and practice for school staff following the MGDAS programme?
- Data presented here is from qualitative interview insights.
- Staff highlighted a number of barriers to attending the sessions, such as balancing competing priorities and finding the time to attend.
- Flexibility and accessibility in delivery and having support from senior leadership were discussed as facilitators to taking part.
- Staff spoke about how attendance varied depending on session type. Higher attendance was noted by staff in schools where participation was mandatory.
- Implementers highlighted that delivering the programme was easier when schools showed a high level of commitment and referred to how providing effective oversight was helpful to schools.
- However, they also mentioned that programme constraints and a lack of lead-in time (i.e. the time between initial approach to schools and the programme start date) to recruitment could make it harder for schools.
- Regarding barriers and facilitators to making changes following MGDAS, schools referred to contextual barriers beyond their control that prevented change, but having an SLT that was supportive of change was seen as helpful.

In interviews, school staff described that one of the facilitators of the training was being able to attend alongside their colleagues, as part of **a team effort**. They expressed that this was important to ensure everyone had a shared level of understanding and could work towards the same goals in becoming more trauma-informed in their practice:

> *"So that everybody is on the same… is coming from the same perspective because we're all at different stages, I think, in terms of our learning to deal with these challenging situations. But we all need to have the same vision of where we would like to be, if we were getting it right all the time, what that would look like"* (School staff).

Staff also felt that being in the sessions together helped to create an atmosphere of support and mutual care, as they could discuss how they felt about the sessions afterwards and know that others had shared

the same experience: *"I think it's just good to have someone to talk to straight afterwards and just to say, 'We are actually sharing similar kinds of feelings.' And I think it's kind of talk therapy, to be fair"* (School staff). This also aided knowledge sharing between staff following the sessions.

Staff reported **mixed views on attendance**, with some staff finding it easier to attend sessions than others. Some school staff reported that sessions were *"widely attended"* and that they had personally attended most, if not all, of the sessions that were on offer for MGDAS: *"I think, off the top of my head, I've attended all of the ones for the pastoral team and one of the ones for SLT and all bar, maybe, three of the ones for the whole-staff ones"* (School staff).

Most staff described that they had support from the wider school system, as their **schools had allocated them set time** to attend the MGDAS training. Some schools made the sessions *"compulsory"* or as part of an *"INSET day"* to ensure that as many staff as possible would attend: *"The school really made sure that people really had no excuse not to be there"* (School staff). Staff also mentioned that having the sessions already booked in from the start of the year made it easier for them to plan their time.

> *"I think that's what worked well in terms of our school here was that it was timetabled in. They gave us the time. 'This is a disaggregated INSET day, so you'll do these twilight sessions.' It was planned from the beginning of the year, and then you can plan around it"* (School staff).

Implementers shared that advance scheduling did not necessarily work for every school, as some found it difficult to commit to particular dates for the training sessions or to allow their staff to commit more than one hour to training sessions. The scheduling also worked better for some staff than others, as a small number of schools had staff use time that had already been put aside for lesson *"planning days"*. Staff affected by this found it more inconvenient to attend the training and looked forward to the sessions less: *"I'll be honest, at the start, I was a bit like, 'Oh, this is in the way of our planning Day'. So I wasn't all that excited for it"* (School staff).

Without this directive that the sessions were mandatory, other school staff reported that their school **struggled to get many staff members attending sessions**, with one staff member sharing: *"I don't think people are willing to give their spare time after school. And it wasn't within the CPD [Continuing Professional Development] timetable, and so people didn't want to do it"* (School staff).

Staff spoke about how they found it difficult to make the time to attend sessions or review content due to the pressure of **balancing competing priorities** within the school. Both school staff and implementers described how urgent or safeguarding issues would have to take priority for staff and mean anything pre-planned would have to be postponed: *"It could be in five minutes' time, I might get a phone call to a safeguarding issue. I might have to drop everything and deal with that"* (School staff). This would still be the case even if staff had been granted the time to attend the MGDAS training by their school leadership. Implementers found this aspect in particular *"very challenging"* to deal with. School staff reflected that they utilised support from their colleagues to ensure they still received the content when they did not have the time to attend themselves: *"Sometimes, when I couldn't attend because I had a parent meeting or an incident happened in school, my colleagues attended, and they filled me in. So, we [are] able to use each other"* (School staff).

Implementers echoed school staff members' views on how difficult it was for them to fit the training sessions into their busy daily schedules alongside **other competing priorities**. This meant that it was not always possible for all necessary staff to attend the training sessions (at all or in their entirety): *"Some staff come*

in late, some staff are leaving early, some staff are disappearing in the middle for half an hour to go on duty" (Implementer). Implementers also shared feedback from staff that competing priorities could make it difficult for them to implement learning from the sessions, as they often could not find the time to take a trauma-informed approach with pupils:

> *"They often had class groups of 30 pupils, so being attentive to somebody being a bit off, and what that then meant, how they would sensitively pull that student apart, and creating that time to look into that was difficult"* (Implementer).

Implementers saw school staff workload and capacity issues as being the result of the wider system within which schools are operating: *"That is the biggest barrier, is the absolute stress and strain on the school system"* (Implementer).

Staff spoke about how **flexibility and accessibility in delivery** made it easier for them to take part in the MGDAS programme. Some staff members mentioned that they appreciated the mode of delivery in that sessions were online, which allowed them the flexibility to join from anywhere in their school: *"I think being [remote], again, has its advantages, which allows us to work in a setting across the academy, work in different rooms"* (School staff). Implementers, however, felt that although online delivery offered flexibility and efficiency, more of the training sessions should have been **in person rather than online**, particularly the whole-school training sessions: *"There's a lot of people saying that they find two hours or three hours online quite challenging"* (Implementer). Staff also shared concerns that it was more difficult for themselves and colleagues to be engaged when sessions were online, especially for the larger sessions, where they were unable to have discussions and where it was a *"bit more difficult to observe"* participation.

> *"To do a whole-staff online session and to get the buy-in from staff when it's a busy school, I think, is really, really difficult because they can kind of see it as they just need to sit and listen, and they're not then kind of engaging with the trainer as much as they would if it was face-to-face"* (School staff).

It was felt by implementers that in-person training elements would also help with building relationships and connections between themselves and school staff. School staff members agreed that they would be more engaged with an in-person facilitator, as it would have enabled them to build a level of rapport: *"I am a really like, personable person. I like to see someone and get a feel for it"* (School staff). Staff also suggested that the type of content being covered in MGDAS sessions would have had more impact if delivered face-to-face: *"Some of the things that you're saying, it's not resonating with everyone through a computer screen"* (School staff).

Implementers echoed that some issues with session engagement emerged due to the **struggles with online delivery,** such as logistical and technical issues. These included school staff members' cameras, microphones or chat functions not working; school staff members sharing computers to join sessions; and implementers trying to deliver online sessions to big groups of staff.

"Sometimes they're in a room, on a big TV, there's, like, 10 people, and you're trying to communicate with these 10 people and teach them things, and it's not always easy to hear them or see their facial expressions or for them to interact, necessarily" (Implementer).

In terms of session delivery, staff commented that the **training facilitators aided delivery** by being friendly and *"approachable"*, and they were happy to answer any staff questions or share advice, as well as being understanding of staff feelings:

> *“The facilitators were just so calm and friendly and considerate, and I remember one session, it was just for myself and the pastoral team. And we’d had such a horrible day, and the facilitators were just like, ‘Give yourself a minute; it’s OK’”* (School staff)

Staff described that this helped them feel validated and like *“you’re not alone”*. They also felt that having the same trainer for sessions was useful for building rapport.

School staff members emphasised the **importance of the oversight team** at KCA in being *“very accommodating”* by helping them arrange (or rearrange) sessions at short notice. They felt that the programme manager in particular had a good understanding of the school and was always available to call if they had queries: *“I’d say [programme manager] at KCA was just phenomenal in her understanding of the school and organisation and very politely chasing up”* (School staff). Session deliverers would also change sessions depending on the school requirements:*“KCA have been extremely flexible in terms of when we want to deliver things and for how long, and adapted the sessions so that they fitted with our calendar rather than theirs”* (School staff). Implementers agreed that they tried to be flexible where possible in response to schools’ needs around the timing and length of the training sessions: *“We’ve been incredibly flexible: early morning sessions, twilight sessions. That’s key because the school day is so busy for people in school”* (Implementer). This approach helped staff feel that facilitators *“understood the pressures”* of working in a school. Implementers indicated that ensuring that the **training content was appealing** and delivered in a stimulating manner had facilitated programme implementation through encouraging school staff member engagement: *“I think, if it was just reading off PowerPoint slides, it would be really difficult. I don’t think you’d keep them engaged, whereas the knowledge of the facilitator is really key”* (Implementer). From a training facilitator's perspective, having well-organised materials was important to facilitate delivery. They also felt that the supervision they had received from the programme delivery team was valuable in helping them fine-tune their delivery skills: *“It helps us think through any resistance or any lack of engagement and how we pick that up next time, so it’s a journey of continuous improvement”* (Implementer).

School staff members agreed that sessions were *“really professionally delivered”* by facilitators, who knew the content well and were *“passionate”* about the subject of trauma-informed practice: *“All the people that have presented and come have been really intelligent, really well informed, really knowledgeable, but also enthusiastic about the work”* (School staff). Staff members indicated that this helped them understand the content better, as it made the theory *“more accessible”* and gave them a sense of what they could do to support their pupils. However, one senior staff member shared that some staff in their school had feedback that the delivery was *“didactic”* and not as effective in parts, so staff members felt like they were in a lecture rather than an interactive session. However, this appears to be a minority view, as staff in other schools remarked that the delivery did not feel like it was *“reading from a script kind of training”*.

Implementers referenced the importance of having **effective oversight of the programme delivery** to aid implementation. This included working closely with schools to schedule training sessions, ensuring that school staff were able to attend and ensuring that school staff had the technology to participate. Implementers also mentioned the utility of effective database and project management systems in facilitating programme implementation.

Implementers also mentioned that the nationally recognised expertise of and positive relationships between the partner organisations involved in the MGDAS programme also facilitated programme implementation.

Implementers spoke about having **difficulties engaging schools** to be a part of the programme. This was again linked to lack of school staff time and capacity in general but also to the demands and time associated with the programme, specifically the length and number of training sessions that school staff needed to attend as part of the programme: *"I think that's been one of the barriers to schools feeling like they could possibly engage with this because it is quite a big commitment"* (Implementer).

Implementers also noted the **lack of *"lead-in"* time** that they had had for the programme, which they felt was a barrier, as it negatively affected school engagement with the recruitment process. For example, it had given implementers limited relationship-building time with schools and senior leaders to facilitate recruitment prior to the start of the programme. The programme had also been introduced partway through the school year rather than at the start of it due to funding and timeframe restrictions that meant the project was unable to approach schools before they had planned their CPD schedules for the year, which implementers felt was not helpful for schools in terms of planning the roll-out of the programme: *"I think the time that funding came through and the time that the programme was running was probably a little bit too late for schools to actually schedule it in"* (Implementer).

Implementers mentioned **difficulties with school recruitment** for the programme. They described having high recruitment targets for schools participating in the study, which led to them having to expand school recruitment nationally beyond specific target areas. Implementers also explained that they had initially hoped to rely on known contacts in LAs to encourage schools to take part in MGDAS. However, they found that the degree of LA influence over schools' involvement in the programme varied, with influence more difficult in relation to MATs, which made it harder to implement MGDAS because schools had more autonomy over the choice to participate (or not) than initially anticipated:

> *"It relatively quickly became apparent that those local authority contacts didn't have as much influence in schools as we thought. Because of that, we realised that there were going to be some challenges in recruiting the numbers that we needed to be able to run the programme. So, we then had to make a decision to roll out the programme to a wider audience."* (Implementer)

Another barrier implementers referenced was the **MGDAS programme constraints** and how these affected their implementation of the training sessions. These included having a set number of hours of training per school that school staff had to participate in, trying to engage schools without having an existing relationship with school SLTs and, as previously mentioned, having limited time to build relationships with schools at the outset of the programme. Ultimately, programme constraints meant that there was limited scope for implementers to adapt the programme around individual schools' needs in the way in which they usually would outside of MGDAS.

> *"What has been really difficult with this one is that we've gone in with a package, which we'd never do in the real world, and we've gone, 'Now you need to make this work. You need to give us 16 hours', I think it is, of training dates and time. That has been really difficult."* (Implementer)

Implementers suggested that having more time for co-creation of the programme with schools and stakeholders would be valuable: *"Class teachers were perhaps not as involved as we would have wanted to. So, I think that might be something to consider"* (Implementer).

Non-teaching staff referenced how feeling like the **programme content was not** aimed at them acted as a barrier to their engagement. A few staff members felt that MGDAS sessions were more tailored towards teaching or SLT and PI colleagues and that non-teaching school staff members were under the impression

that the training would not be beneficial to them due to their job roles, with one senior staff member saying about their colleagues: *"I think sometimes they think, 'Oh, it's just for teachers. It's not really very relevant to us; we're in the office'"* (School staff). One non-teaching member of staff, however, mentioned that they appreciated the opportunity that MGDAS gave for them to take part in training, as regular CPD *"wouldn't involve support staff".* Senior leadership staff felt that support staff could benefit from *"more sessions"* that were specific for them, with basic principles of trauma-informed practice that they could use, for example, when dealing with parents: *"I think it's really, really important that they perhaps have their own session, you know, about meeting and greeting and perhaps build it in some other training about de-escalation techniques for when parents come in, and they're irate"* (School staff).

Staff suggested that the **use of *"constant reminders"* and repetition** of the content was needed following MGDAS to facilitate consolidation of their learning and to help them embed the content within the culture of the school, as it could not be assumed that all staff members would change their practices automatically:

> *"It's that presumption that everyone else will follow, and it's not as simple as that. You have to really take the time to reinforce, and over-embed, and go back, and quality assure. And so, we've got a lot of work to do there"* (School staff)

Staff also found that the materials and resources provided were a useful reminder, as they could take this away to refer to and use after the sessions: *"I like it when they send handouts and things that you can refer back to because, you know, we're in a busy school day, your kind of memory's sometimes not that great"* (School staff).

However, when staff felt the content was too repetitive in terms of *"going over the same information"* again, this impacted learning, as they did not feel they were gaining any new knowledge. This suggests that a balance needs to be struck between reminding staff about the content and ensuring this does not come at the expense of new learning. Some staff described how they wanted **more *"practical tools"* or ideas** from the sessions, rather than the perceived repetition, and would have preferred to have been practising putting the content to use:

> *"I think that sometimes it felt like we were covering the same ground over and over again. I don't personally think that that's necessarily a bad thing because I do think we need to be reminded of it, but, like I've just said, because you learn in the abstract, but you need to practice it in reality, in the moment that it happens"* (School staff)

Implementers commented on **schools' positive attitudes and commitment** towards the programme as a facilitator to implementation and engagement with MGDAS: *"I think there's a lot of staff who have a huge amount of goodwill and desire to support students, and therefore will enjoy doing this kind of training"* (Implementer). Commitment of the school SLT was seen as particularly important to enable whole-school change to take place and to give the programme credibility within schools: *"We know that that needs to happen in schools if you're going to bring about real systemic change because, ultimately, the leadership team [is] responsible for them drafting potential changes to policies and procedures"* (Implementer). Implementers described typically beginning their engagement with schools by having discussions with the SLT and ensuring that SLT members attended at least the first training session to facilitate this positive attitude.

Most school staff members echoed these thoughts of the implementers regarding attitudes and commitment, saying that they could **see the value and importance of the training** and mentioned that they

left sessions feeling encouraged to take forward what they had learnt and put it into practice. They believed this would have a positive impact on their pupils: *"It kind of put that positive thought in my head that, 'No, it's not all about teaching. There's a different side to it. It's about making a positive difference. It's about being there for the children'"* (School staff). Recognising this, therefore, facilitated engagement.

Conversely, if staff members did not see the value in the MGDAS training, then this acted as a barrier to engagement, as they were less likely to use the learning or attend the sessions. School staff members referred to how they were aware of colleagues who were more reluctant to change their ways of working: *"You will sometimes get those that have got a little bit more resistance, you know, to moving forward"* (School staff). They also mentioned that they saw staff who still needed to fully embed the lessons from MGDAS into their practice, particularly when dealing with pupils and trying to get to know them: *"I think some need to improve on how they deal with relationships"* (School staff). Some staff perceived that they were already using trauma-informed practice and therefore did not need more training: *"Sometimes, the discussion was, 'Why are we doing this training? We certainly don't need the training'"* (School staff). Staff shared how their colleagues sometimes held the view that being trauma-informed was not their responsibility, as they preferred to focus only on the teaching side of their job, but this was something interviewed staff disagreed with, stating:

> *"What I see in my training is you're not there to just educate. You're there to help the children as well. And it's not the right job for you if you think you're just going to teach the subject that you're teaching"* (School staff).

Some staff described how the impact of one's own trauma could lead to a reluctance to attend MGDAS sessions due to concerns that the content would trigger memories of their own experiences: *"There's a lot of people that probably didn't want to be involved because it's close to home"* (School staff). Staff did not feel that *"enough consideration"* was taken by their own school if staff did not want to go through the training if affected in this way.

These views regarding some staff members' reluctance to attend or engage with the programme demonstrate the complexity of delivering training to schools and staff with varied needs and perceptions about trauma-informed practice.

Staff described various **contextual challenges** that made **implementing any changes following MGDAS more difficult**. One of these was that individual schools were **restricted in making changes** if they were part of a wider academy trust, as policy decisions came from a level above the school, and, therefore, school staff had limited influence. Some schools saw that they still had the ability to *"make amendments here and there"* despite this, but for most schools in an academy trust, changes could not be made to a huge extent: *"In regards to our policies, they come from [academy trust], which are super thorough, and we've got loads of policies for loads of things as well. So I don't think we can truly impact those policies too much"* (School staff). In the case of these schools, non-leadership staff expressed uncertainty and saw that any changes that had happened were coincidental: *"I think there's been changes in policies, but I'm not sure that they've been directly kind of linked to the More Good Days project"* (School staff). Staff mentioned that where their school could not alter policies, work was being done in areas in which they could impact change, such as wellbeing support for staff.

School staff members were sometimes unsure whether changes to policies had been made yet or whether their senior leadership had plans to do this, particularly if they were not in a role that gave them *"an oversight of the whole school"*.

Staff also referred to contextual barriers that were beyond the scope of their school or academy trust leadership, in that sometimes schools' efforts towards more trauma-informed practice differed from directives from the *"government level"*; however, they did not elaborate on which government initiatives they were referring to. School staff felt that approaches needed to change at this higher level to facilitate any major shift.

> *"Sometimes, you know, what we're trying to do and what [the Department for] Education tell us to do are just conflicting. So, I think that changes need to happen higher up, and there needs to be a realisation higher up that there needs to be a different way of doing education"* (School staff).

School staff explained how **supportive senior management and leadership** were effective in helping them both attend training and enact any changes they wished to make. Senior staff members were aware that they needed to give a *"consistent message"* about embedding trauma-informed practice and give reminders to their colleagues. Teaching and pastoral staff described the importance of having an SLT that was strongly committed to the programme and would put measures in place to ensure they could attend sessions. School staff also spoke about their senior leadership driving changes and making it *"quite clear that the school is working to become more trauma-informed"*. Staff felt that this drive filtered down through the school and helped create a culture of change: *"You're going to follow what the principal wants, aren't you? She's the one that leads the school. So, having her being involved in the sessions and promoting the session, it's definitely going to have a positive impact"* (School staff).

Staff also described that their senior leadership were *"approachable"* and receptive to suggestions and willing to listen to staff if they felt something was missing from their school practices or policies following the training: *"We're more aware now, and we do regularly have the training, so we can actually make suggestions, and they are taken onboard and they are listened to"* (School staff).

In the small number of schools where there was a **perceived lack of support** from senior leadership, staff deemed that this impacted their ability to make changes following the MGDAS programme. Some staff felt that although senior leadership did attend their own training sessions, they were not as supportive of school pastoral staff and were not present for discussions regarding potential changes. One staff member shared the view of their pastoral colleagues: *"They said to me, like, 'Why are we having these conversations about how to improve things if the SLT aren't there?'"* (School staff). Having the positive support from school senior leadership was therefore rated as an important facilitator to change, whereas perceived unsupportive leadership could negatively impact the ability to change.

Changes

Summary of key points in this section:

- IPE research questions addressed in this section are: 4. What changes to policy and practice do school staff implement following the MGDAS programme?, 8. What is the experience of pupils of support from schools following the MGDAS programme, and do different groups of pupils (i.e. ethnic groups, eligibility for FSMs, SEND/EHCP) have different experiences of support? and 9. To what extent, and how, does the MGDAS programme impact school culture?
- Data presented here is mainly from qualitative insights, with some quantitative data from the SDSs.
- Staff described feeling better equipped to help pupils and reported improved connections between staff and pupils.

- Overlapping themes were reported by pupils in interviews, although it was not always as clear that these were a result of MGDAS. Pupils described experiencing supportive relationships with staff, but they also described barriers to supportive relationships.
- Both staff and pupils described an increase in wellbeing-focused events and activities, with pupils going on to talk about signposting and referrals to onward support and wellbeing spaces.
- Staff also talked about more general changes to ways of working (e.g. use of reflective practice) and that the training had helped them to affirm that their school practices and culture were already or nearly trauma-informed. Some staff, however, described a lack of change in culture, policies and practices following the training.

School staff described making **changes to their schools' behaviour policy language** to reflect ideas that they had learned about in the training sessions: *"Instead of discussing behaviour problems, we discuss it as relationship issues because it's about belonging, as well"* (School staff). This included renaming behaviour policies to reflect the importance of the relational context of behaviour, renaming sanctions to emphasise the role of *"reflection"* or *"conversations"* in resolving behavioural issues, and adding concepts to existing behaviour policies, such as *"self-regulation"*. Staff described how their school had changed the way in which it referred to more "traditional" sanctions, such as isolation rooms, so that students would view them in a less punitive manner.

> *"The function of the room didn't really change, but even just re-labelling it, I think, changed our feeling, and hopefully the kids' feeling, that this wasn't a space where you were punished, and you were isolated from everybody because nobody wanted you in their room"* (School staff)

Staff indicated that following the training sessions, they were now looking at behaviour management through a more relational and trauma-informed lens: *"It's about, again, the relationship, so this is what's happened, this is the bit that we can't have happening, so what do you need to do differently? What do we need to do differently?"* (School staff). This included incorporating more opportunities for reflection, conversation and time-outs prior to sanctions.

Relevant insights from quantitative data

This shift described by staff is reflected to an extent in the quantitative findings. According to descriptive analysis, data from the SDS showed a decline in the use of sanctioning language in school policies across both intervention and control schools between baseline and endline. At baseline, the two groups were similar, with around 51.6% of intervention schools and 56.7% of control schools reporting that such language was used "often" or "all of the time". Where there was data at both time points, the baseline was slightly higher at 66.6% and 62.6%, respectively. By endline, the figures had decreased to 50% of intervention schools and 46.6% of control schools, possibly indicating that the reduction reflects a wider shift across schools rather than a programme-specific effect.

In interviews, staff described the skills and knowledge that they had gained from the training sessions, which had enabled them to **feel better equipped to help** pupils.

> *"What would you say that you've learned from the sessions, you know, that you've been able to really apply to your role, that you would say? I think it's you never truly understand what the child's experienced and what that impact that's had on them. So, actually, how am I being with that child? Am I able to, I'm much, much better myself at seeing a situation and actually de-escalating that and*

> *actually different ways to de-escalate and completely de-personalising the behaviour response to me"* (School staff).

Specifically, staff mentioned drawing on concepts that they had learned in the training sessions to manage pupil behaviour issues, such as *"co-regulation"*, having time to calm down before having a conversation about an incident, having a *"restorative conversation"*, *"mirroring"* and identifying a *"trauma response"*. They felt that these were concepts that the wider staff group now had shared knowledge of, which was useful. They also referenced having more of an understanding of the issues that can be behind pupils' behavioural difficulties.

Staff described having a more "*planned*" or structured approach to having difficult conversations with pupils and with parents and carers: *"I go in with more of a planned approach whereby before, naturally, I would have taken a kid away and just calmed the situation down, but now I think about the steps a little bit more"* (School staff). They also indicated that they were now having more regular discussions or using new shared language and terminology in their conversations with other staff members about how to support pupils: *"I wouldn't say it has shifted policy, but I would say it has definitely shifted what we're talking about and the way that we talk about handling, particularly, high-profile kids"* (School staff).

Staff alluded to the importance of **being able to build positive relationships with pupils**. They felt that connections and relationships between at least some staff members and pupils had improved following the training sessions.

> *"I can only think that a lot of the More Good Days At School work that we're doing, it's helped maybe staff handle students in a better way, or at least maybe students feel more connected to their staff as a result"* (School staff).

This included having less conflict, more friendliness and more understanding of or empathy for pupils.

> *"I think it's just finding why, the reason why, having a kind of empathetic understanding, and yeah, just seeing the students just as an individual, just as a human being, rather than focusing too much on behavioural issues or anything like that"* (School staff).

Staff mentioned that since the training sessions, they had been **delivering more mental health and wellbeing-focused sessions for pupils**, including psychoeducation, mentoring and fun activities. They referenced streamlining, honing or expanding their pupil wellbeing support offer following the training sessions: *"We will align that training to our position as the reasons why we're making the decisions and the reasons why we're shaping what we do because of that there. So, that with the staff, it actually connects the dots"* (School staff).

Staff also mentioned that **activities and support for staff wellbeing** had also become **more prominent** since the training sessions: *"I think that staff are taking it more seriously now and knowing that everyone's mental wellbeing is important"* (School staff). This included understanding that staff may also have experienced trauma: *"In order for staff to successfully support pupils, there needs to be that support in place for staff as well. So I found that really helpful, that big eye-opener in regards to that"* (School staff).

Relevant insights from quantitative data

In terms of insights about mental health provision and support data from the SDS, intervention schools were more likely than controls, according to descriptive analysis, to expand both universal and targeted

approaches, indicating that the programme helped build additional capacity to respond to pupils' needs. Similarly, intervention schools showed greater increases in approaches to engaging parents and caregivers, a core element of trauma-informed practice that supports a relational, whole-school culture. These gains suggest the programme strengthened schools' ability to involve families in promoting pupil wellbeing, above sector-wide changes.

A different trend emerged for staff wellbeing. Although intervention schools started from a stronger baseline with more structured support, control schools reported larger increases over time. This implies that sector-wide developments may have driven improvements in staff wellbeing more than the programme itself, with intervention schools having less scope to expand further.

Pupils described in interviews about being **signposted or referred to support** by school staff: *"If you tell like a normal teacher who's like not pastoral or anything, then they'll pass it on to pastoral"* (Pupil). They mentioned having classes or group sessions about mental health, wellbeing or self-esteem and having one-to-one support at school to address specific mental health difficulties: *"It was very helpful for me because I felt comfortable talking with them, and they also helped me a lot. They gave me ideas to do when I'm stressed"* (Pupil). Pupils also referenced their parents and carers being informed by school staff about their support needs. However, pupils were uncertain if these were recent changes (i.e. following MGDAS) or approaches that had always been in place.

Pupils mentioned having **designated wellbeing spaces** at their schools within which they could access wellbeing support from school staff: *"You feel really safe because you've got all the people that you need up there all the time"* (Pupil). They also described using these spaces when they needed some time away from lessons or the rest of the school: *"You can just chill there for a while to make you feel at peace and calm"* (Pupil). Pupils also used time-out passes or requested *"self-regulation"* time when they needed some space or support: *"You'd get a timeout card, which you get 10 minutes out of a lesson to sort yourself out or go talk to someone"* (Pupil). However, it is not clear whether these spaces were already in place before MGDAS, or if they had been implemented or improved since then. The impact of MGDAS on increasing wellbeing spaces for pupils should therefore be interpreted with caution.

School staff described ways in which the training sessions had influenced **changes to their ways of working**. This included increased use of reflective practice in their work: *"Being part of the senior leadership team, I would say the biggest impact has probably come from some of the meetings that we've had, where we've been looking at reflective practice"* (School staff). Staff described reflective practice in terms of all staff members having a voice: having conversations with each other about difficult, upsetting or safeguarding situations; receiving feedback or an alternative perspective; and having space to *"offload"* and be listened to. They noted that staff meetings had become more useful and efficient as a result of reflective practice: *"Anything you can do to reduce an hour's-long conversation where you don't really get any actions out of it, down to a 20–25-minute impactful conversation, you're going to make people happier and less stressed just by doing that"* (School staff).

Staff also referenced their understanding of the importance of recognising their own emotional states and ability to regulate their emotions while at work. This included looking after your own wellbeing as a staff member and managing burnout: *"I think it's almost like you've got to look after yourself first before you try and help someone else, really"* (School staff).

Finally, staff indicated that they were taking a *"gentler"* approach to behaviour management following the training sessions: *"Maybe, explore other things before you go heavy-handed with the sanctions because*

everyone deserves a chance to be heard and everyone deserves that bit of nurturing. So, we've softened that approach" (School staff). This included being less quick to jump to the use of sanctions, helping pupils change their behaviour, exploring what might be underpinning pupils' behaviour and showing pupils that staff are there to help and listen to them.

Relevant insights from quantitative data

Descriptive analysis of the data from the SDS (restricted to the subsample of schools with both baseline and endline responses) suggests that intervention schools reduced their reliance on traditional sanctions, such as warnings, detentions and exclusions, while consolidating approaches like reflecting, addressing needs and using restorative practices. The proportion of schools responding "often" or "all of the time" to the use of warnings decreased from 91% to 73% in intervention schools (from 100% to 91% in control schools). The use of detentions fell from 91% to 64% in intervention schools (remained at 73% in control schools), and exclusions declined from 36% to 27% in intervention schools (increased from 18% to 27% in control schools). High-frequency use of reflections increased from 55% to 82% in intervention schools (with similar increases of 55% to 82% in control schools), addressing needs increased from 55% to 91% in intervention schools (remained at 73% in control schools) and the use of restorative practices increased from 64% to 91% in intervention schools (reduced from 100% to 64% in control schools). Intervention schools also showed stronger gains in fostering pupils' senses of safety and in reframing behaviours as expressions of trauma or unmet needs, rather than simply reactions to immediate triggers. High-frequency consideration of pupils' senses of safety increased from 58% to 92% in intervention schools (63% to 80% in control schools), and emphasis on immediate triggers decreased from 42% to 25% in intervention schools (increased from 44% to 47% in control schools).

At the same time, both intervention and control schools strengthened relational and inclusive practices, such as showing empathy, fostering belonging, differentiating support and providing needs-based support. However, the patterns differed: intervention schools emphasised empathy (66.7% to 100% in intervention; 56.3% to 80% in control) and differentiation (75% to 100% in intervention; 75.1% to 86.6% in control), while control schools reported larger increases in fostering belonging (91.7% to 83.3% in intervention; 87.5% to 100% in control). These broader shifts likely reflect sector-wide developments in inclusive education, but the distinctive changes in intervention schools, such as less reliance on sanctions, greater use of restorative and reflective practices, and reframed approaches to behaviour, point to the programme's specific influence in embedding trauma-informed strategies-based support. However, the patterns differed: intervention schools emphasised empathy-wide developments in inclusive education, but the distinctive changes in intervention schools, such as less reliance on sanctions, greater use of restorative and reflective practices, and reframed approaches to behaviour, point to the programme's specific influence in embedding trauma-informed strategies.

In interviews, and as referenced in the barriers and facilitators section, **not all staff members had noticed any changes** to their school's culture, policies and practices following the training sessions. They felt that more time or effort was needed for change to be realised; that their school culture was already positive, relationally-focused or trauma-informed; or that staff were already doing the practices advocated by the training sessions.

> *"I think what we need to do is make sure that we try and keep a slot on our all-day insets to incorporate and build into the culture of the school some of these things that we've learnt this year.*

So, it's not a criticism of More Good Days At School, it's more of how we implement it, I think" (School staff).

Moreover, it was not always possible for staff to tease apart the impact of the training sessions from the impact of other similar initiatives that they had going on in their schools.

> *"Because we're always so busy and there's always something else and another training, and there's always something we're doing; I think every week we have to stay behind for some sort of training. So, I can't think of anything to hand"* (School staff).

Although some staff wellbeing changes had been noted, staff commented on further changes that they would like to see, including more priority placed on staff wellbeing, such as regular supervision meetings for staff to discuss challenging situations.

Some perceived changes in support from their schools were noted by pupils. However, we cannot say definitively from the pupils' perspectives that these changes were directly the result of the MGDAS staff training. This is because it was not always clear from pupils' perspectives whether changes that they had noticed had happened since the outset of MGDAS or whether these changes had been instigated prior to the programme. Moreover, pupils were not aware that the MGDAS staff training had been taking place, meaning that it was difficult for them to identify this as a causal influence with certainty.

The following changes were noted by pupils:

- More frequent check-ins from staff about pupils' wellbeing: *"They go in and say, 'Is everybody OK, do you need anything?' or anything like that. So that's really a change since Year 7"* (Pupil).
- New staff members with mental health–focused roles: *"We had a different teacher before, but [they] had to move away. So we had another one come in, and [they've] been doing a great job so far. We've been able to do some things for Children's Mental Health Week"* (Pupil).
- New wellbeing support initiatives: *"I think they've brought in loads of new stuff all the time about like wellbeing and things"* (Pupil).
- A shift away from detention as an immediate punishment to reflection time instead: *"It's similar to detention; you stay behind after school, but instead of, like, just dead silence, either you talk to the teacher, and you chat about what happened and why, or you reflect just by yourself or stuff like that"* (Pupil).
- Perceived improvements in school safeguarding practices: *"Safeguarding has become a lot better"* (Pupil).
- More focus in general on speaking about mental health and wellbeing at school, such as through assemblies: *"I feel like Year 7 and Year 8s, it wasn't really spoken about; I feel like it's more now"* (Pupil).
- Improvements in staff members' ways of interacting with pupils, such as staff members trying not to shout at pupils or being more friendly: *"Some teachers you can see improvement with the ways of teaching and ways of speaking, not shouting to students to stop and in a good tone and try not to shout at them so they won't feel uncomfortable"* (Pupil).

- Staff helping pupils understand why they have been given a particular punishment and trying to understand pupils' perspectives and situations when a behaviour issue occurs: *"If you've got like things going on at home or something and you lash out in lessons, like, they'll help you, and they'll understand"* (Pupil).

To explore this further, we triangulated pupils' perceptions of change with those of staff members at their schools. Some staff members referenced changes that had taken place or were in the process of taking place at their schools following the MGDAS staff training. These included changes in behaviour policies (in terms of taking a more relational, reflective or restorative approach, rather than a punitive approach) and an increased focus on pupils' mental health and wellbeing: *"Strategically we've changed a lot. So, again, we don't go straight to punitive; we have reflection meetings first, so students have a chance to come and talk about if they're removed from a lesson"* (School staff). Staff also noted changes in staff members' ways of interacting with pupils, such as staff having more *"sympathy"* and *"compassion"* towards pupils, and pupils being encouraged to be more *"open"* with staff. These reflect changes identified by pupils.

However, sometimes staff also found it difficult to link these changes directly to the MGDAS staff training, as they felt that these changes had already been in play prior to the training: *"I couldn't say for sure that anything that's come through on the trauma-informed training has gone into practice because we sort of do it already"* (School staff)). They also felt that more time was needed for changes to happen and become embedded, commenting that it is a *"long process"* or a *"journey"*. For example, staff felt that some staff members still had room to improve on how they related to pupils: *"I think some teachers are old school, and like, 'We're here to learn, and it's my way or the highway'. But, yeah, I think some need to improve on how they deal with relationships"* (School staff). Staff were also sceptical about whether meaningful change would actually happen to *"rigid"* behaviour policies: *"So, they follow those steps, and sometimes it doesn't allow for that relational practice"* (School staff).

Staff felt that the MGDAS staff training would have had knock-on effects for pupils but that pupils might not recognise these because they would be more *"behind the scenes"* or because the school had already undergone a lot of change anyway.

> *"It might be that they've discussed doing an internal isolation for a student. And actually, when everyone's got their heads together and thought in a more relational way and a more trauma-informed way, that actually that's not happened for that student and the sanction's been different or the support's been different"* (School staff).

Relevant insights from quantitative data

When asked to describe their school's culture at baseline in the SDS, intervention schools reported via open responses that they had varied approaches. One school stated that it had issues with *"increased poor behaviour"* but that this was *"misunderstood by some staff",* and pupils' mental wellbeing was often *"overlooked as the root cause".* Another school stated that although all staff had previously had trauma-informed training and this was supported by SLT, *"naturally, some people's natural view does not align directly with the school"*. Similarly, a different school described how *"not all staff are yet convinced about [a trauma-informed] approach and want to see more visible punitive action taken to 'solve' behaviour"*. Schools did, however, generally describe that they had good mental health knowledge *("There is a good awareness of mental health in school for both students and staff"*) and a willingness to be supportive of mental health (*"Staff are very proactive in identifying pupils who require additional support"*).

At endline, intervention schools most commonly described having an open approach to mental health understanding and awareness. One school described that it had *"committed to having a full mental health action plan in place*" for the next few years. Schools also felt that their culture promoted values of kindness and care (*"We have a supportive and caring ethos, where each individual and contribution is valued"*). However, some schools felt that the impact of MGDAS on their school culture was still developing; one school stated that the lessons from the training "*[needed] to be more embedded"*, which is supported by the qualitative interview findings, as described previously in this section.

Most schools did not list specific changes that had been made as a result of MGDAS. Those that did, highlighted how changes had been made in approaches to behaviour management and emotion regulation; for example, one school mentioned incorporating *"restorative conversations to support emotional safety and positive behaviour in the classroom"*. Schools also stated changes in the use of reflective practice (e.g. *"fortnightly reflective practice sessions with SLT to be introduced to middle leaders"*). Some schools described how they were making changes to staff wellbeing, such as by introducing *"a staff wellbeing charter, created in collaboration with staff wellbeing leads and unions"*.

An additional change to culture detailed in interviews was that school staff felt that the training sessions had *"affirmed"* the good practice that their schools were already doing and confirmed or provided reassurance that they were already taking the right approach: *"What it has done is it's affirmed everything that we believe in, do you know what I mean? All of this has said, we're doing things the right way"* (School staff).

They also felt that the training sessions had enhanced their schools' existing practices by providing labels or a common language for their approaches and actions (e.g. trauma-informed), an evidence-based or scientific rationale, and a structure or framework for their current activities, including bringing together existing practices.

> *"It was nice to have almost like, 'Here's the evidence for why we do it. Here's the justification. Here's the language around it', which I think is sometimes the thing that people struggle with is that they know what they're doing is good practice, but they wouldn't necessarily be brave enough to call it 'trauma-informed'"* (School staff).

Staff members indicated that while their school culture had potentially already been positive or trauma-informed, or shifting in that direction prior to MGDAS, the training sessions had helped *"fine-tune"* and advance their school's journey or cultural shift.

> *"I think that the school has a very warm culture, but in terms of how things are, yeah, there's that continuing improvement that we see happening. And again, it's calmer. Again, students will look to resolve that conflict and will be open to that restorative when something serious happens"* (School staff).

These findings from both the qualitative interviews and quantitative surveys suggest that staff members were feeling hopeful about how MGDAS could impact their school culture, but they still felt that there was some way to go before full-scale or widespread changes could be seen.

Long-term impact and sustainability

Summary of key points in this section:

- IPE research questions addressed in this section are: 5. What are the facilitators and barriers to changing policy and practice for school staff following the MGDAS programme (in terms of sustainable change)?
- School staff talked about what would influence the future sustainability and impact of MGDAS.
- They discussed that long-term staff buy-in was needed, along with refresher trainings and reminders, to ensure that the programme learnings became embedded.
- Staff expressed concerns that changes to the staff body in school may affect culture and, therefore, affect programme sustainability.
- Staff also talked about the importance of parental and social support in ensuring they could implement a trauma-informed approach following MGDAS.

Staff members reflected on **the importance of *"staff buy-in"*** when considering the long-term legacy of the programme. Some interviewees reported that staff would be willing to continue making changes if they could see that they were having the desired impact:

> *"It speaks for itself, with the data and the school now than what it was. It speaks for itself; it works. The teachers say, 'I can teach my lesson. I'm not dealing with behaviour; I can teach my lesson'. And that's a very impactful statement, isn't it? And I think that's what gets us the buy-in"* (School staff).

Staff members felt that seeing these results would help them ensure that the approach changes would remain *"embedded in our culture".* Staff members stressed the importance of senior leadership support as a factor and emphasised the need for someone *"at the top"* to be pushing to *"elicit any change"*, setting the agenda and scheduling any further training to ensure its long-term success:

> *"I think staff need motivation, and call it what you want, a motivational speaker or a kick up the backside, I don't know what it is, but there's got to be something from somewhere or someone who's kind of leading it. And you look up to your principal, and your deputies, and SLT, and if it's for the good of the children, then absolutely. But I believe it's got to come from the top"* (School staff).

Staff members shared concerns that their *"SLT having a brand-new initiative that they want to go on"* could impact the sustainability of the MGDAS programme, as this would then take priority for school leadership, rather than focusing on implementing the lessons learnt from the current training: *"There's always something we're learning all the time. If there's no continuation over a long term, then it's difficult for that one thing to have an effect"* (School staff). Staff felt the number of initiatives being driven by SLT could be *"overwhelming"* at times, meaning initiatives like MGDAS are potentially forgotten about as they cease to be at the forefront of the minds of staff: *"If we're developing culture and we want this to be an active part of the school, it has to be the thing that we are always talking about. That's the thing that kills off initiatives, I think"* (School staff).

Staff members also spoke about **the importance of refresher trainings and feedback** when considering the sustainability of the programme. Staff members stressed that one-off trainings are often forgotten, so they recommended that staff learning be reinforced with regularly scheduled refresher trainings: *"If it's front of mind, then it's more relevant. You're going to remember it"* (School staff).

While staff members were clear that further training was necessary, their views varied on how often it was needed; some staff thought *"yearly"* training on the topic would be sufficient, whereas others thought reminder sessions should take place *"every six months"*. Some staff members did not specify frequency but thought that refreshers should take place *"little and often, just little reminders here"* so that trauma-informed approaches could become *"a natural part of your day, as opposed to something you come up against three times a year for an hour"* (School staff). Staff members suggested short, interactive *"practical skill"* trainings – such as scheduled reflective practice and feedback pop-ins – would help to consolidate learning: *"Maybe have some little clinics or just some little training slots and make it interactive, make it fun, make it engaging, make it something that, you know, learning but fun"* (School staff). Staff members suggested reminders could also be in the form of weekly focus points, presentations or posters displayed in classrooms.

Staff shared how **changes to the staff body could affect school culture** and could impact the sustainability of the programme. Staff described that a *"high turnover of staff"* and new starters could mean that gaps in training form in the staff body, which can affect whether training is embedded into the school culture long-term:

> *"I feel like the change to staff body, like a massive change where everyone that was given this training, or many people that were given this training, are then no longer in the school; I think that could affect school culture again, where it all starts to change again and there's a shift, and we move away from this"* (School staff).

Staff also felt that any changes to leadership could also impact long-term change, as they may not understand the value of any practice changes that have already happened: *"It's the risk of a change of leadership which might not see what you've done or understand the developments that have taken place"* (School staff).

Staff members advised that a **lack of parental and social support** could impact the legacy of the programme. They felt that parents often did not understand the change towards being more trauma-informed and could sometimes disagree with staff when they are trying to implement change: *"I think our parents are a massive barrier to some of those changes and to the implementation of things because they don't agree, or they hear the child's version, and it's not the true version"* (School staff). Staff felt that sharing more information with parents would help mitigate this.

Some staff also felt that wider societal changes would be needed to support staff members in enacting change, which was beyond the scope of the MGDAS programme in its current form:

> *"I think a school like ours is certainly going to keep doing what we're doing and keep making those impacts, but, until you feed out into the community, I think there's only so much success or change that we're going to see"* (School staff).

Relevant insights from quantitative data

Despite these concerns listed above, schools still reported in the SDSs that they were implementing more approaches to engage parents and caregivers by endline. At baseline, and according to descriptive analysis, both groups were already engaging parents and caregivers, though with slightly different emphases (e.g. control schools more often provided written information, while intervention schools more often offered one-to-one support). By endline, intervention schools were substantially more likely to expand their

approaches than controls. Among the subsample that had baseline and endline data, 73% of intervention schools increased the number of approaches they used to engage parents and caregivers (average increase: 1.18 additional approaches), compared with 45% of control schools (average increase: 0.64 approaches). No intervention schools reduced their approaches, whereas 18% of control schools decreased the number of approaches they used. This suggests that the programme contributed to strengthening schools' capacity to involve parents and caregivers above sector-wide changes. As engaging parents and carers is a core element of a whole-school supportive and relational culture, this aligns with the programme's intended outcomes.

Cost information

Summary results

The table below provides a summary of the intervention's estimated costs. Total costs are estimated at £2,172,731 (2023–2025 prices) for delivery across 31 schools, involving approximately 3,720 staff participants. This is equivalent to £70,090 per school or £584 per staff participant. Of the per-school cost, £52,327 would be incurred by the school, while £17,764 would be incurred by the delivery team. The majority of costs were driven by staff participation in training, with setup costs forming only a small proportion of overall expenditure.

Table 22: Summary cost results.

	Total costs	Cost per staff participant
Setup	£38,788	£10
Recurring	£2,133,943	£574
Total	**£2,172,731**	**£584**

Prerequisites

The costing exercise assumed that certain resources were already available within schools and therefore did not generate additional costs. Schools were expected to provide suitable spaces for training sessions, such as halls or meeting rooms, and to make use of standard IT equipment, including laptops and projectors. It was also assumed that existing administrative capacity would be sufficient to circulate communications and manage timetabling and that established safeguarding and school policies were already in place, such as oversight by the designated safeguarding lead and compliance with data protection requirements. Finally, workforce information, including staff role counts, was assumed to be readily available from schools to support planning and monitoring.

Setup costs

Setup costs were one-off expenditures incurred at the start of the programme. For the delivery team, this covered equipment and materials required to launch delivery, such as digital equipment and preparatory materials. Schools also reported some additional one-off expenditures linked to programme launch, such as minor administrative or resource costs.

Recurring costs

Recurring costs were those incurred each time the programme was delivered. For the delivery team, these included staff delivery time, a proportion of office rental allocated to programme delivery, and travel and related expenses for staff members delivering sessions in person. For schools, recurring costs included staff participation in training, cover costs to release teachers for training and coordination activities required to organise and oversee delivery. These estimates were based on survey data collected from intervention schools averaged across available responses to provide representative annual costs.

Method

At the end of the first academic year of the intervention (Year 1; 2023/24) and again at the end of the second year (Year 2; 2024/25), schools were asked to complete a cost survey. The survey collected information on staff roles, time spent coordinating the training, use of cover staff and the total number of staff members in each role. In total, cost survey data was available from 17 intervention schools: four schools completed the survey at Year 1 only, six schools at both Years 1 and 2, and seven schools at Year 2 only, retrospectively reporting costs across the whole trial. To estimate the cost of coordination and cover staff, data from schools that provided information for either Year 1 or Year 2 were combined, and average values were calculated to provide a representative annual cost per school.

Staff participation costs were estimated using publicly available national pay scales and employer contribution rates. Support staff were costed using the National Joint Council (NJC) pay spine (Local Government Association, 2024), teaching staff using the DfE’s classroom teacher pay scales, and senior leadership roles using the national leadership pay ranges (DfE, 2024b). PI leads were costed using a combination of UPS midpoints and teaching and learning responsibility allowances.

Data on staff role counts were available from 12 intervention schools. These figures were averaged across schools and cross-referenced with national workforce data to ensure consistency. Training was structured so that all staff had access to seven hours of sessions, while SLTs and PI leads had access to an additional 18.5 hours (25.5 hours in total). For costing purposes, senior leadership was defined broadly to include non-teaching leadership roles, PI leads, assistant and deputy headteachers, headteachers, heads of year and lead practitioners.

Costs were calculated by multiplying the average number of staff in each role by the maximum training hours available and applying the relevant hourly rate (inclusive of employer on-costs). This provided a comprehensive estimate of the total resources required for staff participation based on full compliance assumptions.

The programme delivery team maintained detailed records of all costs incurred during implementation. These records were collated and summarised, and the resulting information was shared with the evaluation team to ensure transparency and to provide a robust basis for the cost analysis.

The detailed breakdown of these assumptions is presented in Table 23, while the monetary values by the programme delivery team and schools are summarised in Table 24.

Table 23: Full list of cost assumptions and results

Category	Description
Staff	
Labour cost assumptions	Support staff roles were costed using the National Joint Council pay spine (Local Government Association, 2024). Teaching staff were costed using the Department for Education's published pay scales (Department for Education, 2024), including the main pay range (M1–M6), upper pay scale (UPS), unqualified teacher rates, trainee teacher rates and leadership ranges. Pastoral and inclusion (PI) leads were costed using UPS midpoints plus teaching and learning responsibility 2 allowances. Employer on-costs (National Insurance and pension contributions) were applied using statutory rates for 2023/24 and 2024/25. Hourly rates were derived by dividing annual salary plus on-costs by typical working weeks and contracted weekly hours for the role.
Staff training time	Assuming full compliance, all staff were assumed to attend seven hours of training. Senior leadership teams and PI leads were assumed to attend an additional 18.5 hours (25.5 hours total). Costs were calculated by multiplying average staff counts per role by training hours and hourly rates. Based on full compliance, staff participation in training averaged £47,964 per school. If all 31 schools complied, the total cost would be £1,486,894.
School coordination	Schools reported time spent coordinating training, averaging £1,713 per school. As this activity was required each year, it was treated as a recurring cost.
Cover staff	Schools reported average cover costs of £2,370 per school to release teachers for training.
Intervention trainers	Delivery staff costs were reported at £493,886 in total, equivalent to £15,929 per school.
Programme	
Design and preparation	Programme design and preparation were undertaken by the delivery team and embedded within staff delivery costs.
Buildings and facilities	
Office rent	A proportion of office rental was allocated by the programme delivery, averaging £842 per school. Training was delivered on school premises during the school day, so no additional school facility costs were incurred.
Materials and equipment	
Equipment and materials	Initial purchases of digital equipment and resources by the programme delivery team cost £30,108 in total, equivalent to £972 per school.
Other inputs	
Travel and expenses	Programme delivery staff travel costs averaged £19 per school.
Other school costs	Other school costs, such as hiring costs for cover staff, averaged £280 per school.

Table 24: Cost of implementing the programme, by item and delivery partner

Cost item	Setup or recurring?	Programme delivery team	Schools	Total
Staff				
Intervention trainers and delivery leads	Recurring	£493,886	–	£493,886
Staff coordination	Recurring	–	£53,103 (1,713 × 31)	£53,103
School cover costs	Recurring	–	£73,470 (2,370 × 31)	£73,470
School staff participation in training	Recurring	–	£1,486,894 (47,964 × 31)	£1,486,894
Buildings and facilities				
Proportion of office rent	Recurring	£26,094	–	£26,094
Materials and equipment				
Equipment and materials	Setup	£30,108	–	£30,108
Other costs				
Travel and related expenses	Recurring	£596	–	£596
School other costs	Setup	–	£8,680 (280 × 31)	£8,680
Total costs				
Total setup costs	–	£30,108	£8,680	£38,788
Total recurring costs	–	£520,576	£1,613,467	£2,133,943
Grand total	–	£550,684	£1,622,147	£2,172,731
Number of participants (staff)	–	–	3,720 staff (120 x 31)	3,720
Cost per participant (staff)	–	–	–	£584
Setup costs per school	–	–	–	£1,252
Recurring cost per school	–	–	–	£68,838
Total cost per school	–	–	–	£70,090
Total cost for the delivery team (per school)				£17,764
Total cost for the school (per school)				£52,327

Conclusion

Table 25: Key conclusions (authored by the YEF)

Key conclusions
MGDAS had no impact on children's externalising behaviours. After the programme, children in MGDAS schools had similar levels of externalising behaviours compared to their counterparts in the control group. This result has a very low security rating.
MGDAS demonstrated no impact on prosocial behaviour, bullying perpetration, children's perceptions of the impact of their emotional or behavioural difficulties, their perceived relationships with teachers, their perception of peer support, and their perception of the relevance of schoolwork and their sense of agency over it. It had a small negative impact on children's internalising behaviour, school attendance, suspensions and likelihood of being a victim of bullying. It had a small positive impact on children's aspirations and goals. These are secondary outcomes and should be treated with more caution. There is also statistical uncertainty regarding these outcomes.
MGDAS demonstrated a moderate positive impact on teachers' understanding of the underlying causes of behaviour, their confidence in using trauma-informed practices, the extent to which responses to behaviour used trauma-informed practice, teachers' empathy and on reducing teacher burnout. Teachers in MGDAS schools were also slightly more likely to report fulfilment from helping others and have positive feelings about their roles. There was no impact on teacher stress from exposure to trauma. These are secondary outcomes and should be treated with more caution, and all of these outcomes are statistically uncertain.
The trial suffered from very high attrition. 49.9% of children in the trial at randomisation were not included in the final analysis. All impact findings should therefore be treated with significant caution.
Teachers' attendance at training was low. On average, teachers attended only 29% of sessions. High workload, competing pressures, and long and frequent sessions impacted staff attendance. Some staff reported feeling better equipped to support children, while others described a lack of change in culture, policies and practices.

Impact evaluation and implementation process evaluation integration

The overarching research question for the MGDAS evaluation was: how effective is the implementation of a whole-school approach to trauma-informed practice, including senior leadership support, frontline practitioner training and ongoing reflective practice (context), in improving safe social connections (mechanism) and reducing externalising difficulties (primary outcome), thereby reducing the likelihood of young people becoming involved in crime and violence in the future (long-term outcome)?

Evidence to support the logic model

In this section, we provide a high-level summary of the findings based on the logic model (see Appendix C). In the next section, we provide a more detailed interpretation of the quantitative and qualitative findings. Overall, the evaluation provided mixed support for the original logic model. In terms of inputs, 25 schools were retained to the end of the programme and generally completed the sessions, though staff attendance was limited and varied across elements, with smaller targeted groups engaging more successfully than whole-school sessions. For change mechanisms, the direction of effects for staff outcomes generally suggested greater alignment with trauma-informed practice in intervention schools, although these differences were not statistically significant. Elements of the logic model related to staff wellbeing and burnout also showed non-significant effects, suggesting that factors such as having positive feelings about one's role, compassion satisfaction, burnout and secondary traumatic stress were not impacted by the intervention. For the impact evaluation, no evidence was found that pupils experienced

stronger safe social connections with staff or peers in the intervention arm. However, qualitative data suggested that staff felt better equipped to support pupils and reported improved connections with them. Pupils also described supportive relationships with staff, but also barriers such as staff being perceived as unfair or unapproachable. At the level of outcomes, there was no statistically significant effect of the intervention on the primary outcome of externalising difficulties, and most pupil outcomes (including mental health difficulties, prosocial behaviour, attendance, bullying perpetration and educational inclusion) were also non-significant. There was limited evidence of increased bullying victimisation in intervention schools, though this appeared dependent on model specification, as the finding was non-significant when covariates were removed from the model and after multiple comparison correction. Evidence from the SDS data found that intervention schools made shifts towards trauma-informed practice, reducing reliance on punitive sanctions and instead consolidating approaches like reflective conversations, needs-based responses and restorative practices. However, it is noted that only around a third of schools completed the survey at baseline and endline, so the data may be biased in terms of schools that enacted more organisational change. The outcomes of involvement in criminal activities and school attainment were not explored in the current study; however, the data from the trial will be linked to the National Pupil Database to explore longer-term outcomes.

Interpretation

Programme implementation was characterised by modest overall attendance and marked variation across schools. Whole-school elements were consistently less well attended, reflecting the difficulty of securing broad staff participation, whereas targeted sessions for SLTs and PI staff achieved stronger engagement. Qualitative evidence highlighted barriers such as heavy workload, competing priorities and limited time, but also pointed to strategies that facilitated attendance, including compulsory scheduling within directed time and leadership endorsement. While the training was largely delivered as intended in retained schools, adaptations were frequently introduced to accommodate school needs, such as shorter or split sessions, recordings, and additional reflective practice groups. The sequencing of the training was also amended to adapt to school schedules. These modifications illustrate efforts to balance fidelity with feasibility, maintaining coverage of core content while tailoring delivery to local contexts.

Implementation was further shaped by systemic pressures and structural constraints. Barriers included the substantial time commitment required, late programme introduction, limited lead-in for relationship building, and challenges associated with online delivery, compounded by high staff workload and capacity limitations. The highly pressured school context, which affected schools' ability to engage with the intervention and evaluation, presents fundamental questions for school-based trauma-informed practice programmes. The highly pressured context is precisely why programmes like MGDAS are needed, to enable staff to promote a safe environment for pupils and to respond to behaviour in a trauma- and relationally-informed approach. Such programmes are also needed to provide wellbeing support to school staff. However, the same pressures that create this need also act as barriers to delivery, potentially meaning that those most in need are least able to engage.

Existing evidence demonstrates that one-off trainings or fragmented initiatives are insufficient to achieve sustained change, with organisational infrastructure, leadership commitment, and longer-term support required to embed practice across whole schools (Maynard et al., 2019; Gaffney et al., 2021; Avery et al., 2020). The findings from this evaluation therefore reinforce the wider literature: intensive trauma-informed programmes can provide essential support for staff and pupils, but insights from the qualitative data found that their effectiveness depends on systemic investment and sustained implementation, which remain

difficult to achieve in highly pressured school contexts. High attrition and low attendance in the trial further reflected the pressures schools were operating under, highlighting the challenges of sustaining engagement with intensive programmes in real-world settings.

In terms of the findings from the impact evaluation, there was no statistically significant effect of the intervention on the primary outcome of pupil SDQ externalising difficulties within the primary analysis model. The absence of a statistically significant effect on the primary outcome was consistent across multiple analytical models – this absence of a statistically significant effect persisted when covariates were removed, indicating that the result was not dependent on model specification. Missing data analyses (including MCAR tests, predictors of missingness in the model, MI and non-response weighting) produced estimates that aligned with the complete case analysis, suggesting that the absence of a statistically significant effect is unlikely to be explained by biases arising from missingness. Compliance analysis using a CACE framework also showed no statistically significant effect, indicating that non-compliance with the training programme did not mask a potential impact. Additional robustness checks, including models incorporating midline externalising scores, moderation by staff empathy-related behaviours, and exploratory mediation via safe social connections, again yielded non-significant results. Subgroup analyses (interactions by ethnicity, FSM eligibility and SEND status) were also non-significant when correcting for multiple comparisons, indicating that the absence of a statistically significant effect was consistent across different pupil groups. These findings suggest that the lack of intervention effect is robust across alternative specifications, missing data assumptions, compliance scenarios and pupil subgroups, and therefore provide evidence that the programme did not reduce externalising difficulties.

The majority of secondary pupil and staff outcomes also showed non-significant effects of the intervention. For staff outcomes, effects for intervention schools generally pointed to greater alignment with trauma-informed practice (i.e. positive effect sizes on the Response to Problem Behaviour subscale of the ARTIC-35), but these effects were not statistically significant. This pattern may reflect the relatively low staff response rate or the possibility that schools require more time to fully embed new practices. The qualitative evidence demonstrates how changes were enacted in practice. Staff consistently reported feeling better equipped to support pupils, drawing on concepts such as co-regulation, restorative conversations and recognition of trauma responses. Staff further emphasised a shift away from punitive sanctions towards strategies centred on de-escalation and reflection. These accounts align closely with the direction of the quantitative findings, indicating that trauma-informed approaches were beginning to embed within everyday professional practice.

The intervention effect on the understanding causes of problem behaviour subscale of the ARTIC-35 was positive and aligned with increased understanding of the processes and consequences of trauma, but it did not reach statistical significance after adjustment for multiple comparisons. Nevertheless, the qualitative evidence points to a similar pattern: staff consistently described how the programme consolidated their understanding of trauma and attachment and fostered a shared language for interpreting behaviour as linked to pupils' circumstances rather than deliberate misbehaviour. These accounts suggest that, even in the absence of a statistically reliable effect, the reframing of behaviour was experienced by staff as a meaningful outcome of the training.

For pupils, there was evidence that bullying victimisation was higher in schools that received the intervention; however, this should be interpreted with caution as the effect was not statistically significant when covariates were removed from the model and after adjusting for multiple comparison correction. The qualitative data provided limited insight into this issue, although these points cannot be directly linked to

the MGDAS programme. Some pupils reported that staff were not always proactive or available in responding to bullying or misbehaviour, sometimes interpreting bullying behaviour as joking between peers and not consistently acting on requests for support. Although not directly evidenced by this evaluation, it is possible that adopting a less punitive and more relational approach to managing bullying behaviour may have contributed to the persistence of such behaviour. The evidence indicates that the apparent increase in bullying victimisation should be treated as tentative, and further research is needed to understand how trauma-informed approaches interact with school responses to bullying.

Beyond these outcomes, qualitative evidence pointed to changes not captured in the quantitative analysis but relevant to the logic model. Staff and pupils described increases in wellbeing-focused activities, signposting and referrals to support services, indicating strengthened mental health provision. From the SDS data, intervention schools also reported expanding both universal and targeted approaches to pupil support and showed greater increases in engaging parents and caregivers compared to control schools. These developments align with the programme's aim of building a relational, whole-school culture and understanding of young people's needs, even though they were not reflected in significant quantitative effects.

Quantitative data did not show improvements in staff wellbeing, with non-significant findings on ProQoL scales of burnout, compassion satisfaction and secondary traumatic stress. As revealed in the data from the SDS, although control schools reported larger increases in staff wellbeing provision over time, intervention schools appeared to start from a stronger baseline with more structured support, leaving less scope for expansion. Qualitative accounts indicated that staff valued the programme's emphasis on wellbeing and described greater awareness of the importance of supporting staff mental health, including recognition that staff may themselves have experienced trauma, alongside some local initiatives such as reflective practice sessions or wellbeing charters. However, schools reported limited concrete changes and highlighted the need for greater prioritisation and structural supports, such as regular supervision opportunities. These findings suggest that while the programme raised awareness and prompted some initiatives, broader improvements in staff wellbeing were constrained by sector-wide pressures and limited capacity, with change more strongly driven by wider developments in the education sector than by the intervention itself. Given that staff wellbeing is a central mechanism in trauma-informed practice, the limited evidence of change in this domain represents a blocker to the programme's theory of change, underscoring the need for systemic investment to support staff if pupil-level outcomes are to be achieved.

It is noted that the generalisability of the findings is constrained by the high level of attrition observed in the evaluation. There was an overall attrition rate of 49.9%, and attrition was higher in the intervention arm (56.7%) compared with the control arm (44.1%). Losses were attributable to school withdrawal, pupil changes in schools and challenges in securing endline data collection. Importantly, attrition differed significantly by ethnicity, with Asian pupils disproportionately retained and White pupils disproportionately lost to follow-up. Although subgroup analyses did not identify statistically significant differences in primary outcome effects by ethnicity, the uneven retention across groups raises concerns about representativeness and the external validity of the results. These patterns suggest that while the intervention was implemented across a diverse set of schools, the analytical sample may not fully reflect the demographic composition of the original trial population. The disproportionate withdrawal in the intervention arm and differential retention by ethnicity introduce potential bias and limit the extent to which findings can be generalised across racial and ethnic subgroups and to the wider school population. Future evaluations of trauma-informed practice programmes should therefore prioritise strategies to minimise attrition and

ensure equitable retention across demographic groups, in order to strengthen the external validity and applicability of results.

Overall, quantitative evidence showed no reduction in pupil externalising difficulties. Qualitative findings suggest that staff felt better equipped to use trauma-informed strategies such as de-escalation and restorative approaches, but they also highlighted that embedding these practices across the whole school was a longer-term process. Staff noted that entrenched behaviour policies, variation in staff attitudes and the “behind the scenes” nature of many changes limited the visibility and immediate impact for pupils.

The qualitative data also underscored the importance of senior leadership commitment in enabling systemic change. Where leadership actively supported attendance, reinforced consistent messaging, and drove policy adaptation, staff felt more able to enact change. However, in many schools, competing priorities, high workload and limited capacity constrained implementation. Overall, these findings suggest that while the programme influenced specific aspects of staff practice, translation into pupil-level outcomes and consistent change across all trauma-informed domains requires sustained senior leadership support, policy alignment and sufficient time for practices to embed.

The results of this evaluation add to a growing but mixed body of evidence on trauma-informed practice in schools (Gaffney et al., 2021). Consistent with prior studies, the intervention was successfully implemented in terms of session delivery, but staff attendance was low and varied, reflecting well-documented challenges of workforce capacity and competing priorities in educational settings (Maynard et al., 2019; Gaffney et al., 2021). For staff outcomes, effects for intervention schools generally pointed toward greater alignment with trauma-informed practice, although these differences were not statistically significant and did not translate into consistent changes in pupil outcomes. This aligns with earlier reviews that highlight the difficulty of achieving measurable pupil-level change through organisational-level interventions, particularly when implementation fidelity and engagement are mixed (Avery et al., 2020; Gaffney et al., 2021).

Limitations and lessons learned

There were several limitations that should be considered when interpreting the findings. First, attrition reduced the completeness of the dataset at both the school and pupil levels. At the school level, eight of the 62 randomised schools (12.9%) withdrew before endline data collection, with higher dropout in intervention compared to control schools (intervention: six schools, 19.4%; control: two schools, 6.5%). School withdrawal was primarily due to staffing and capacity constraints, leadership changes, or wider organisational pressures during the study period, which limited schools’ ability to engage fully with research activities. At the pupil level, attrition was substantial. Of the 13,900 randomised pupils, 6,941 (49.9%) were lost to follow-up, with higher attrition in the intervention arm (56.7%) than in the control arm (44.1%). Pupil-level loss to follow-up reflected a combination of school withdrawal, pupils moving schools, and difficulties engaging schools to complete endline measures. The study was originally powered to detect an MDES of 0.17, assuming 6,372 pupils per arm and 15% attrition. However, the final analytical sample comprised 2,779 pupils in the intervention arm and 4,180 in the control arm. This reduction in sample size, alongside differential attrition between trial arms, is likely to have reduced the statistical power to detect intervention effects.

Second, due to the primary focus on pupil data collection, follow-up on SDS responses and cost data was limited. However, to facilitate completion by school staff, the survey was shortened for the final data collection period.

Third, there were challenges in recording and analysing staff attendance data at training sessions, and the data associated with two sessions had to be discarded due to the possibility that trainers had not prompted attendance form completion. Additionally, in 8% of cases, forms could not be matched to a training session with the same date, name, and at the same school; this could be due to user error in completing the forms and could be avoided in future evaluations by creating unique identifiers for each session which are automatically completed using (e.g. a URL with a pre-filled field).

Fourth, the aim was to select four intervention schools across at least three LAs as qualitative case study schools. Selection sought to achieve variation in urban/rural setting, ethnic diversity, proportion of pupils eligible for FSMs and number of pupils with SEND support. While pupil interviews were drawn from these case study schools, recruitment of school staff for interviews was not restricted to the four case study schools due to time and capacity constraints. Staff participation was open to all intervention schools, which introduced potential selection bias, as participation was voluntary and may have been more likely among schools engaged with the programme.

Fifth, the timeframe for the evaluation also posed constraints. As the intervention focused on staff practice and broader cultural change, its effects on pupils were unlikely to emerge quickly. The relatively short period between baseline and follow-up may therefore have limited the extent to which changes at pupil level could be observed. In addition, the multi-component nature of trauma-informed practice programmes makes it challenging to define what constitutes meaningful change and to determine what effects are reasonable to expect within a given timeframe.

Future research and publications

Following the conclusion of the YEF-funded study, Anna Freud will link the pupil-level quantitative data to national records from the National Pupil Database via the Office for National Statistics Secure Research Service. This will allow examination of longer-term impacts on school exclusions and academic attainment, addressing future research questions about the sustained effects of the intervention. These analyses fall outside the timeframe of the current study and will be conducted by Anna Freud as non-costed work, with the intention of producing a peer-reviewed journal article. Further research could investigate the structural and contextual factors that contribute to observed disparities in outcomes, including differences in school resources, teacher–student relationships, classroom practices and broader community influences. Understanding how these factors interact with pupil characteristics and school-level dynamics could help identify the mechanisms driving inequities and inform the design of more targeted and equitable approaches in future school-based interventions. A critical area could be assessing the extent to which trauma-informed practice reduces the racial disparities in exclusionary practices (e.g. suspensions, exclusions) and whether the adoption of trauma-informed practice results in a reduction of punitive measures for students of colour, whose behaviour is often pathologised rather than viewed through a trauma lens (Griffiths, 2023). It is noted that, because suspensions and exclusions are relatively rare events, examining these patterns would require very large samples to ensure adequate statistical power. Future research should also examine implementation factors that facilitate culturally responsive trauma-informed practices, as published descriptions of trauma-informed schools frequently lack components related to reducing the likelihood of trauma revictimisation for racially minoritised students through school policies and procedures reform (Brown et al., 2024).

References

Appleton, J. J., Christenson, S. L., Kim, D. and Reschly, A. L. (2006). 'Measuring cognitive and psychological engagement: Validation of the Student Engagement Instrument', *Journal of School Psychology*, 44(5), pp. 427–445.

Avery, J. C., Morris, H., Galvin, E., Misso, M., Savaglio, M. and Skouteris, H. (2021). 'Systematic review of school-wide trauma-informed approaches', *Journal of Child & Adolescent Trauma*, 14(3), pp. 381–397.

Baglivio, M. T., Wolff, K. T., DeLisi, M. and Jackowski, K. (2020). 'The role of adverse childhood experiences (ACEs) and psychopathic features on juvenile offending criminal careers to age 18', *Youth Violence and Juvenile Justice,* 18(4), pp. 337–364.

Baker, C. N., Brown, S. M., Overstreet, S., Wilcox, P. D. and New Orleans Trauma Informed Schools Learning Collaborative (2021). 'Validation of the Attitudes Related to Trauma-Informed Care Scale (ARTIC)', *Psychological Trauma: Theory, Research, Practice, and Policy*, 13(5), pp. 505–513.

Berliner, L. and Kolko, D. J. (2016). 'Trauma-informed care: A commentary and critique', *Child Maltreatment,* 21(2), pp. 168–172.

Blankenstein, N. E., Vandenbroucke, A. R., de Vries, R., Swaab, H., Popma, A. and Jansen, L. (2022). 'Understanding aggression in adolescence by studying the neurobiological stress system: A systematic review', *Motivation Science,* 8(2), p. 133.

Braun, V. and Clarke, V. (2020). 'One size fits all? What counts as quality practice in (reflexive) thematic analysis?', *Qualitative research in psychology*, 18(3), pp. 328–352.

Brown, S. M., Baker, C. N., Hoover, S., Gould, L. F., Drucker, K., VanDeMark, A., Porche, M. V. and Overstreet, S. (2024). 'Common elements of trauma-informed schools and attention to racial equity: A scoping review', *School Mental Health*, 16, pp. 869–892. doi: 10.1007/s12310-024-09721-2.

Cresswell, C. W. and Plano Clark, V. L. (2018). *Designing and Conducting Mixed Methods Research*. 3rd ed. London: Sage.

Department for Education (2024a). 'Suspensions – ethnicity facts and figures', *Ethnicity Facts & Figures*. Available at: https://www.ethnicity-facts-figures.service.gov.uk/education-skills-and-training/absence-and-exclusions/pupil-exclusions/latest/ (Accessed: December 2025).

Department for Education (2024b). *School Teachers' Pay and Conditions Document 2024 and Guidance on School Teachers' Pay and Conditions*. London: DfE. Available at: https://www.gov.uk/government/publications/school-teachers-pay-and-conditions. (Accessed: October 2025)

Department for Education (2025). '*Special educational needs in England: Academic year 2024/25'*. Available at: https://explore-education-statistics.service.gov.uk/find-statistics/special-educational-needs-in-england/2024-25 (Accessed: October 2025).

Department for Education (2025). 'Schools, pupils and their characteristics: Academic year 2024/25'. Available at: https://explore-education-statistics.service.gov.uk/find-statistics/school-pupils-and-their-characteristics/2024-25 (Accessed: October 2025).

Elia, C., Karamanos, A., Thompson, R., Lewis, G., Zammit, S., Hickman, M., Heron, J. and Cornish, R. (2020). 'Weight misperception and psychological symptoms from adolescence to young adulthood: Longitudinal study of an ethnically diverse UK cohort', *BMC Public Health*, 20, p. 712. https://doi.org/10.1186/s12889-020-08823-1.

Espelage, D. L. and Holt, M. K. (2001). 'Bullying and victimization during early adolescence: Peer influences and psychosocial correlates', *Journal of Emotional Abuse*, 2(2–3), pp. 123–142.

Felitti, V. J., Anda, R. F., Nordenberg, D., Williamson, D. F., Spitz, A. M., Edwards, V., Koss, M. P. and Marks, J. S. (1998). 'Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults: The Adverse Childhood Experiences (ACE) Study', *American Journal of Preventive Medicine,* 14(4), pp. 245–258.

Freedman, D. (1984). 'On bootstrapping two-stage least-squares estimates in stationary linear models', *The Annals of Statistics*, 12(3), pp. 827–842. http://www.jstor.org/stable/2240963.

Gaffney, H., Jolliffee, D. and White, H. (2021). *Trauma Informed Care: Toolkit Technical Report*. London: Youth Endowment Fund.

Gillett-Swan, J. and Lundy, K. (2021). '"Trauma-informed" ideas in English Education: Discussing the scientific evidence base and exploring the discursive and practice effects'. *International Journal of School Social Work,* 6(1), Article 4. https://doi.org/10.4148/2161-4148.1069.

Goodman, R., Meltzer, H. and Bailey, V. (1998). 'The Strengths and Difficulties Questionnaire: A pilot study on the validity of the self-report version.', *European Child and Adolescent Psychiatry*, 7(3), pp. 125–30.

Goodman, A., Lamping, D.L. and Ploubidis, G.B. (2010). 'When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the Strengths and Difficulties Questionnaire (SDQ): Data from British parents, teachers and children', *Journal of Abnormal Child Psychology*, 38, pp. 1179–1191. doi: 10.1007/s10802-010-9434-0.

Hedges, L. V. and Hedberg, E. C. (2007). 'Intraclass correlation values for planning group-randomized trials in education'. *Educational Evaluation and Policy Analysis,* 29(1), pp. 60–87. https://doi.org/10.3102/0162373707299706.

IPPR (2017). *'Half of expelled pupils suffer mental health issues in "burningly unjust system", think-tank finds'*. Available at: https://www.ippr.org/media-office/half-of-expelled-pupils-suffer-mental-health-issues-in-burningly-injust-system-think-tank-finds.

Lang, J. M., Campbell, K. and Vanderploeg, J. J. (2015). *Advancing Trauma-Informed Systems for Children*. Farmington, CT: Child Health and Development Institute.

Lewis, S. J., Arseneault, L., Caspi, A., Fisher, H. L., Matthews, T., Moffitt, T. E., Odgers, C. L., Stahl, D., Teng, J. Y. and Danese, A. (2019). 'The epidemiology of trauma and post-traumatic stress disorder in a representative cohort of young people in England and Wales', *The Lancet Psychiatry,* 6(3), pp. 247–256. https://doi.org/10.1016/S2215-0366(19)30031-8.

Little, R.J.A., 1988. 'A test of missing completely at random for multivariate data with missing values'. *Journal of the American Statistical Association*, 83(404), pp. 1198–1202.

Local Government Association (LGA) (2024). *National Joint Council (NJC) pay scales 2024/25*. London: LGA. Available at: https://www.local.gov.uk/our-support/workforce-and-hr-support/njc-pay-scales.

Maynard, B. R., Farina, A., Dell, N. A. and Kelly, M. S. (2019). 'Effects of trauma-informed approaches in schools: A systematic review', *Campbell Systematic Reviews,* 15(1–2), 10.1002/cl2.1018.

McCrory, E. J., De Brito, S. A., Sebastian, C. L., Mechelli, A., Bird, G., Kelly, P. A. and Viding, E. (2011). 'Heightened neural reactivity to threat in child victims of family violence', *Current Biology,* 21(23), pp. R947–R948.

Neil, L., Viding, E., Armbruster-Genc, D., Lisi, M., Mareschal, I., Rankin, G., McCrory, E. J. and Martin, P. (2022). 'Trust and childhood maltreatment: Evidence of bias in appraisal of unfamiliar faces', *Journal of Child Psychology and Psychiatry,* 63(6), pp. 655–662.

Pumariega, A.J., Jo, Y., Beck, B., Fernandez, J. S., Bryant, B.E., Mian, A. I. and Rothe, E. M. (2022). 'Trauma and US minority children and youth', *Current Psychiatry Reports*, 24, pp. 285–295. doi: 10.1007/s11920-022-01336-1.

Rainer C., Le, H. and Abdinasir, K. (2023). *Behaviour and mental health in schools*. Available at: https://cypmhc.org.uk/wp-content/uploads/2023/06/Behaviour-and-Mental-Health-in-Schools-Full-Report.pdf.

Ritchie, J. and Spencer, L. (1994). 'Qualitative Data Analysis for Applied Policy Research.' In: A. Bryman and R. Burgess (eds) *Analysing Qualitative Data*. London: Routledge, pp. 173–194.

Stamm, B. H. (2010). *The Concise ProQOL Manual*, 2nd Ed. Pocatello, ID: ProQOL.org.

Substance Abuse and Mental Health Services Administration (SAMHSA) (2014). *SAMHSA's Concept of Trauma and Guidance for a Trauma-Informed Approach*. Rockville MD: Substance Abuse and Mental Health Services Administration. HHS Publication No. (SMA) 14-4884.

Valliant, R., Dever, J. and Kreuter, F. (2013). *Practical Tools for Designing and Weighting Survey Samples*. New York: Springer Science Business Media.

van Buuren, S. and Groothuis-Oudshoorn, K. (2011). 'MICE: Multivariate imputation by chained equations in R.'*, Journal of Statistical Software*, 45(3), pp. 1–67. https://doi.org/10.18637/jss.v045.i03.

Van Voorhees, E. and Scarpa, A. (2004). 'The effects of child maltreatment on the hypothalamic-pituitary-adrenal axis', *Trauma, Violence, & Abuse,* 5(4), pp. 333–352.

Villadsen, A., Libuy, N. and Fitzsimons, E. (2025). *Adverse and Positive Childhood Experiences and Their Associations with Children's Involvement in Violence: Analyses of Data from the Millennium Cohort Study*. Youth Endowment Fund. (Accessed: 8 December 2025).

Watson, K.R. and Astor, R. A. (2025). 'A critical review of empirical support for trauma-informed approaches in schools and a call for conceptual, empirical and practice integration', *Review of Education,* 13(1), e70025. https://doi.org/10.1002/rev3.70025.

Ye, C., Adam, M. and Borm, G. F. (2014). 'Estimating treatment effects in randomised controlled trials with non-compliance: A simulation study', *BMJ Open,* 4(6), e005354. https://doi.org/10.1136/bmjopen-2014-005354.

Youth Endowment Fund (2025). *Racial Disproportionality in Youth Justice — Final Report*. Available at: https://youthendowmentfund.org.uk/wp-content/uploads/2025/02/YEF_Racial_Disproportionality_FINAL.pdf (Accessed: December 2025).

Appendix A: Recruitment documents: Memorandum of Understanding, Pupil Information Sheet and Pupil Privacy Notice

Memorandum of Understanding

More Good Days At School: Building relationships to promote health, happiness and learning

Between

Anna Freud

4-8 Rodney Street

London, NW 9JH

And

[school]

Background

Anna Freud has been commissioned by the Youth Endowment Fund (YEF) to deliver the "More Good Days At School: Building relationships to promote health, happiness and learning" study. This study will implement and evaluate mental health support across secondary schools with the aim of contributing to the evidence-based for school-based mental health support. The study will run from December 2023 – September 2025. A full timeline of key dates for schools can be found in Appendix 1.

Purpose of this document

This Memorandum of Understanding (MoU) outlines what will be required of schools participating in the "More Good Days At School: Building relationships to promote health, happiness and learning" study. In order to take part in the study, a senior leader at the school must add their signature to this form and return by email to MGDAS@annafreud.org within two weeks of receiving it. Please put your postcode and the name of your school as the email subject heading (e.g., MoU – NW1 9JH – Freud High School). This document should be read in conjunction with other documents provided, such as the information sheets and guidance documents.

Requirements of schools

We are conducting a randomised controlled trial, which means that schools will be randomly allocated to either the intervention group (receiving whole-school trauma-informed staff training in addition to existing support), or the control group (continuing with existing support). Random allocation means that schools will **not** be able to choose which group they are in. Schools will only be informed of their group allocation after both baseline data collection and randomisation have been completed (February 2024). If the school is allocated to the intervention group, the school will be contacted by the training providers (Knowledge Change Action and Warren Larkin Associates Ltd.) to arrange and conduct the training sessions. Data will be

collected from both the intervention and the control groups. For a full timeline of the project, please review Appendix 1. Schools will not be allowed to switch to a different group after they have been allocated, as this would compromise the integrity of the research.

If your school decides to take part in the study, we will ask the following:

1. Year 8 and 9 pupils will be asked to complete a survey 3 times over the next two years about how pupils are doing, what it's like at school, how they get on with people at school, what they think about wellbeing support at school, as well as demographic information.
2. All school staff members will be asked to complete a survey 3 times over the next two years that asks about their well-being, what practices they adopt in supporting young people and their experiences with it, as well as demographic information.
3. Pupils and school staff from a small number of schools will be asked to speak to a researcher one-to-one about their experiences during the study. This discussion will be audio-recorded.

Participation is voluntary, and everyone will be asked to provide consent before taking part. If pupils and staff choose not to take part, it will not affect their rights. Parents/carers will have the option for their child to opt-out of the study before baseline data collection. Until July 2025, pupils and staff can withdraw at any time and their information with be deleted unless permission is given to keep the data. If pupils or staff have taken part in a one-to-one discussion, they will be able to withdraw their data until 9 months later.

Results from this project will be published in reports, such as on Anna Freud's, UCL's, and Youth Endowment Fund's websites. The final report will not contain any personal information about the people who took part in the study, and it will not be possible to identify individuals from the report.

Role of the lead contact at school

The lead contact at the school will be responsible for communicating the requirements of the project to all participating school staff once the school has been randomly allocated to its research group.

The lead contact will also be asked to do the following:

- Answer queries from the evaluation team within a reasonable timeframe.
- Notify the research team if there are any issues that could prevent the execution of the research trial.
- If the school has to withdraw from the study for unavoidable reasons, they will notify the research team.
- Store any data securely and confidentially, and in accordance with the Data Protection Act (1998), and the General Data Protection Regulation (GDPR).
- Send information to the school's Safeguarding Lead to ensure students are supported in the event they become distressed at any point during the study.
- Complete and send the adverse events information form to Anna Freud in the unlikely event that an adverse event may occur.
- Complete two types of surveys at multiple time points: one describing wellbeing support provided at your school, and one describing costs incurred as a result of taking part in the MGDAS

programme activities (such as staff hours and facilities), if the school is randomised to receive the training.

- Securely send pupil school data to the evaluation team in July 2024 and July 2025.
    - This includes the pupils' Unique Pupil Number, date of birth, gender, ethnicity, whether they are looked after by the local authority, whether they have any special educational needs, whether they have a disability, school attendance, attainment, and exclusions.

Ethics approval

UCL Ethics Committee has approved the study. All research is looked at by an independent group of people, called a Research Ethics Committee, to protect your interests and safety. This research has been reviewed and given a favourable outcome by University College London Research Ethics Committee (no. 14037/012).

Data security

The UK GDPR basis for processing these data is legitimate interest (Article 6(1)(f)) and research purposes (Article 9(2)(j)) and DPA (2018) Schedule 1 part 1 paragraph 4. This allows us to process personal data for research or evaluation purposes where appropriate care is taken to manage the data securely.

Personal data (e.g., names, contact details, survey responses, discussion responses audio recording) will be held securely until the end of the project (winter 2025).
After the study has ended (winter 2025), we will link the data we have collected with other data from the Department for Education (DfE). This is so that we, and future researchers, can look at long-term impacts of this programme using other sources of information from educational and criminal datasets. The DfE will replace all identifying information about the young people who have taken part in the study (their name, gender, date of birth, home address) with the young person's unique Pupil Matching Reference number in the DfE's National Pupil Database. Once this has been done, it is no longer possible to identify any individual person from the study data. This process is called pseudonymisation. Once this has been done, we (Anna Freud and UCL) will delete identifying information, except for consent forms. Consent forms (with names on them) will be held as long as the data are held, so for at least 10 years (December 2035).

Once we (Anna Freud and UCL) have finished our study (winter 2025) and transfer the data to the DfE, we hand over control to the Youth Endowment Fund (YEF) for protecting the personal information. The DfE will transfer the pseudonymised information to the YEF archive, which is stored in the Office for National Statistics' Secure Research Service. The YEF is the 'controller' of the information in the YEF archive. By maintaining the archive and allowing approved researchers to access the information in the archive, the YEF is performing a task in the public interest and this gives the YEF a lawful basis to use personal information.
The Secure Research Service is an extremely protected and controlled environment - it would be technically almost impossible (and also illegal) for anyone to link the information back to an individual person. For more information on the Office for National Statistics' Secure Research Service, you can visit their website: https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/secureresearchservice

Information in the YEF archive can only be used by approved researchers to explore whether "More Good Days At School: Building relationships to promote health, happiness and learning", and other programmes funded by YEF, had an impact over a longer period of time. Using the unique Pupil Matching Reference

numbers added to the data by the DfE, it will be possible to link the records held in the YEF archive to other public datasets such as education and criminal justice datasets. This will help approved researchers to find out the long-term impact of the projects funded by YEF because they'll be able to see, for example, whether being part of a project reduces a child's likelihood of being excluded from school or becoming involved in criminal activity.

The YEF will keep information in the YEF archive for as long as it is needed for research purposes. Data protection laws permit personal information to be kept for longer periods of time where it is necessary for research and archiving in the public interest and for statistical purposes. The YEF will carry out a review every 5 years to assess whether there is a continued benefit to storing the information in the archive, based on its potential use in future research.
Once information goes into the YEF archive it can no longer be deleted as that would affect the quality of the archived data for use in future research.

The information from audio recordings will be kept for no longer than 9 months. The audio recordings will be shared with a company called The Transcription Service for the purposes of writing up the spoken answers into words. There will be a data sharing agreement in place before they write-up the audio recordings. We will make sure that they keep the data safe, in line with UK data protection law. When the recording has been written up, we will replace your name with a number and the recording will be deleted. In group discussions, we ask everyone to respect each other's privacy and to not repeat what is discussed with other people. As soon as the audio recordings have been written up in text, the audio recording will be deleted, and it will not be possible to identify individual people from the write up.

Safeguarding

Occasionally, someone may feel upset about a question or issue that arises during the study. If a pupil or staff member feels upset by any of the questions they are asked as part of this study, they should tell a researcher (using the contact details on the information sheet) or your school's safeguarding or pastoral lead. There are also external support services such as The Samaritans (Tel. 116 123, www.samaritans.org), Childline (Tel. 0800 1111, www.childline.org.uk), or Education Support (Tel. 08000 562561, https://www.educationsupport.org.uk/get-help/help-for-you/helpline/).

We will treat the information that pupils or staff share with us as confidential, but we may have to break confidentiality if they tell us something that makes us concerned about them or others being at risk. If this happens then we will try to discuss the issue with them first.

Contact details

The Project Lead (Julian) is based at Anna Freud and UCL and can be contacted by email if you have any questions about this project, if there are any problems, or if you have any complaints: Julian.Childs@ucl.ac.uk. The general e-mail address for the project research team is MGDAS@annafreud.org.

If you agree to take part, please tick each point below to show you agree to the terms and conditions of the study. Please also ensure that the school name and address is completed on the first page. A senior leader at the school must add their signature to this form and return by email to MGDAS@annafreud.org within two weeks of receiving it. Please put your postcode and the name of your school as the email subject heading (e.g., MoU – NW1 9JH – Freud High School).

☐ **I have read the points in the Memorandum of Understanding (MoU).**

☐ **I agree to all the points in the MoU and am happy for my school to be involved.**

Signature:

Name of signatory:

Job role:

Date:

Appendix 1

Project timelines

Baseline data collection: December 2023 – February 2024		
	Task	**Date**
1.	Send details of additional school contact (if the lead is unavailable) and Safeguarding Lead to Anna Freud.	By 21 December
2.	Send the parents/carers information sheet link (via e-mail) including opt-out consent form link to parents/carers of all pupils in Years 8 and 9.	Send by 21 December and collect until 8 January
3.	Organise introduction of the study to all pupils in Years 8 and 9 via assemblies (using our pupil study animation video) and send the pupil information sheet (via print-out or e-mail).	8-15 January
4.	Send to all staff members the study protocol training PowerPoint.	8-15 January
5.	Sign the Memorandum of Understanding and send to Anna Freud.	Within 2 weeks of receiving it
6.	Anna Freud to inform lead contact which students have opted out.	8 January 2024

7.	Allocate tutor time (approx. 45 minutes) to share the baseline survey with all pupils in Years 8 and 9. Send tutors the guidance on completing measures in schools document and pupil survey link. Inform tutors which students have opted out.	8 January until 2 February
8.	Distribute staff survey link and regular reminders to all school staff members.	8-15 January
9.	Lead contact to complete survey describing wellbeing support provided at your school.	15 January until 2 February
10.	Anna Freud to inform schools the outcome of the randomisation (intervention or support as usual).	By 19 February
11.	If assigned intervention group, the programme team will be in contact regarding the training.	19 February onwards

Please find below a general outline of the tasks to be completed for the mid-programme follow up in 2024 and the end of programme follow up in 2025. More detailed tasks and specific dates will be provided closer to the time.

Mid-programme follow up: June 2024 – February 2025		
1.	Mid-programme follow-up data collected (pupil and staff survey)	June - July 2024
2.	School contact to complete survey describing wellbeing support provided at your school	July 2024
3.	Local school data (including cost survey) received (please see Memorandum of Understanding)	July 2024
4.	School staff and pupil interviews in a small number of schools	January – February 2025
End of programme follow up: June – July 2025		
1.	End of programme follow-up data collected (pupil, staff and support description survey)	June - July 2025
2.	School contact to complete survey describing wellbeing support provided at your school	June - July 2025
3.	Local school data (including cost survey) received (please see Memorandum of Understanding)	July 2025

Version 2, 30th October 2023

Pupil agreement to take part

Primary Researcher: Julian Edbrooke-Childs, email: Julian.Childs@ucl.ac.uk
Research Team: MGDAS@annafreud.org
Data Protection Officer contact: dpo@annafreud.org; data-protection@ucl.ac.uk
UCL Research Ethics Committee approval number: 14037/012

Summary

- Please read this document carefully and the document about how we use and protect your information.
- Your school is taking part in a project about how to best support young people to be healthy and happy.
- All schools taking part are continuing with existing wellbeing support they provide. In half of the schools, staff will also be trained to work with pupils differently.
- To look at the impact of this new way of working with pupils, we will:
  - Ask you to complete 3 surveys over 2 years;
  - Look at information from the school's records;
  - You will have the option of speaking to a researcher one-to-one;
  - After the end of the project (winter 2025), the information we collect will be linked with other information using a secure, specialist service. This is so that we, and future researchers, can look at long-term impacts of this programme using other sources of information from educational and criminal datasets.

1. What is this project?

Hello, and thank you for reading this information sheet. My name's Julian. I'm leading a research project to look at the impact of support for young people to be healthy and happy at school.

I'm passionate about this project for many reasons. Everyone is different, and all of us are impacted by the things we've experienced. This study is looking at how schools support young people's wellbeing and help pupils to be healthy and happy no matter what they've experienced in the past or are currently struggling with in life.

To look at this, we are working with lots of different schools. All of them are continuing with the existing wellbeing support they provide to young people. In half of the schools, staff will be trained to work with pupils differently by learning more about how things people have experience might impact on them.

We don't know if this different way of working with pupils is better, worse, or the same as existing wellbeing support. This is what we want to find out in the research. The findings from this research will be used to inform how schools support the wellbeing of their pupils in the future.

This information sheet contains more details about who we are, what we're doing, and why we're doing it. It also explains how we'll use your personal information if you agree to take part in our study.

All projects like this are looked at by an independent group of people, called a Research Ethics Committee, to protect your interests and safety. This project has been approved by the University College London (UCL) Research Ethics Committee (ID: 14037/012). The project is being conducted by:

- UCL (https://www.ucl.ac.uk/evidence-based-practice-unit/evidence-based-practice-unit-0) and
- Anna Freud (https://www.annafreud.org/).

The project is funded by the Youth Endowment Fund (https://youthendowmentfund.org.uk/).

2. Why have I been invited to take part?

You have been invited to take part in this study because your school is involved in the "More Good Days At School: Building relationships to promote health, happiness and learning" project.

Pupils in Year 8 (12-13 years) or Year 9 (13-14 years) during the 23/24 academic year at your school are eligible to take part.

3. Do I have to take part?

No! If you do not want to take part in the study, you don't have to. We would like as many people as possible to take part in order make sure we have the best information about what makes a difference for young people. If you choose not to take part, it will not affect your rights or support.

Please think about whether now is a good time for you to take part or not.

If you choose to take part and change your mind later, please speak to your parent or guardian who can contact us. You will not have to give a reason why. You can do this until July 2025. Until then, you can withdraw at any time and your information with be deleted unless you tell us we can keep it.

4. What happens if I take part?

We have already checked if it's OK to take part with your parent or guardian. If you take part, we will ask you to fill in a survey online at school (e.g., in tutor time) that will take no longer than 30 minutes. We will ask you to complete it 3 times over two academic years.

The survey will ask about:

- How you're doing
- What it's like at school
- How you get on with people at school
- What you think about wellbeing support at school
- Your name and contact details
- Your gender and ethnicity

In a small number of schools, researchers will be speaking to pupils one-to-one (at school) about your experiences during the study, and this discussion will be audio-recorded.

Your school will share some of the information that they have about you, which will include:

- Information to help us understand if the different way of working is more or less helpful for pupils with certain characteristics:
  - Date of birth
  - Gender
  - Ethnicity
  - Whether you are looked after by the local authority
  - Whether you have any special educational needs
  - Free school meal eligibility
  - Whether you have a disability
- Information to help us understand the potential longer-term impacts of the different way of working:
  - Your educational record

After the study has ended (21st November 2025), we will link the data we have collected with other types of information (described below).

5. What will happened with the results of the study?

We (Anna Freud and UCL) will write a report about what we find, but the report won't include your name or any other information that could be used to identify you. Results will be shared with schools so that parents/guardians and pupils can view them. The report will be shared with other people; e.g., on Anna Freud, UCL, and YEF websites.

6. Keeping you safe

If you feel upset by any of the questions we ask you, you should tell your parent or guardian, staff you trust at your school, or the researchers. You can find the researchers' contact details in the box on the first page. You can also use the following free sources of support:

- Anna Freud Centre Crisis Messenger: Text AFC to 85258. It is a free 24/7 text service for anyone in crisis anytime, anywhere.
- The Mix: Call 0808 808 4994. The Mix are there to help you take on any challenge that you're facing.
- Samaritans: Free to call service 24 hours a day, call them on 116 123
- Childline: Free to call service 24 hours a day, call them on 0800 1111

We will keep what you tell us private (and we will not share what you tell us with your parent or guardian) unless we think that you or someone else might be at risk of harm. If this happens then we will try talk to you first to tell you why we want to talk to another person or organization.

7. Who can I speak to if I have any questions?

The Primary Researcher (Julian) is based at Anna Freud and UCL and can be contacted by email if you have any questions about this project, if there are any problems, or if you have any complaints: Julian.Childs@ucl.ack.uk. Please also talk to your parent or guardian if you have any complaints and they can help to take it further.

Version 4, 16th January 2025

More Good Days At School: Building relationships to promote health, happiness and learning

How we use and protect your information

Primary Researcher: Julian Edbrooke-Childs, email: Julian.Childs@ucl.ac.uk
Research Team: MGDAS@annafreud.org
Data Protection Officer contact: dpo@annafreud.org; data-protection@ucl.ac.uk
UCL Research Ethics Committee approval number: 14037/012

This information sheet contains more details about how we'll use and protect your personal information if you agree to take part in our study.

When we (UCL and Anna Freud) collect and use participants' personal information as part of the study, we are the controllers of the personal information, which means we decide what personal information to collect and how it is used.

The UK GDPR basis for processing these data is legitimate interest (Article 6(1)(f)) and research purposes (Article 9(2)(j)) and DPA (2018) Schedule 1 part 1 paragraph 4. This allows us to process your personal data for research or evaluation purposes where appropriate care is taken to manage the data securely.

1. What will happen to my information after the project?

a) Information controlled by Anna Freud and UCL

Surveys, data from schools, and educational records
After the study ends (winter 2025), we will share the information we have gathered about everyone who has taken part with the Department for Education (DfE). This is so we can look at longer-term impacts on educational records. The DfE will replace all identifying information about the young people who have taken part in the study (their name, gender, date of birth) with a unique code. This is so the information we have collected can be linked to other information about educational records. Once this has been done, it is no longer possible to identify any individual young person from the study data. This is done using a very secure and protected system that is widely used for this type of thing. Once this has been done, we (Anna Freud and UCL) will delete identifying information, except for consent forms.

Consent forms

Consent forms (with names on them) will be held as long as the data are held, so for at least 10 years (end of 2035). They will be held securely and separately from other information about you.

Audio recordings
The information from audio recordings (if you want to meet with one of the researchers to talk about your experiences during the study) will be kept for no longer than 9 months. We will securely share this audio recording with another organisation (the Transcription Centre: (https://www.transcriptioncentre.co.uk/) who will write up the discussion into text (i.e., as a transcript). As soon as the audio recordings have been written up in as text, the audio recording will be deleted and it will not be possible to identify you from the write-up. Transcripts will be held for at least 10 years (end of 2035).

b) Information controlled by Youth Endowment Fund (YEF) (the project funder)

The Youth Endowment Fund, or YEF for short, is giving us money to do this study. When we (Anna Freud and UCL) finish the study, we'll give your information to the YEF. Before your information goes into the YEF archive, the Department for Education will take out your name and other personal details like your address. This means that no one who looks at the information in the YEF archive will know who you are. They will keep your information in a safe place called the YEF archive. Information will be kept in the YEF archive for as long as it is needed for future research. You can find more information about the YEF archive on the YEF's website: https://youthendowmentfund.org.uk/evaluation-data-archive/.

In the future, people can ask to use the YEF archive to do more studies to find out whether "More Good Days At School: Building relationships to promote health, happiness and learning", and other projects like it, have helped young people. Only researchers who are approved by the YEF will be able to look at the archive. The information from this study will then be able to be linked to other information like educational records and criminal records, to see any longer-term impacts of the project. This information cannot be used for anything other than looking at the impact of programmes like the one in this study. It is illegal to use the information for other reasons, such as by the police.

We will only use your information if the law says it's OK. Because this study is interesting and important to lots of people, the law says we can use your information to do this kind of work.

2. Will my information be shared with anyone else?

In order to ensure that all our (Anna Freud and UCL) records are accurate, we will share information about which surveys (if any) are missing from you with your school. This is so the team can make sure that everything we need from you has been completed. We will also be sharing a survey access code linked to your name with your school, and your school will share the survey access code with you for the final survey. This is so that we can link your responses to the final survey with the previous surveys. We will not share with the school, nor will anyone at the school be able to access, any of the information in your surveys.

As part of the study, we are inviting you to talk to a researcher about your experiences during the study. This will be audio recorded. If you choose to do this, we will let the school know so that they can seek consent from your parents/carers. We will not share with the school, nor will anyone at the school be able to access, your audio recording. The only time that we may have to break confidentiality is if you tell us

something that makes us concerned about you or others being at risk. If this happens then we will try to discuss the issue with you first.

3. What are my data rights?

You have the same rights as adults when it comes to your data. Some of these include:

- Knowing what data we have about you
- The right to correct anything that's wrong
- Being able to ask us to stop using your data
- And more

4. Who can I speak to if I have any questions?

If you feel that we may not be handling your data appropriately or if you have any queries or concerns about this, you (or your parent or guardian) can contact us at:

- Anna Freud: Rachel Hart is the information governance manager and can be contacted on dpo@annafreud.org
- UCL: data-protection@ucl.ac.uk

You and your parent or guardian can also, at any time, make a complaint about our processing of your data to the Information Commissioner: https://ico.org.uk/global/contact-us.

Version 3, 12th March 2025

Appendix B: Effect size estimation

Appendix B Table 1: Effect size estimation

			Intervention group		Control group			
Outcome	Unadjusted differences in means	Adjusted differences in means	n (missing)	Variance of outcome	n (missing)	Variance of outcome	Pooled variance	Pooled standard deviation
SDQ Externalising Difficulties (Adjusted model)	-0.120	0.0564	2985 (3439)	16.090	4429 (3047)	15.520	15.747	3.968
SDQ Externalising Difficulties (Sensitivity model)	-0.126	0.020	2985 (3439)	16.094	4429 (3047)	15.553	15.771	3.971
SDQ Internalising Difficulties (Adjusted model)	0.162	0.158	2984 (3440)	13.740	4429 (3047)	13.933	13.856	3.722
SDQ Impact Score (Adjusted model, no baseline outcome data or random intercept for pupils)	0.059	0.029	2841 (3583)	2.755	4248 (3228)	2.484	2.592	1.610
SDQ Prosocial Behaviour (Adjusted model)	0.079	-0.015	2987 (3437)	4.493	4430 (3046)	4.410	4.443	2.108
SEI Teacher-Student Relationships (Adjusted model)	0.094	-0.091	2737 (3687)	27.194	4078 (3398)	26.910	27.019	5.198
SEI Peer Support for Learning (Adjusted model)	0.064	0.066	2777 (3647)	12.970	4087 (3389)	12.790	12.861	3.586
SEI Control and Relevance of Schoolwork (Adjusted model)	0.205	-0.066	2746 (3678)	25.189	4086 (3390)	25.050	25.104	5.010
SEI Future Aspirations and Goals (Adjusted model)	0.208	0.162	2773 (3651)	10.028	4116 (3360)	10.234	10.154	3.186
IBS Bullying Perpetration (Adjusted model)	0.003	0.054	2772 (3702)	32.640	4038 (3438)	33.280	33.031	5.747
IBS Bullying Victimisation (Adjusted model)	0.353	0.363	2761 (3663)	18.823	4102 (3374)	17.089	17.764	4.215
IBS Bullying Victimisation (Sensitivity model)	0.328	0.327	2761 (3663)	18.911	4102 (3374)	17.221	17.884	4.229
School attendance percentage (Adjusted model)	-0.244	-0.427	3386 (3038)	103.756	4506 (2970)	73.153	86.239	9.286

ARTIC-35 Underlying Causes	0.267	0.094	245 (729)	0.628	457 (727)	0.556	0.579	0.761
ARTIC-35 Self Efficacy	0.134	0.119	245 (729)	0.737	456 (728)	0.845	0.808	0.899
ARTIC-35 Response to Problem Behaviour	0.357	0.109	245 (729)	0.881	456 (728)	0.709	0.767	0.876
ARTIC-35 Reactions to Work	0.160	0.043	245 (729)	0.829	456 (728)	0.714	0.753	0.868
ARTIC-35 Empathy and Control (On the Job Behaviour)	0.262	0.090	245 (729)	0.575	456 (728)	0.615	0.601	0.776
ProQoL Compassion Satisfaction	1.05	0.266	244 (730)	31.274	459 (725)	30.867	31.009	5.569
ProQoL Burnout	-0.529	-0.580	238 (736)	32.343	454 (730)	32.854	32.678	5.716
ProQoL Secondary Traumatic Stress	0.365	-0.055	237 (737)	25.098	452 (732)	24.979	25.020	5.002

Note. Variance was calculated from all available endline data. ARTIC = Attitudes Related to Trauma-Informed Care; IBS = Illinois Bully Scale; ProQoL = Professional Quality of Life; SDQ = Strengths and Difficulties Questionnaire; SEI = Student Engagement Instrument.

Appendix C: Full model output

Appendix C Table 1: Adjusted Primary Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	t	df error	P	Effects	Group
Intercept	2.846	0.322	0.950	2.215	3.478	8.839	6935.000	0.000	fixed	
Group (Intervention)	0.056	0.124	0.950	-0.187	0.300	0.454	6935.000	0.650	fixed	
SDQ Extern (Baseline)	0.586	0.010	0.950	0.567	0.605	59.601	6935.000	0.000	fixed	
Cohort (Cohort 2)	-0.260	0.269	0.950	-0.787	0.267	-0.966	6935.000	0.334	fixed	
SDQ Intern (Baseline)	0.030	0.011	0.950	0.008	0.052	2.647	6935.000	0.008	fixed	
Gender (Female)	-0.022	0.081	0.950	-0.181	0.137	-0.271	6935.000	0.787	fixed	
Gender (Other)	-0.125	0.278	0.950	-0.670	0.420	-0.450	6935.000	0.653	fixed	
Ethnicity (Asian)	-0.395	0.110	0.950	-0.610	-0.180	-3.597	6935.000	0.000	fixed	
Ethnicity (Black)	-0.425	0.160	0.950	-0.740	-0.111	-2.650	6935.000	0.008	fixed	
Ethnicity (Mixed)	-0.237	0.176	0.950	-0.582	0.108	-1.349	6935.000	0.177	fixed	
Ethnicity (Other)	-0.260	0.195	0.950	-0.643	0.123	-1.332	6935.000	0.183	fixed	
LA (1)	-0.089	0.384	0.950	-0.841	0.664	-0.231	6935.000	0.818	fixed	
LA (2)	0.084	0.417	0.950	-0.734	0.901	0.201	6935.000	0.841	fixed	
LA (3)	-0.188	0.239	0.950	-0.656	0.280	-0.788	6935.000	0.431	fixed	
LA (4)	-0.700	0.458	0.950	-1.598	0.198	-1.528	6935.000	0.127	fixed	
LA (5)	-0.248	0.365	0.950	-0.964	0.468	-0.679	6935.000	0.497	fixed	
LA (6)	-0.423	0.196	0.950	-0.807	-0.038	-2.155	6935.000	0.031	fixed	
LA (7)	0.111	0.313	0.950	-0.503	0.724	0.354	6935.000	0.723	fixed	
LA (8)	-0.225	0.245	0.950	-0.704	0.255	-0.919	6935.000	0.358	fixed	
LA (9)	-0.398	0.315	0.950	-1.015	0.218	-1.266	6935.000	0.205	fixed	
LA (11)	-0.285	0.209	0.950	-0.694	0.124	-1.367	6935.000	0.172	fixed	
LA (12)	-0.492	0.324	0.950	-1.128	0.144	-1.515	6935.000	0.130	fixed	
SD (Intercept)	0.312		0.95						random	School
SD (Observations)	3.059		0.95						random	Residual

Note. SDQ = Strengths and Difficulties Questionnaire, extern = externalising, intern = internalising, LA = local authority.

Appendix C Table 2: Sensitivity Primary Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	t	df error	p	Effects	Group
Intercept	2.672	0.308	0.950	2.068	3.276	8.671	7382.000	0.000	fixed	
Group (Intervention)	0.020	0.123	0.950	-0.222	0.261	0.160	7382.000	0.873	fixed	
SDQ Extern (Baseline)	0.600	0.009	0.950	0.582	0.617	68.226	7382.000	0.000	fixed	
Cohort (Cohort 2)	-0.241	0.268	0.950	-0.766	0.284	-0.899	7382.000	0.369	fixed	
LA (1)	0.093	0.369	0.950	-0.631	0.817	0.252	7382.000	0.801	fixed	
LA (2)	0.327	0.410	0.950	-0.477	1.131	0.798	7382.000	0.425	fixed	
LA (3)	-0.254	0.235	0.950	-0.715	0.207	-1.081	7382.000	0.280	fixed	
LA (4)	-0.418	0.451	0.950	-1.302	0.466	-0.926	7382.000	0.354	fixed	
LA (5)	-0.055	0.361	0.950	-0.763	0.653	-0.153	7382.000	0.878	fixed	
LA (6)	-0.331	0.194	0.950	-0.712	0.049	-1.706	7382.000	0.088	fixed	
LA (7)	0.181	0.308	0.950	-0.423	0.785	0.587	7382.000	0.558	fixed	
LA (8)	-0.027	0.239	0.950	-0.495	0.440	-0.114	7382.000	0.909	fixed	
LA (9)	-0.566	0.311	0.950	-1.176	0.044	-1.817	7382.000	0.069	fixed	
LA (11)	-0.376	0.206	0.950	-0.779	0.027	-1.827	7382.000	0.068	fixed	
LA (12)	-0.305	0.321	0.950	-0.934	0.323	-0.952	7382.000	0.341	fixed	
SD (Intercept)	0.315	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	3.065	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. Extern = externalising, LA = local authority.

Appendix C Table 3: Internalising Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	t	df error	P	Effects	Group
Intercept	2.551	0.272	0.95	2.018	3.084	9.38	6935	0.000	fixed	
Group (Intervention)	0.158	0.105	0.95	-0.048	0.363	1.506	6935	0.132	fixed	
SDQ Intern (Baseline)	0.548	0.01	0.95	0.528	0.568	54.277	6935	0.000	fixed	
Cohort (Cohort 2)	-0.309	0.23	0.95	-0.759	0.141	-1.344	6935	0.179	fixed	
Gender (Female)	1.145	0.077	0.95	0.994	1.296	14.895	6935	0.000	fixed	
Gender (Other)	0.997	0.266	0.95	0.475	1.519	3.744	6935	0.000	fixed	
Ethnicity (Asian)	-0.433	0.102	0.95	-0.633	-0.233	-4.239	6935	0.000	fixed	
Ethnicity (Black)	-0.509	0.152	0.95	-0.807	-0.21	-3.341	6935	0.001	fixed	
Ethnicity (Mixed)	-0.119	0.168	0.95	-0.448	0.211	-0.705	6935	0.481	fixed	
Ethnicity (Other)	-0.385	0.186	0.95	-0.749	-0.02	-2.068	6935	0.039	fixed	
LA (1)	-0.058	0.336	0.95	-0.716	0.601	-0.171	6935	0.864	fixed	
LA (2)	-0.53	0.358	0.95	-1.231	0.171	-1.481	6935	0.139	fixed	
LA (3)	-0.335	0.2	0.95	-0.726	0.057	-1.676	6935	0.094	fixed	
LA (4)	-0.048	0.391	0.95	-0.814	0.718	-0.123	6935	0.902	fixed	
LA (5)	-0.632	0.31	0.95	-1.24	-0.024	-2.039	6935	0.041	fixed	
LA (6)	-0.467	0.164	0.95	-0.789	-0.144	-2.84	6935	0.005	fixed	
LA (7)	-0.181	0.261	0.95	-0.693	0.332	-0.692	6935	0.489	fixed	
LA (8)	-0.244	0.206	0.95	-0.647	0.159	-1.185	6935	0.236	fixed	
LA (9)	-0.749	0.264	0.95	-1.267	-0.232	-2.839	6935	0.005	fixed	
LA (11)	-0.423	0.175	0.95	-0.767	-0.08	-2.414	6935	0.016	fixed	
LA (12)	-0.771	0.274	0.95	-1.307	-0.235	-2.819	6935	0.005	fixed	
SD (Intercept)	0.233	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	2.932	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. Intern = internalising SDQ = Strengths and Difficulties Questionnaire, LA = local authority.

Appendix C Table 4: Impact Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	t	df error	p	Effects	Group
Intercept	0.056	0.119	0.950	-0.176	0.289	0.476	6628.000	0.634	fixed	
Group (Intervention)	0.029	0.044	0.950	-0.058	0.115	0.646	6628.000	0.518	fixed	
Cohort (Cohort 2)	-0.071	0.099	0.950	-0.265	0.122	-0.722	6628.000	0.470	fixed	
SDQ Intern (Baseline)	0.141	0.005	0.950	0.131	0.152	26.965	6628.000	<0.001	fixed	
Gender (Female)	0.326	0.039	0.950	0.249	0.403	8.316	6628.000	<0.001	fixed	
Gender (Other)	0.341	0.137	0.950	0.072	0.611	2.486	6628.000	0.013	fixed	
Ethnicity (Asian)	-0.231	0.050	0.950	-0.330	-0.132	-4.594	6628.000	<0.001	fixed	
Ethnicity (Black)	-0.150	0.078	0.950	-0.303	0.002	-1.930	6628.000	0.054	fixed	
Ethnicity (Mixed)	-0.021	0.087	0.950	-0.191	0.149	-0.239	6628.000	0.811	fixed	
Ethnicity (Other)	-0.089	0.095	0.950	-0.276	0.098	-0.933	6628.000	0.351	fixed	
LA (1)	-0.012	0.156	0.950	-0.317	0.293	-0.076	6628.000	0.939	fixed	
LA (2)	-0.151	0.154	0.950	-0.453	0.150	-0.984	6628.000	0.325	fixed	
LA (3)	-0.024	0.083	0.950	-0.186	0.139	-0.285	6628.000	0.775	fixed	
LA (4)	0.260	0.168	0.950	-0.069	0.589	1.547	6628.000	0.122	fixed	
LA (5)	-0.115	0.132	0.950	-0.374	0.144	-0.871	6628.000	0.384	fixed	
LA (6)	-0.121	0.068	0.950	-0.255	0.013	-1.766	6628.000	0.077	fixed	
LA (7)	-0.053	0.108	0.950	-0.266	0.160	-0.488	6628.000	0.625	fixed	
LA (8)	0.066	0.086	0.950	-0.102	0.234	0.770	6628.000	0.441	fixed	
LA (9)	-0.160	0.111	0.950	-0.377	0.057	-1.444	6628.000	0.149	fixed	
LA (11)	-0.113	0.073	0.950	-0.256	0.030	-1.544	6628.000	0.123	fixed	
LA (12)	-0.229	0.115	0.950	-0.455	-0.003	-1.983	6628.000	0.047	fixed	
SD (Intercept)	0.062	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	1.488	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. SDQ = Strengths and Difficulties Questionnaire, intern = internalising, LA = local authority.

Appendix C Table 5: Prosocial Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	T	df error	p	Effects	Group
Intercept	3.368	0.231	0.950	2.915	3.821	14.589	6936.000	0.000	Fixed	
Group (Intervention)	-0.015	0.089	0.950	-0.189	0.159	-0.168	6936.000	0.866	Fixed	
SDQ Prosocial (Baseline)	0.426	0.011	0.950	0.406	0.447	40.059	6936.000	0.000	Fixed	
Cohort (Cohort 2)	0.179	0.190	0.950	-0.193	0.551	0.943	6936.000	0.346	Fixed	

SDQ Intern (Baseline)	-0.021	0.006	0.950	-0.034	-0.009	-3.423	6936.000	0.001	Fixed	
Gender (Female)	0.735	0.049	0.950	0.639	0.831	14.981	6936.000	0.000	Fixed	
Gender (Other)	0.128	0.164	0.950	-0.194	0.449	0.778	6936.000	0.436	fixed	
Ethnicity (Asian)	0.161	0.066	0.950	0.031	0.291	2.431	6936.000	0.015	fixed	
Ethnicity (Black)	0.029	0.096	0.950	-0.158	0.217	0.306	6936.000	0.759	fixed	
Ethnicity (Mixed)	0.076	0.104	0.950	-0.129	0.280	0.727	6936.000	0.467	fixed	
Ethnicity (Other)	0.044	0.116	0.950	-0.183	0.272	0.380	6936.000	0.704	fixed	
LA (1)	0.064	0.264	0.950	-0.453	0.582	0.243	6936.000	0.808	fixed	
LA (2)	0.304	0.292	0.950	-0.268	0.876	1.042	6936.000	0.298	fixed	
LA (3)	0.112	0.173	0.950	-0.228	0.451	0.645	6936.000	0.519	fixed	
LA (4)	0.711	0.325	0.950	0.073	1.349	2.185	6936.000	0.029	fixed	
LA (5)	0.228	0.260	0.950	-0.281	0.737	0.879	6936.000	0.379	fixed	
LA (6)	0.334	0.142	0.950	0.055	0.612	2.351	6936.000	0.019	fixed	
LA (7)	-0.066	0.227	0.950	-0.512	0.380	-0.291	6936.000	0.771	fixed	
LA (8)	0.350	0.176	0.950	0.004	0.696	1.985	6936.000	0.047	fixed	
LA (9)	0.407	0.228	0.950	-0.039	0.854	1.788	6936.000	0.074	fixed	
LA (11)	0.304	0.150	0.950	0.010	0.597	2.029	6936.000	0.042	fixed	
LA (12)	0.520	0.233	0.950	0.063	0.976	2.232	6936.000	0.026	fixed	
SD (Intercept)	0.251	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	1.808	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. SDQ = Strengths and Difficulties Questionnaire, intern = internalising, LA = local authority.

Appendix C Table 6: SEI TSR Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	T	df error	P	Effects	Group
Intercept	10.816	0.670	0.950	9.504	12.129	16.155	5925.000	0.000	fixed	
Group (Intervention)	-0.091	0.244	0.950	-0.569	0.388	-0.373	5925.000	0.709	fixed	
SEI TSR (Baseline)	0.476	0.011	0.950	0.454	0.499	41.652	5925.000	0.000	fixed	
Cohort (Cohort 2)	1.327	0.516	0.950	0.315	2.339	2.570	5925.000	0.010	fixed	
SDQ Intern (Baseline)	-0.013	0.017	0.950	-0.045	0.020	-0.761	5925.000	0.446	fixed	
Gender (Female)	0.001	0.128	0.950	-0.251	0.252	0.005	5925.000	0.996	fixed	
Gender (Other)	-0.110	0.448	0.950	-0.988	0.769	-0.245	5925.000	0.807	fixed	
Ethnicity (Asian)	0.005	0.176	0.950	-0.339	0.349	0.028	5925.000	0.977	fixed	
Ethnicity (Black)	0.021	0.261	0.950	-0.491	0.534	0.082	5925.000	0.935	fixed	
Ethnicity (Mixed)	0.100	0.276	0.950	-0.441	0.642	0.363	5925.000	0.716	fixed	
Ethnicity (Other)	-0.285	0.316	0.950	-0.905	0.335	-0.901	5925.000	0.368	fixed	
LA (1)	1.003	0.709	0.950	-0.387	2.393	1.415	5925.000	0.157	fixed	
LA (2)	0.822	0.791	0.950	-0.730	2.373	1.039	5925.000	0.299	fixed	
LA (3)	0.723	0.475	0.950	-0.208	1.655	1.522	5925.000	0.128	fixed	
LA (4)	0.437	0.883	0.950	-1.293	2.168	0.495	5925.000	0.620	fixed	
LA (5)	1.182	0.709	0.950	-0.209	2.572	1.666	5925.000	0.096	fixed	
LA (6)	2.106	0.390	0.950	1.342	2.870	5.405	5925.000	0.000	fixed	
LA (7)	0.097	0.622	0.950	-1.122	1.316	0.156	5925.000	0.876	fixed	
LA (8)	1.142	0.482	0.950	0.198	2.086	2.371	5925.000	0.018	fixed	
LA (9)	0.871	0.627	0.950	-0.357	2.099	1.390	5925.000	0.165	fixed	
LA (11)	-0.127	0.411	0.950	-0.933	0.678	-0.310	5925.000	0.757	fixed	
LA (12)	1.347	0.638	0.950	0.098	2.597	2.113	5925.000	0.035	fixed	
SD (Intercept)	0.695	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	4.421	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. SEI = Student Engagement Instrument, TSR = teacher-student relationships, SDQ = Strengths and Difficulties Questionnaire, intern = internalising, LA = local authority.

Appendix C Table 7: Suspensions Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	z	df error	P	Effects	Group
Intercept	-1.325	0.528	0.950	-2.360	-0.291	-2.512	Inf	0.012	fixed	
Group (Intervention)	0.172	0.195	0.950	-0.211	0.555	0.879	Inf	0.379	fixed	
Suspensions (Baseline)	2.953	0.140	0.950	2.678	3.229	21.036	Inf	0.000	fixed	
Cohort (Cohort 2)	-1.153	0.471	0.950	-2.077	-0.229	-2.446	Inf	0.014	fixed	
SDQ Intern (Baseline)	0.017	0.015	0.950	-0.013	0.048	1.122	Inf	0.262	fixed	
Gender (Female)	-0.441	0.125	0.950	-0.685	-0.196	-3.534	Inf	0.000	fixed	
Gender (Other)	0.619	0.305	0.950	0.022	1.216	2.033	Inf	0.042	fixed	
Ethnicity (Asian)	-0.684	0.170	0.950	-1.016	-0.351	-4.031	Inf	0.000	fixed	
Ethnicity (Black)	-0.671	0.273	0.950	-1.205	-0.136	-2.461	Inf	0.014	fixed	
Ethnicity (Mixed)	-0.046	0.237	0.950	-0.511	0.419	-0.196	Inf	0.845	fixed	

Ethnicity (Other)	-0.248	0.260	0.950	-0.758	0.262	-0.954	Inf	0.340	fixed	
LA (3)	-0.775	0.371	0.950	-1.503	-0.048	-2.089	Inf	0.037	fixed	
LA (4)	-0.355	0.535	0.950	-1.403	0.693	-0.664	Inf	0.507	fixed	
LA (5)	-1.625	0.564	0.950	-2.732	-0.519	-2.880	Inf	0.004	fixed	
LA (6)	-1.121	0.299	0.950	-1.707	-0.534	-3.746	Inf	0.000	fixed	
LA (7)	-1.217	0.456	0.950	-2.111	-0.322	-2.665	Inf	0.008	fixed	
LA (8)	-1.906	0.702	0.950	-3.281	-0.530	-2.715	Inf	0.007	fixed	
LA (9)	-2.183	0.803	0.950	-3.757	-0.608	-2.717	Inf	0.007	fixed	
LA (11)	-0.472	0.295	0.950	-1.051	0.108	-1.596	Inf	0.110	fixed	
LA (12)	-1.976	0.640	0.950	-3.230	-0.722	-3.088	Inf	0.002	fixed	
SD (Intercept)	0.407	NA	0.950	NA	NA	NA	NA	NA	random	School

Note. SDQ = Strengths and Difficulties Questionnaire, intern = internalising, LA = local authority.

Appendix C Table 8: SEI PSAS Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	T	df error	P	Effects	Group
Intercept	11.445	0.42	0.95	10.621	12.268	27.235	6004	0.00	fixed	
Group (Intervention)	0.066	0.13	0.95	-0.189	0.321	0.506	6004	0.613	fixed	
SEI PSAS (Baseline)	0.352	0.013	0.95	0.327	0.378	27.142	6004	0.00	fixed	
Cohort (Cohort 2)	0.523	0.283	0.95	-0.031	1.077	1.85	6004	0.064	fixed	
SDQ Intern (Baseline)	-0.121	0.013	0.95	-0.147	-0.095	-9.146	6004	0.00	fixed	
Gender (Female)	-0.016	0.09	0.95	-0.193	0.161	-0.178	6004	0.859	fixed	
Gender (Other)	-0.261	0.314	0.95	-0.877	0.354	-0.832	6004	0.405	fixed	
Ethnicity (Asian)	0.16	0.12	0.95	-0.076	0.396	1.331	6004	0.183	fixed	
Ethnicity (Black)	0.118	0.181	0.95	-0.237	0.473	0.65	6004	0.516	fixed	
Ethnicity (Mixed)	-0.203	0.196	0.95	-0.586	0.181	-1.035	6004	0.301	fixed	
Ethnicity (Other)	0.009	0.22	0.95	-0.421	0.44	0.043	6004	0.966	fixed	
LA (1)	0.096	0.405	0.95	-0.699	0.891	0.237	6004	0.812	fixed	
LA (2)	0.714	0.435	0.95	-0.139	1.566	1.641	6004	0.101	fixed	
LA (3)	0.613	0.248	0.95	0.127	1.099	2.474	6004	0.013	fixed	
LA (4)	0.274	0.47	0.95	-0.647	1.195	0.583	6004	0.56	fixed	
LA (5)	0.998	0.384	0.95	0.246	1.749	2.601	6004	0.009	fixed	
LA (6)	0.844	0.205	0.95	0.441	1.246	4.111	6004	0	fixed	
LA (7)	0.075	0.324	0.95	-0.559	0.709	0.231	6004	0.817	fixed	
LA (8)	0.58	0.254	0.95	0.083	1.078	2.288	6004	0.022	fixed	
LA (9)	0.456	0.329	0.95	-0.188	1.1	1.388	6004	0.165	fixed	
LA (11)	0.518	0.217	0.95	0.091	0.944	2.38	6004	0.017	fixed	
LA (12)	1.119	0.339	0.95	0.454	1.784	3.301	6004	0.001	fixed	
SD (Intercept)	0.304	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	3.187	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. SEI = Student Engagement Instrument. PSAS = peer support at school, SDQ = Strengths and Difficulties Questionnaire, intern = internalising, LA = local authority.

Appendix C Table 9: SEI CRS Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	t	df error	p	Effects	Group
Intercept	11.401	0.639	0.95	10.149	12.653	17.848	5893	0.000	fixed	
Group (Intervention)	-0.066	0.229	0.95	-0.514	0.383	-0.286	5893	0.775	fixed	
SEI CRS (Baseline)	0.455	0.012	0.95	0.432	0.478	38.985	5893	0.000	fixed	
Cohort (Cohort 2)	1.082	0.484	0.95	0.133	2.03	2.235	5893	0.025	fixed	
SDQ Intern (Baseline)	-0.044	0.016	0.95	-0.077	-0.012	-2.711	5893	0.007	fixed	
Gender (Female)	0.425	0.126	0.95	0.179	0.672	3.385	5893	0.001	fixed	
Gender (Other)	0.444	0.429	0.95	-0.397	1.285	1.035	5893	0.301	fixed	
Ethnicity (Asian)	0.731	0.172	0.95	0.393	1.068	4.24	5893	0.000	fixed	
Ethnicity (Black)	1.199	0.258	0.95	0.693	1.705	4.646	5893	0.00	fixed	
Ethnicity (Mixed)	0.572	0.267	0.95	0.048	1.096	2.138	5893	0.033	fixed	
Ethnicity (Other)	0.507	0.306	0.95	-0.094	1.107	1.653	5893	0.098	fixed	
LA (1)	0.597	0.675	0.95	-0.728	1.921	0.883	5893	0.377	fixed	
LA (2)	0.194	0.75	0.95	-1.276	1.663	0.258	5893	0.796	fixed	
LA (3)	-0.065	0.444	0.95	-0.934	0.804	-0.146	5893	0.884	fixed	
LA (4)	0.408	0.824	0.95	-1.207	2.022	0.495	5893	0.621	fixed	
LA (5)	0.774	0.665	0.95	-0.529	2.078	1.165	5893	0.244	fixed	
LA (6)	1.277	0.364	0.95	0.563	1.991	3.506	5893	0.000	fixed	
LA (7)	0.101	0.583	0.95	-1.042	1.243	0.173	5893	0.863	fixed	

LA (8)	0.48	0.451	0.95	-0.403	1.364	1.066	5893	0.286	fixed	
LA (9)	0.199	0.583	0.95	-0.944	1.342	0.341	5893	0.733	fixed	
LA (11)	0.010	0.386	0.95	-0.747	0.767	0.026	5893	0.98	fixed	
LA (12)	1.193	0.596	0.95	0.025	2.361	2.001	5893	0.045	fixed	
SD (Intercept)	0.637	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	4.339	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. SEI = Student Engagement Instrument. CRS = control and relevance of schoolwork, SDQ = Strengths and Difficulties Questionnaire, intern = internalising, LA = local authority.

Appendix C Table 10: SEI FG Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	t	df error	p	Effects	Group
Intercept	8.346	0.332	0.950	7.695	8.997	25.120	6043.000	0.000	fixed	
Group (Intervention)	0.162	0.103	0.950	-0.041	0.364	1.567	6043.000	0.117	fixed	
SEI FG (Baseline)	0.411	0.012	0.950	0.388	0.435	34.360	6043.000	0.000	fixed	
Cohort (Cohort 2)	0.479	0.227	0.950	0.035	0.924	2.115	6043.000	0.035	fixed	
SDQ Intern (Baseline)	-0.029	0.011	0.950	-0.050	-0.008	-2.750	6043.000	0.006	fixed	
Gender (Female)	0.552	0.080	0.950	0.396	0.709	6.923	6043.000	0.000	fixed	
Gender (Other)	-0.155	0.280	0.950	-0.705	0.394	-0.555	6043.000	0.579	fixed	
Ethnicity (Asian)	0.398	0.105	0.950	0.193	0.603	3.805	6043.000	0.000	fixed	
Ethnicity (Black)	0.371	0.160	0.950	0.056	0.686	2.312	6043.000	0.021	fixed	
Ethnicity (Mixed)	0.154	0.172	0.950	-0.184	0.491	0.891	6043.000	0.373	fixed	
Ethnicity (Other)	0.332	0.195	0.950	-0.049	0.713	1.707	6043.000	0.088	fixed	
LA (1)	0.543	0.327	0.950	-0.098	1.184	1.661	6043.000	0.097	fixed	
LA (2)	0.321	0.346	0.950	-0.357	0.999	0.927	6043.000	0.354	fixed	
LA (3)	0.284	0.195	0.950	-0.098	0.665	1.458	6043.000	0.145	fixed	
LA (4)	0.252	0.374	0.950	-0.482	0.986	0.673	6043.000	0.501	fixed	
LA (5)	0.328	0.308	0.950	-0.275	0.931	1.068	6043.000	0.286	fixed	
LA (6)	0.692	0.162	0.950	0.375	1.009	4.278	6043.000	0.000	fixed	
LA (7)	0.291	0.254	0.950	-0.207	0.790	1.146	6043.000	0.252	fixed	
LA (8)	0.373	0.200	0.950	-0.020	0.766	1.861	6043.000	0.063	fixed	
LA (9)	0.399	0.259	0.950	-0.109	0.906	1.540	6043.000	0.124	fixed	
LA (11)	0.198	0.172	0.950	-0.140	0.536	1.150	6043.000	0.250	fixed	
LA (12)	0.808	0.269	0.950	0.281	1.335	3.006	6043.000	0.003	fixed	
SD (Intercept)	0.209	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	2.846	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. SEI = Student Engagement Instrument. FG = future goals , SDQ = Strengths and Difficulties Questionnaire, intern = internalising, LA = local authority.

Appendix C Table 11: IBS Perp Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	t	df error	p	Effects	Group
Intercept	2.742	0.657	0.950	1.453	4.030	4.171	5803.000	0.000	fixed	
Group (Intervention)	0.054	0.261	0.950	-0.457	0.566	0.207	5803.000	0.836	fixed	
IBS Perp (Baseline)	0.434	0.013	0.950	0.409	0.459	34.225	5803.000	0.000	fixed	
Cohort (Cohort 2)	0.097	0.561	0.950	-1.002	1.197	0.174	5803.000	0.862	fixed	
SDQ Intern (Baseline)	0.021	0.020	0.950	-0.017	0.059	1.082	5803.000	0.280	fixed	
Gender (Female)	-1.129	0.150	0.950	-1.423	-0.835	-7.523	5803.000	0.000	fixed	
Gender (Other)	-0.250	0.513	0.950	-1.256	0.755	-0.488	5803.000	0.625	fixed	
Ethnicity (Asian)	-0.206	0.202	0.950	-0.601	0.189	-1.022	5803.000	0.307	fixed	
Ethnicity (Black)	0.461	0.303	0.950	-0.133	1.055	1.522	5803.000	0.128	fixed	
Ethnicity (Mixed)	0.380	0.321	0.950	-0.249	1.010	1.184	5803.000	0.236	fixed	
Ethnicity (Other)	0.289	0.372	0.950	-0.441	1.019	0.776	5803.000	0.437	fixed	
LA (1)	-0.225	0.769	0.950	-1.733	1.283	-0.293	5803.000	0.770	fixed	
LA (2)	-0.050	0.853	0.950	-1.723	1.622	-0.059	5803.000	0.953	fixed	
LA (3)	0.264	0.503	0.950	-0.722	1.250	0.524	5803.000	0.600	fixed	
LA (4)	-0.286	0.934	0.950	-2.117	1.544	-0.307	5803.000	0.759	fixed	
LA (5)	-0.672	0.766	0.950	-2.173	0.829	-0.878	5803.000	0.380	fixed	
LA (6)	-0.957	0.417	0.950	-1.774	-0.140	-2.297	5803.000	0.022	fixed	
LA (7)	0.356	0.661	0.950	-0.939	1.652	0.539	5803.000	0.590	fixed	
LA (8)	-0.355	0.512	0.950	-1.358	0.649	-0.693	5803.000	0.488	fixed	
LA (9)	-0.883	0.662	0.950	-2.181	0.415	-1.334	5803.000	0.182	fixed	
LA (11)	-0.461	0.436	0.950	-1.316	0.395	-1.056	5803.000	0.291	fixed	
LA (12)	-0.404	0.683	0.950	-1.744	0.936	-0.591	5803.000	0.554	fixed	
SD (Intercept)	0.707	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	5.123	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. IBS = Illinois Bully Scale, perp = perpetration. SDQ = Strengths and Difficulties Questionnaire, intern = internalising, LA = Local Authority.

Appendix C Table 12: IBS Victim Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	T	df error	p	Effects	Group
Intercept	1.905	0.385	0.950	1.150	2.659	4.950	6039.000	0.000	fixed	
Group (Intervention)	0.363	0.149	0.950	0.071	0.655	2.437	6039.000	0.015	fixed	
IBS Victim (Baseline)	0.395	0.013	0.950	0.370	0.420	30.796	6039.000	0.000	fixed	
Cohort (Cohort 2)	-0.583	0.328	0.950	-1.225	0.059	-1.780	6039.000	0.075	fixed	
SDQ Intern (Baseline)	0.139	0.016	0.950	0.108	0.169	8.916	6039.000	0.000	fixed	
Gender (Female)	-0.272	0.104	0.950	-0.477	-0.068	-2.613	6039.000	0.009	fixed	
Gender (Other)	-0.080	0.353	0.950	-0.772	0.612	-0.227	6039.000	0.820	fixed	
Ethnicity (Asian)	-0.335	0.137	0.950	-0.604	-0.067	-2.450	6039.000	0.014	fixed	
Ethnicity (Black)	-0.066	0.208	0.950	-0.473	0.341	-0.318	6039.000	0.751	fixed	
Ethnicity (Mixed)	0.059	0.226	0.950	-0.383	0.502	0.263	6039.000	0.792	fixed	
Ethnicity (Other)	-0.117	0.250	0.950	-0.607	0.374	-0.466	6039.000	0.641	fixed	
LA (1)	0.654	0.457	0.950	-0.243	1.550	1.430	6039.000	0.153	fixed	
LA (2)	-0.384	0.491	0.950	-1.347	0.579	-0.781	6039.000	0.435	fixed	
LA (3)	-0.304	0.283	0.950	-0.858	0.250	-1.076	6039.000	0.282	fixed	
LA (4)	0.524	0.534	0.950	-0.522	1.571	0.982	6039.000	0.326	fixed	
LA (5)	-0.935	0.442	0.950	-1.801	-0.069	-2.117	6039.000	0.034	fixed	
LA (6)	-0.878	0.236	0.950	-1.340	-0.416	-3.724	6039.000	0.000	fixed	
LA (7)	0.127	0.371	0.950	-0.600	0.854	0.342	6039.000	0.732	fixed	
LA (8)	-0.351	0.290	0.950	-0.919	0.217	-1.211	6039.000	0.226	fixed	
LA (9)	-0.949	0.373	0.950	-1.681	-0.218	-2.543	6039.000	0.011	fixed	
LA (11)	-0.493	0.248	0.950	-0.979	-0.007	-1.987	6039.000	0.047	fixed	
LA (12)	-0.647	0.390	0.950	-1.411	0.117	-1.660	6039.000	0.097	fixed	
SD (Intercept)	0.348	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	3.646	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. IBS = Illinois Bully Scale, victim = victimisation. SDQ = Strengths and Difficulties Questionnaire, intern = internalising, LA = local authority.

Appendix C Table 13: IBS Victim Model Without Covariates Summary.

Parameter	Estimate	SE	CI	CI low	CI high	T	df error	p	Effects	Group
Intercept	2.256	0.373	0.950	1.524	2.988	6.045	6418.000	0.000	fixed	
Group (Intervention)	0.327	0.150	0.950	0.033	0.621	2.179	6418.000	0.029	fixed	
IBS Victim (Baseline)	0.452	0.011	0.950	0.430	0.473	41.800	6418.000	0.000	fixed	
Cohort (Cohort 2)	-0.597	0.332	0.950	-1.247	0.053	-1.801	6418.000	0.072	fixed	
LA (1)	0.826	0.446	0.950	-0.048	1.700	1.852	6418.000	0.064	fixed	
LA (2)	-0.101	0.491	0.950	-1.064	0.862	-0.205	6418.000	0.838	fixed	
LA (3)	-0.426	0.283	0.950	-0.982	0.130	-1.503	6418.000	0.133	fixed	
LA (4)	0.716	0.534	0.950	-0.332	1.764	1.339	6418.000	0.181	fixed	
LA (5)	-0.719	0.445	0.950	-1.591	0.152	-1.618	6418.000	0.106	fixed	
LA (6)	-0.868	0.238	0.950	-1.334	-0.401	-3.646	6418.000	0.000	fixed	
LA (7)	0.307	0.373	0.950	-0.423	1.038	0.825	6418.000	0.410	fixed	
LA (8)	-0.195	0.288	0.950	-0.759	0.370	-0.676	6418.000	0.499	fixed	
LA (9)	-1.104	0.375	0.950	-1.838	-0.369	-2.946	6418.000	0.003	fixed	
LA (11)	-0.615	0.248	0.950	-1.102	-0.128	-2.475	6418.000	0.013	fixed	
LA (12)	-0.573	0.393	0.950	-1.343	0.196	-1.460	6418.000	0.144	fixed	
SD (Intercept)	0.362	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	3.688	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. IBS = Illinois Bully Scale, SEN = special educational needs. LA = local authority.

Appendix C Table 14: Attendance Percentage Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	T	df error	p	Effects	Group
Intercept	1.026	1.495	0.950	-1.904	3.956	0.687	7356.000	0.492	fixed	
Group (Intervention)	-0.427	0.309	0.950	-1.032	0.178	-1.382	7356.000	0.167	fixed	
Attend perc (Baseline)	0.987	0.013	0.950	0.961	1.012	77.016	7356.000	0.000	fixed	
Cohort (Cohort 2)	-0.395	0.818	0.950	-1.998	1.208	-0.483	7356.000	0.629	fixed	
SDQ Intern (Baseline)	-0.023	0.023	0.950	-0.068	0.021	-1.030	7356.000	0.303	fixed	

Gender (Female)	-0.566	0.177	0.950	-0.914	-0.219	-3.194	7356.000	0.001	fixed	
Gender (Other)	-1.235	0.589	0.950	-2.390	-0.080	-2.096	7356.000	0.036	fixed	
Ethnicity (Asian)	0.544	0.234	0.950	0.085	1.004	2.324	7356.000	0.020	fixed	
Ethnicity (Black)	0.156	0.350	0.950	-0.530	0.842	0.446	7356.000	0.656	fixed	
Ethnicity (Mixed)	0.613	0.372	0.950	-0.116	1.342	1.647	7356.000	0.100	fixed	
Ethnicity (Other)	0.113	0.410	0.950	-0.691	0.918	0.276	7356.000	0.783	fixed	
LA (3)	0.312	0.567	0.950	-0.799	1.422	0.550	7356.000	0.582	fixed	
LA (4)	-1.752	0.987	0.950	-3.688	0.183	-1.775	7356.000	0.076	fixed	
LA (5)	-0.460	0.991	0.950	-2.402	1.482	-0.465	7356.000	0.642	fixed	
LA (6)	1.007	0.476	0.950	0.073	1.941	2.114	7356.000	0.035	fixed	
LA (7)	0.991	0.725	0.950	-0.429	2.412	1.368	7356.000	0.171	fixed	
LA (8)	0.907	0.958	0.950	-0.972	2.786	0.947	7356.000	0.344	fixed	
LA (9)	0.521	0.731	0.950	-0.912	1.954	0.712	7356.000	0.476	fixed	
LA (11)	0.276	0.481	0.950	-0.668	1.219	0.573	7356.000	0.567	fixed	
LA (12)	0.926	1.088	0.950	-1.207	3.060	0.851	7356.000	0.395	fixed	
SD (Intercept)	0.752	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	6.738	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. SDQ = Strengths and Difficulties Questionnaire, intern = internalising, LA = local authority.

Appendix C Table 15: ARTIC Underlying Causes Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	T	df error	p	Effects	Group
Intercept	1.566	0.191	0.950	1.191	1.941	8.198	646.000	0.000	fixed	
Group (Intervention)	0.094	0.055	0.950	-0.014	0.202	1.702	646.000	0.089	fixed	
Underlying Causes (Baseline)	0.690	0.030	0.950	0.632	0.748	23.249	646.000	0.000	fixed	
Cohort (Cohort 2)	-0.101	0.110	0.950	-0.318	0.116	-0.913	646.000	0.361	fixed	
LA (1)	0.028	0.165	0.950	-0.296	0.352	0.170	646.000	0.865	fixed	
LA (2)	-0.110	0.207	0.950	-0.517	0.297	-0.530	646.000	0.596	fixed	
LA (3)	-0.115	0.084	0.950	-0.281	0.050	-1.367	646.000	0.172	fixed	
LA (4)	0.147	0.259	0.950	-0.361	0.655	0.567	646.000	0.571	fixed	
LA (5)	-0.190	0.144	0.950	-0.472	0.092	-1.321	646.000	0.187	fixed	
LA (6)	0.026	0.077	0.950	-0.125	0.178	0.343	646.000	0.732	fixed	
LA (7)	0.072	0.150	0.950	-0.223	0.367	0.480	646.000	0.631	fixed	
LA (8)	-0.057	0.100	0.950	-0.255	0.140	-0.570	646.000	0.569	fixed	
LA (9)	0.007	0.160	0.950	-0.307	0.320	0.041	646.000	0.967	fixed	
LA (11)	-0.208	0.075	0.950	-0.356	-0.061	-2.776	646.000	0.006	fixed	
LA (12)	-0.320	0.148	0.950	-0.610	-0.029	-2.163	646.000	0.031	fixed	
SD (Intercept)	0.038	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	0.556	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. ARTIC = Attitudes Related to Trauma-Informed Care, LA = local authority.

Appendix C Table 16: ARTIC Self Efficacy Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	T	df error	p	Effects	Group
Intercept	2.289	0.236	0.950	1.825	2.753	9.686	644.000	0.000	fixed	
Group (Intervention)	0.119	0.078	0.950	-0.033	0.271	1.533	644.000	0.126	fixed	
Self Efficacy (Baseline)	0.587	0.030	0.950	0.527	0.647	19.278	644.000	0.000	fixed	
Cohort (Cohort 2)	-0.043	0.157	0.950	-0.351	0.266	-0.272	644.000	0.786	fixed	
LA (1)	-0.005	0.223	0.950	-0.443	0.432	-0.024	644.000	0.981	fixed	
LA (2)	-0.086	0.282	0.950	-0.641	0.468	-0.306	644.000	0.759	fixed	
LA (3)	0.157	0.131	0.950	-0.100	0.414	1.202	644.000	0.230	fixed	
LA (4)	0.190	0.377	0.950	-0.551	0.930	0.504	644.000	0.615	fixed	
LA (5)	-0.152	0.206	0.950	-0.556	0.253	-0.736	644.000	0.462	fixed	
LA (6)	0.146	0.112	0.950	-0.074	0.366	1.300	644.000	0.194	fixed	
LA (7)	0.030	0.210	0.950	-0.383	0.442	0.141	644.000	0.888	fixed	
LA (8)	-0.227	0.148	0.950	-0.518	0.063	-1.535	644.000	0.125	fixed	
LA (9)	-0.201	0.217	0.950	-0.627	0.225	-0.927	644.000	0.355	fixed	
LA (11)	-0.096	0.112	0.950	-0.316	0.125	-0.852	644.000	0.394	fixed	
LA (12)	-0.354	0.209	0.950	-0.764	0.056	-1.694	644.000	0.091	fixed	
SD (Intercept)	0.109	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	0.701	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. ARTIC = Attitudes Related to Trauma-Informed Care, LA = local authority.

Appendix C Table 17: ARTIC Responses Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	T	df error	p	Effects	Group
Intercept	1.881	0.228	0.950	1.433	2.330	8.239	644.000	0.000	Fixed	
Group (Intervention)	0.109	0.072	0.950	-0.032	0.250	1.523	644.000	0.128	Fixed	
Responses (Baseline)	0.663	0.032	0.950	0.600	0.726	20.775	644.000	0.000	Fixed	
Cohort (Cohort 2)	-0.180	0.144	0.950	-0.462	0.103	-1.250	644.000	0.212	Fixed	
LA (1)	-0.005	0.208	0.950	-0.413	0.403	-0.024	644.000	0.981	Fixed	
LA (2)	-0.208	0.262	0.950	-0.722	0.307	-0.792	644.000	0.429	Fixed	
LA (3)	-0.214	0.118	0.950	-0.446	0.018	-1.808	644.000	0.071	Fixed	
LA (4)	0.139	0.321	0.950	-0.492	0.770	0.431	644.000	0.666	Fixed	
LA (5)	-0.121	0.189	0.950	-0.491	0.250	-0.639	644.000	0.523	Fixed	
LA (6)	0.016	0.103	0.950	-0.186	0.217	0.151	644.000	0.880	Fixed	
LA (7)	0.052	0.191	0.950	-0.322	0.426	0.272	644.000	0.786	fixed	
LA (8)	0.023	0.135	0.950	-0.242	0.288	0.172	644.000	0.864	fixed	
LA (9)	0.036	0.202	0.950	-0.361	0.433	0.177	644.000	0.859	fixed	
LA (11)	-0.199	0.102	0.950	-0.400	0.002	-1.946	644.000	0.052	fixed	
LA (12)	-0.288	0.192	0.950	-0.665	0.089	-1.501	644.000	0.134	fixed	
SD (Intercept)	0.089	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	0.667	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. ARTIC = Attitudes Related to Trauma-Informed Care, LA = local authority.

Appendix C Table 18: ARTIC Reactions Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	T	df error	p	Effects	Group
Intercept	2.151	0.243	0.950	1.675	2.628	8.866	643.000	0.000	Fixed	
Group (Intervention)	0.043	0.075	0.950	-0.104	0.190	0.572	643.000	0.567	Fixed	
Reactions (Baseline)	0.607	0.034	0.950	0.540	0.674	17.734	643.000	0.000	Fixed	
Cohort (Cohort 2)	-0.028	0.152	0.950	-0.326	0.271	-0.181	643.000	0.857	Fixed	
LA (1)	-0.204	0.218	0.950	-0.632	0.225	-0.934	643.000	0.351	Fixed	
LA (2)	0.052	0.275	0.950	-0.489	0.593	0.188	643.000	0.851	Fixed	
LA (3)	-0.081	0.124	0.950	-0.325	0.163	-0.651	643.000	0.515	Fixed	
LA (4)	-0.069	0.338	0.950	-0.732	0.594	-0.205	643.000	0.838	Fixed	
LA (5)	-0.051	0.199	0.950	-0.442	0.339	-0.258	643.000	0.797	Fixed	
LA (6)	-0.083	0.108	0.950	-0.295	0.128	-0.773	643.000	0.440	Fixed	
LA (7)	-0.096	0.205	0.950	-0.499	0.307	-0.466	643.000	0.641	Fixed	
LA (8)	-0.108	0.142	0.950	-0.387	0.171	-0.758	643.000	0.449	Fixed	
LA (9)	-0.069	0.212	0.950	-0.486	0.347	-0.326	643.000	0.744	Fixed	
LA (11)	-0.335	0.107	0.950	-0.546	-0.125	-3.126	643.000	0.002	Fixed	
LA (12)	-0.261	0.203	0.950	-0.659	0.137	-1.288	643.000	0.198	Fixed	
SD (Intercept)	0.094	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	0.701	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. ARTIC = Attitudes Related to Trauma-Informed Care, LA = local authority.

Appendix C Table 19: ARTIC on the Job Behaviour Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	T	df error	p	Effects	Group
Intercept	1.994	0.223	0.950	1.557	2.431	8.957	645.000	0.000	Fixed	
Group (Intervention)	0.090	0.064	0.950	-0.035	0.216	1.411	645.000	0.159	Fixed	
On Job Behaviour (Baseline)	0.679	0.030	0.950	0.619	0.738	22.398	645.000	0.000	Fixed	
Cohort (Cohort 2)	-0.207	0.131	0.950	-0.463	0.049	-1.585	645.000	0.113	Fixed	
LA (1)	-0.041	0.182	0.950	-0.398	0.317	-0.223	645.000	0.824	Fixed	
LA (2)	-0.166	0.231	0.950	-0.620	0.287	-0.720	645.000	0.472	Fixed	
LA (3)	-0.071	0.111	0.950	-0.288	0.146	-0.640	645.000	0.522	Fixed	
LA (4)	0.351	0.279	0.950	-0.196	0.898	1.259	645.000	0.208	Fixed	
LA (5)	-0.257	0.171	0.950	-0.594	0.079	-1.500	645.000	0.134	Fixed	
LA (6)	-0.068	0.094	0.950	-0.252	0.116	-0.729	645.000	0.466	Fixed	
LA (7)	0.056	0.172	0.950	-0.282	0.393	0.324	645.000	0.746	Fixed	
LA (8)	0.117	0.124	0.950	-0.125	0.360	0.949	645.000	0.343	Fixed	
LA (9)	-0.017	0.177	0.950	-0.366	0.331	-0.097	645.000	0.922	Fixed	
LA (11)	-0.160	0.095	0.950	-0.345	0.026	-1.689	645.000	0.092	Fixed	
LA (12)	-0.349	0.173	0.950	-0.688	-0.010	-2.019	645.000	0.044	Fixed	
SD (Intercept)	0.099	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	0.560	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. ARTIC = Attitudes Related to Trauma-Informed Care, LA = local authority.

Appendix C Table 20: Model Output for the CACE Model.

Parameter	Estimate	SE	CI	CI low	CI high	t	df error	p	Effects	Group
(Intercept)	2.354	0.419	0.950	1.532	3.175	5.617	6935.000	0.000	fixed	
Compliance	0.489	0.413	0.950	-0.320	1.299	1.185	6935.000	0.236	fixed	
Cohort (Cohort 2)	-0.291	0.232	0.950	-0.746	0.163	-1.256	6935.000	0.209	fixed	
SDQ Extern (baseline)	0.586	0.010	0.950	0.567	0.606	59.723	6935.000	0.000	fixed	
SDQ Intern (baseline)	0.030	0.011	0.950	0.008	0.052	2.647	6935.000	0.008	fixed	
Gender (Male)	0.021	0.080	0.950	-0.136	0.178	0.266	6935.000	0.790	fixed	
Gender (Other)	-0.117	0.277	0.950	-0.659	0.426	-0.421	6935.000	0.674	fixed	
Ethnicity (Black)	-0.030	0.157	0.950	-0.338	0.278	-0.192	6935.000	0.848	fixed	
Ethnicity (Mixed)	0.171	0.182	0.950	-0.185	0.527	0.942	6935.000	0.346	fixed	
Ethnicity (Other)	0.141	0.193	0.950	-0.236	0.519	0.734	6935.000	0.463	fixed	
Ethnicity (White)	0.415	0.107	0.950	0.205	0.626	3.868	6935.000	0.000	fixed	
LA (0)	0.067	0.344	0.950	-0.608	0.741	0.193	6935.000	0.847	fixed	
LA (2)	0.208	0.469	0.950	-0.712	1.128	0.444	6935.000	0.657	fixed	
LA (3)	-0.145	0.375	0.950	-0.881	0.590	-0.388	6935.000	0.698	fixed	
LA (4)	-0.659	0.502	0.950	-1.643	0.325	-1.312	6935.000	0.190	fixed	
LA (5)	-0.184	0.435	0.950	-1.036	0.668	-0.423	6935.000	0.672	fixed	
LA (6)	-0.368	0.350	0.950	-1.053	0.318	-1.051	6935.000	0.293	fixed	
LA (7)	0.183	0.400	0.950	-0.602	0.968	0.456	6935.000	0.648	fixed	
LA (8)	-0.162	0.368	0.950	-0.884	0.561	-0.438	6935.000	0.661	fixed	
LA (9)	-0.343	0.414	0.950	-1.154	0.468	-0.829	6935.000	0.407	fixed	
LA (11)	-0.202	0.360	0.950	-0.907	0.504	-0.560	6935.000	0.576	fixed	
LA (12)	-0.386	0.407	0.950	-1.184	0.413	-0.947	6935.000	0.344	fixed	
SD (Intercept)	0.223	NA	0.95	NA	NA	NA	NA	NA	random	School
SD (Observations)	3.057	NA	0.95	NA	NA	NA	NA	NA	random	Residual

Note. Note that the p-value for compliance differs from that reported in the main body (p=0.380). This is due to bootstrapping, which was used for the results in the main body because it produces more robust standard errors. SDQ = Strengths and Difficulties Questionnaire, intern = internalising, LA = local authority.

Appendix C Table 21: Primary Missing Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	t	df error	p	Effects	Group
Intercept	6.208	1.123	0.950	4.006	8.409	5.526	5884.000	0.000	fixed	
Group (Intervention)	0.053	0.142	0.950	-0.225	0.331	0.374	5884.000	0.708	fixed	
SDQ Extern (Baseline)	0.563	0.011	0.950	0.542	0.584	51.847	5884.000	0.000	fixed	
Cohort (Cohort 2)	-0.538	0.372	0.950	-1.267	0.190	-1.449	5884.000	0.147	fixed	
SDQ Intern (Baseline)	0.030	0.012	0.950	0.006	0.054	2.437	5884.000	0.015	fixed	
Gender (Female)	-0.005	0.089	0.950	-0.180	0.169	-0.059	5884.000	0.953	fixed	
Gender (Other)	-0.182	0.305	0.950	-0.781	0.416	-0.597	5884.000	0.551	fixed	
Ethnicity (Asian)	-0.420	0.117	0.950	-0.650	-0.190	-3.580	5884.000	0.000	fixed	
Ethnicity (Black)	-0.307	0.177	0.950	-0.655	0.041	-1.730	5884.000	0.084	fixed	
Ethnicity (Mixed)	-0.238	0.189	0.950	-0.609	0.133	-1.256	5884.000	0.209	fixed	
Ethnicity (Other)	-0.325	0.212	0.950	-0.740	0.090	-1.537	5884.000	0.124	fixed	
Year (9)	-0.116	0.082	0.950	-0.277	0.044	-1.419	5884.000	0.156	fixed	
School size	0.000	0.000	0.950	-0.001	0.000	-1.146	5884.000	0.252	fixed	
Attendance Perc (Baseline)	-0.020	0.008	0.950	-0.035	-0.004	-2.525	5884.000	0.012	fixed	
Sus (Baseline)	1.051	0.241	0.950	0.579	1.522	4.368	5884.000	0.000	fixed	
SEN (Yes)	0.387	0.119	0.950	0.154	0.620	3.254	5884.000	0.001	fixed	
LA (1)	-0.076	0.406	0.950	-0.873	0.720	-0.188	5884.000	0.851	fixed	
LA (3)	-0.107	0.262	0.950	-0.621	0.408	-0.407	5884.000	0.684	fixed	
LA (4)	-0.752	0.489	0.950	-1.711	0.208	-1.536	5884.000	0.125	fixed	
LA (5)	-0.463	0.462	0.950	-1.370	0.443	-1.003	5884.000	0.316	fixed	
LA (6)	-0.396	0.227	0.950	-0.841	0.049	-1.744	5884.000	0.081	fixed	

LA (7)	0.174	0.332	0.950	-0.476	0.825	0.525	5884.000	0.599	fixed	
LA (8)	0.036	0.347	0.950	-0.643	0.716	0.105	5884.000	0.917	fixed	
LA (9)	-0.325	0.333	0.950	-0.978	0.329	-0.973	5884.000	0.331	fixed	
LA (11)	-0.204	0.231	0.950	-0.657	0.249	-0.882	5884.000	0.378	fixed	
LA (12)	-0.793	0.492	0.950	-1.759	0.172	-1.611	5884.000	0.107	fixed	
SD (Intercept)	0.325	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	3.042	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. SDQ = Strengths and Difficulties Questionnaire, intern = internalising, perc = percent, sus = suspensions, SEN = special educational needs, LA = local authority.

Appendix C Table 22: Model output for the multi-level logistic regression to predict missingness in the primary outcome**.**

Parameter	Estimate	SE	z value	p
Intercept	5.160	2.241	2.303	0.021
Gender (Female)	0.106	0.079	1.334	0.182
Gender (Other)	0.087	0.244	0.357	0.721
Ethnicity (Asian)	-0.105	0.111	-0.944	0.345
Ethnicity (Black)	-0.057	0.164	-0.348	0.728
Ethnicity (Mixed)	-0.089	0.160	-0.555	0.579
Ethnicity (Other)	-0.030	0.191	-0.157	0.875
Year Group (Year 9)	0.329	0.068	4.846	<.001
FSM Eligibility (Yes)	0.107	0.079	1.348	0.178
SEN (Yes)	0.347	0.093	3.725	<.001
Baseline SDQ Extern	0.026	0.011	2.455	0.014
Attendance perc (23/24)	-0.057	0.006	-10.240	<.001
Suspensions (23/24)	0.780	0.160	4.862	<.001
LA (1)	2.459	0.875	2.810	0.005
LA (3)	-1.384	0.770	-1.798	0.072
LA (4)	-0.067	1.182	-0.057	0.955
LA (5)	0.950	0.664	1.431	0.152
LA (6)	-0.934	0.609	-1.533	0.125
LA (7)	-1.126	0.919	-1.225	0.221
LA (8)	-0.692	0.888	-0.779	0.436
LA (9)	-1.105	0.967	-1.142	0.253
LA (11)	1.201	0.735	1.635	0.102
LA (12)	-1.973	0.964	-2.047	0.041
School Size	-0.001	0.000	-2.360	0.018
Average School Baseline Extern	0.142	0.290	0.491	0.624
School FSM perc	-0.024	0.053	-0.451	0.652
School SEN perc	-0.021	0.021	-1.023	0.306
Baseline Intern	-0.005	0.012	-0.405	0.685
Baseline Prosocial	0.013	0.018	0.708	0.479
Baseline SEI TSR	-0.002	0.009	-0.208	0.835
Baseline SEI CRS	-0.004	0.011	-0.397	0.692
Baseline SEI PSAS	-0.001	0.012	-0.081	0.936
Baseline SEI FG	-0.004	0.014	-0.304	0.761
Baseline IBS Victim	-0.013	0.010	-1.371	0.170
Baseline IBS Perp	-0.003	0.007	-0.415	0.678

Note. CRS = control and relevance of schoolwork, extern = externalising, FG = future goals, FSM = free school meals, IBS = Illinois Bully Scale, intern = internalising, LA = local authority, perc = percent, perp = perpetration, PSAS = peer support at school, SDQ = Strengths and Difficulties Questionnaire, SEI = Student Engagement Instrument, SEN = special educational needs, TSR = teacher-student relationships, victim = victimisation.

Appendix C Table 23: Weighted Primary Model.

Parameter	Estimate	SE	CI	CI low	CI high	z	df error	p	Effects	Group
Intercept	2.763	0.244	0.950	2.284	3.242	11.306	Inf	0.000	fixed	
Group (Intervention)	0.041	0.099	0.950	-0.152	0.234	0.415	Inf	0.678	fixed	
SDQ Extern (Baseline)	0.592	0.008	0.950	0.575	0.608	69.975	Inf	0.000	fixed	
Cohort (Cohort 2)	-0.321	0.199	0.950	-0.711	0.069	-1.613	Inf	0.107	fixed	
SDQ Intern (Baseline)	0.038	0.010	0.950	0.019	0.056	3.880	Inf	0.000	fixed	
Gender (Female)	0.042	0.067	0.950	-0.089	0.174	0.629	Inf	0.530	fixed	
Gender (Other)	-0.337	0.282	0.950	-0.890	0.216	-1.194	Inf	0.232	fixed	
Ethnicity (Asian)	-0.366	0.087	0.950	-0.536	-0.195	-4.207	Inf	0.000	fixed	
Ethnicity (Black)	-0.373	0.139	0.950	-0.644	-0.101	-2.689	Inf	0.007	fixed	
Ethnicity (Mixed)	-0.196	0.150	0.950	-0.489	0.098	-1.305	Inf	0.192	fixed	
Ethnicity (Other)	-0.142	0.178	0.950	-0.491	0.208	-0.794	Inf	0.427	fixed	
LA (1)	-0.088	0.508	0.950	-1.083	0.908	-0.172	Inf	0.863	fixed	
LA (2)	0.257	0.359	0.950	-0.447	0.962	0.716	Inf	0.474	fixed	
LA (3)	-0.153	0.171	0.950	-0.487	0.181	-0.897	Inf	0.370	fixed	
LA (4)	-0.649	0.437	0.950	-1.505	0.207	-1.486	Inf	0.137	fixed	
LA (5)	-0.250	0.306	0.950	-0.850	0.349	-0.818	Inf	0.413	fixed	
LA (6)	-0.357	0.148	0.950	-0.647	-0.068	-2.421	Inf	0.015	fixed	
LA (7)	0.004	0.222	0.950	-0.431	0.439	0.017	Inf	0.986	fixed	
LA (8)	-0.244	0.194	0.950	-0.624	0.136	-1.256	Inf	0.209	fixed	
LA (9)	-0.407	0.229	0.950	-0.856	0.042	-1.777	Inf	0.075	fixed	
LA (11)	-0.157	0.171	0.950	-0.491	0.178	-0.919	Inf	0.358	fixed	
LA (12)	-0.519	0.230	0.950	-0.970	-0.069	-2.258	Inf	0.024	fixed	
SD (Intercept)	0.197	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	3.010	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. SDQ = Strengths and Difficulties Questionnaire, intern = internalising, LA = local authority.

Appendix C Table 24: Primary Imputation Model.

Parameter	Estimate	SE	CI	CI low	CI high	Statistic	df error	p
Intercept	4.386	1.047	0.950	2.332	6.440	4.188	1429.468	0.000
Group (Intervention)	0.070	0.124	0.950	-0.173	0.313	0.567	6895.677	0.571
SDQ Extern (Baseline)	0.576	0.010	0.950	0.557	0.596	57.662	6715.578	0.000
Cohort (Cohort 2)	-0.209	0.269	0.950	-0.737	0.318	-0.777	6865.612	0.437
SDQ Intern (Baseline)	0.028	0.011	0.950	0.006	0.051	2.505	6910.402	0.012
Gender (Male)	0.001	0.081	0.950	-0.158	0.161	0.017	6875.894	0.986
Gender (Other)	-0.127	0.277	0.950	-0.671	0.416	-0.459	6898.964	0.646
Ethnicity (Black)	-0.010	0.158	0.950	-0.320	0.299	-0.065	6900.690	0.948
Ethnicity (Mixed)	0.137	0.182	0.950	-0.219	0.494	0.755	6917.088	0.450
Ethnicity (Other)	0.082	0.193	0.950	-0.297	0.461	0.426	6903.938	0.670
Ethnicity (White)	0.379	0.110	0.950	0.164	0.594	3.459	6923.470	0.001
SEN(Yes)	0.339	0.115	0.950	0.112	0.565	2.933	953.427	0.003
Year (Final)	-0.085	0.076	0.950	-0.234	0.063	-1.123	6911.356	0.262
Attendance Perc (Baseline)	-0.015	0.007	0.950	-0.029	0.000	-1.961	394.290	0.051
Sus (Baseline)	0.799	0.237	0.950	0.334	1.264	3.376	369.566	0.001
LA (2)	0.204	0.534	0.950	-0.842	1.251	0.382	6883.015	0.702
LA (3)	-0.034	0.419	0.950	-0.856	0.787	-0.082	6928.182	0.935
LA (4)	-0.558	0.568	0.950	-1.671	0.556	-0.982	6927.754	0.326
LA (5)	-0.121	0.496	0.950	-1.094	0.851	-0.244	6922.923	0.807
LA (6)	-0.276	0.391	0.950	-1.042	0.490	-0.707	6927.423	0.480
LA (7)	0.231	0.458	0.950	-0.666	1.128	0.504	6927.548	0.614
LA (8)	-0.115	0.415	0.950	-0.929	0.699	-0.276	6861.947	0.782
LA (9)	-0.260	0.466	0.950	-1.174	0.655	-0.557	6928.283	0.578
LA (11)	-0.158	0.401	0.950	-0.944	0.628	-0.394	6927.708	0.694
LA (13)	0.091	0.384	0.950	-0.662	0.845	0.238	6921.154	0.812
LA (12)	-0.322	0.466	0.950	-1.234	0.591	-0.691	6909.804	0.490

Note. SDQ = Strengths and Difficulties Questionnaire, intern = internalising, SEN = special educational needs, perc = percent, sus = suspension, LA = local authority.

Appendix C Table 25: Interaction by Ethnicity Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	T	df error	P	Effects	Group
Intercept	2.860	0.329	0.950	2.215	3.505	8.692	6931.000	0.000	fixed	
Group (Intervention)	-0.018	0.160	0.950	-0.331	0.295	-0.114	6931.000	0.909	fixed	
SDQ Extern (Baseline)	0.586	0.010	0.950	0.567	0.605	59.612	6931.000	0.000	fixed	
Ethnicity (Asian)	-0.431	0.134	0.950	-0.695	-0.168	-3.210	6931.000	0.001	fixed	
Ethnicity (Black)	-0.724	0.214	0.950	-1.144	-0.304	-3.382	6931.000	0.001	fixed	
Ethnicity (Mixed)	-0.307	0.219	0.950	-0.736	0.122	-1.401	6931.000	0.161	fixed	
Ethnicity (Other)	-0.014	0.264	0.950	-0.531	0.502	-0.054	6931.000	0.957	fixed	
Cohort (Cohort 2)	-0.244	0.273	0.950	-0.780	0.291	-0.894	6931.000	0.371	fixed	
SDQ Intern (Baseline)	0.030	0.011	0.950	0.008	0.052	2.646	6931.000	0.008	fixed	
Gender (Female)	-0.023	0.081	0.950	-0.182	0.137	-0.277	6931.000	0.781	fixed	
Gender (Other)	-0.133	0.278	0.950	-0.679	0.412	-0.479	6931.000	0.632	fixed	
LA (1)	-0.097	0.389	0.950	-0.859	0.666	-0.249	6931.000	0.804	fixed	
LA (2)	0.054	0.424	0.950	-0.776	0.885	0.128	6931.000	0.898	fixed	
LA (3)	-0.191	0.244	0.950	-0.670	0.288	-0.782	6931.000	0.434	fixed	
LA (4)	-0.657	0.467	0.950	-1.572	0.259	-1.406	6931.000	0.160	fixed	
LA (5)	-0.227	0.371	0.950	-0.955	0.501	-0.611	6931.000	0.541	fixed	
LA (6)	-0.426	0.200	0.950	-0.818	-0.034	-2.131	6931.000	0.033	fixed	
LA (7)	0.109	0.319	0.950	-0.515	0.734	0.343	6931.000	0.732	fixed	
LA (8)	-0.217	0.249	0.950	-0.705	0.271	-0.871	6931.000	0.384	fixed	
LA (9)	-0.395	0.320	0.950	-1.023	0.233	-1.234	6931.000	0.217	fixed	
LA (11)	-0.300	0.213	0.950	-0.718	0.118	-1.408	6931.000	0.159	fixed	
LA (12)	-0.493	0.331	0.950	-1.143	0.156	-1.490	6931.000	0.136	fixed	
Group (Intervention) x Ethnicity (Asian)	0.102	0.209	0.950	-0.308	0.511	0.486	6931.000	0.627	fixed	
Group (Intervention) x Ethnicity (Black)	0.650	0.316	0.950	0.031	1.269	2.057	6931.000	0.040	fixed	
Group (Intervention) x Ethnicity (Mixed)	0.193	0.364	0.950	-0.521	0.908	0.531	6931.000	0.596	fixed	
Group (Intervention) x Ethnicity (Other)	-0.490	0.388	0.950	-1.250	0.271	-1.263	6931.000	0.207	fixed	
SD (Intercept)	0.322	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	3.058	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. SDQ = Strengths and Difficulties Questionnaire, intern = internalising, LA = local authority,

Appendix C Table 26: Interaction by SEN Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	t	df error	P	Effects	Group
Intercept	3.181	0.418	0.950	2.362	4.001	7.608	5976.000	0.000	fixed	
Group (Intervention)	0.069	0.143	0.950	-0.211	0.350	0.484	5976.000	0.628	fixed	
SDQ Extern (Baseline)	0.578	0.011	0.950	0.557	0.599	54.397	5976.000	0.000	fixed	
SEN(Yes)	0.590	0.150	0.950	0.296	0.883	3.941	5976.000	0.000	fixed	
Ethnicity (Asian)	-0.363	0.116	0.950	-0.590	-0.137	-3.140	5976.000	0.002	fixed	
Ethnicity (Black)	-0.387	0.173	0.950	-0.726	-0.049	-2.243	5976.000	0.025	fixed	
Ethnicity (Mixed)	-0.194	0.187	0.950	-0.561	0.173	-1.035	5976.000	0.301	fixed	
Ethnicity (Other)	-0.218	0.208	0.950	-0.627	0.190	-1.048	5976.000	0.295	fixed	
Cohort (Cohort 2)	-0.641	0.368	0.950	-1.363	0.081	-1.739	5976.000	0.082	fixed	
SDQ Intern (Baseline)	0.024	0.012	0.950	0.000	0.048	1.983	5976.000	0.047	fixed	
Gender (Female)	0.005	0.088	0.950	-0.168	0.177	0.052	5976.000	0.958	fixed	
Gender (Other)	-0.213	0.305	0.950	-0.811	0.384	-0.700	5976.000	0.484	fixed	
LA (1)	-0.167	0.395	0.950	-0.941	0.607	-0.422	5976.000	0.673	fixed	
LA (3)	-0.145	0.250	0.950	-0.636	0.346	-0.579	5976.000	0.563	fixed	
LA (4)	-0.654	0.473	0.950	-1.580	0.273	-1.383	5976.000	0.167	fixed	
LA (5)	-0.595	0.453	0.950	-1.482	0.293	-1.314	5976.000	0.189	fixed	
LA (6)	-0.400	0.214	0.950	-0.819	0.020	-1.869	5976.000	0.062	fixed	
LA (7)	0.193	0.325	0.950	-0.444	0.831	0.594	5976.000	0.552	fixed	

LA (8)	0.067	0.338	0.950	-0.595	0.730	0.199	5976.000	0.843	fixed	
LA (9)	-0.365	0.326	0.950	-1.003	0.273	-1.121	5976.000	0.263	fixed	
LA (11)	-0.293	0.221	0.950	-0.726	0.140	-1.326	5976.000	0.185	fixed	
LA (12)	-0.940	0.486	0.950	-1.892	0.013	-1.933	5976.000	0.053	fixed	
Group (Intervention) x SEN (Yes)	-0.304	0.239	0.950	-0.773	0.166	-1.268	5976.000	0.205	fixed	
SD (Intercept)	0.321	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	3.051	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. SEN = Special Educational Needs, SDQ = Strengths and Difficulties Questionnaire, intern = internalising, LA = local authority,

Appendix C Table 27: Interaction by FSM Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	t	df error	P	Effects	Group
Intercept	3.087	0.416	0.950	2.272	3.902	7.425	5931.000	0.000	fixed	
Group (Intervention)	0.061	0.150	0.950	-0.234	0.356	0.406	5931.000	0.685	fixed	
SDQ Extern (Baseline)	0.578	0.011	0.950	0.557	0.599	54.073	5931.000	0.000	fixed	
FSM (Yes)	0.451	0.120	0.950	0.215	0.687	3.752	5931.000	0.000	fixed	
Ethnicity (Asian)	-0.411	0.117	0.950	-0.639	-0.182	-3.524	5931.000	0.000	fixed	
Ethnicity (Black)	-0.476	0.174	0.950	-0.817	-0.135	-2.739	5931.000	0.006	fixed	
Ethnicity (Mixed)	-0.268	0.188	0.950	-0.636	0.101	-1.425	5931.000	0.154	fixed	
Ethnicity (Other)	-0.267	0.209	0.950	-0.677	0.143	-1.276	5931.000	0.202	fixed	
Cohort (Cohort 2)	-0.578	0.364	0.950	-1.291	0.136	-1.587	5931.000	0.113	fixed	
SDQ Intern (Baseline)	0.026	0.012	0.950	0.002	0.050	2.131	5931.000	0.033	fixed	
Gender (Female)	-0.007	0.088	0.950	-0.180	0.166	-0.083	5931.000	0.934	fixed	
Gender (Other)	-0.259	0.307	0.950	-0.861	0.343	-0.843	5931.000	0.399	fixed	
LA (1)	-0.091	0.391	0.950	-0.858	0.676	-0.233	5931.000	0.816	fixed	
LA (3)	-0.139	0.247	0.950	-0.622	0.344	-0.563	5931.000	0.573	fixed	
LA (4)	-0.615	0.467	0.950	-1.531	0.302	-1.315	5931.000	0.188	fixed	
LA (5)	-0.508	0.448	0.950	-1.386	0.370	-1.135	5931.000	0.257	fixed	
LA (6)	-0.368	0.212	0.950	-0.783	0.047	-1.740	5931.000	0.082	fixed	
LA (7)	0.238	0.321	0.950	-0.392	0.867	0.740	5931.000	0.459	fixed	
LA (8)	0.084	0.334	0.950	-0.570	0.738	0.252	5931.000	0.801	fixed	
LA (9)	-0.317	0.321	0.950	-0.947	0.312	-0.989	5931.000	0.323	fixed	
LA (11)	-0.251	0.225	0.950	-0.693	0.190	-1.116	5931.000	0.264	fixed	
LA (12)	-0.807	0.481	0.950	-1.749	0.136	-1.678	5931.000	0.093	fixed	
Group (Intervention) x FSM (Yes)	-0.198	0.183	0.950	-0.557	0.161	-1.080	5931.000	0.280	fixed	
SD (Intercept)	0.313	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	3.051	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. FSM = Free School Meals, SEN = special educational needs, SDQ = Strengths and Difficulties Questionnaire, intern = internalising, LA = local authority.

Appendix C Table 28: Primary Midline Model Summary

Parameter	Estimate	SE	CI	CI low	CI high	t	df error	p	Effects	Group
Intercept	2.989	0.247	0.950	2.506	3.472	12.120	12768.000	0.000	fixed	
Group (Intervention)	-0.098	0.102	0.950	-0.298	0.101	-0.966	12768.000	0.334	fixed	
Time (Endline vs Midline)	-0.568	0.057	0.950	-0.679	-0.457	-10.028	12768.000	0.000	fixed	
SDQ Extern (Baseline)	0.650	0.008	0.950	0.635	0.665	83.886	12768.000	0.000		
Cohort (Cohort 2)	-0.283	0.203	0.950	-0.680	0.114	-1.397	12768.000	0.162	fixed	
SDQ Intern (Baseline)	0.022	0.009	0.950	0.004	0.039	2.433	12768.000	0.015	fixed	
Gender (Female)	0.008	0.064	0.950	-0.118	0.133	0.118	12768.000	0.906	fixed	
Gender (Other)	0.082	0.214	0.950	-0.337	0.501	0.383	12768.000	0.702	fixed	
Ethnicity (Asian)	-0.372	0.087	0.950	-0.542	-0.202	-4.290	12768.000	0.000	fixed	
Ethnicity (Black)	-0.369	0.130	0.950	-0.624	-0.114	-2.839	12768.000	0.005	fixed	
Ethnicity (Mixed)	-0.290	0.141	0.950	-0.567	-0.013	-2.054	12768.000	0.040	fixed	
Ethnicity (Other)	-0.282	0.158	0.950	-0.592	0.027	-1.789	12768.000	0.074	fixed	
LA (1)	-0.131	0.277	0.950	-0.674	0.412	-0.472	12768.000	0.637	fixed	
LA (2)	-0.119	0.299	0.950	-0.705	0.468	-0.397	12768.000	0.691		
LA (3)	-0.109	0.183	0.950	-0.469	0.250	-0.596	12768.000	0.551	fixed	
LA (4)	-0.262	0.318	0.950	-0.885	0.362	-0.822	12768.000	0.411	fixed	
LA (5)	-0.359	0.274	0.950	-0.897	0.178	-1.311	12768.000	0.190	fixed	
LA (6)	-0.398	0.146	0.950	-0.685	-0.111	-2.715	12768.000	0.007	fixed	
LA (7)	0.051	0.246	0.950	-0.431	0.533	0.207	12768.000	0.836	fixed	
LA (8)	-0.105	0.182	0.950	-0.461	0.251	-0.576	12768.000	0.565	fixed	
LA (9)	-0.461	0.237	0.950	-0.926	0.005	-1.941	12768.000	0.052	fixed	
LA (10)	0.309	0.677	0.950	-1.018	1.636	0.457	12768.000	0.648		
LA (11)	-0.162	0.152	0.950	-0.460	0.136	-1.064	12768.000	0.287	fixed	
LA (12)	-0.438	0.237	0.950	-0.902	0.026	-1.849	12768.000	0.065	fixed	
Group (Intervention) x Time (Endline vs Midline)	0.140	0.089	0.950	-0.035	0.314	1.568	12768.000	0.117	fixed	
SD (Intercept)	1.970	NA	0.950	NA	NA	NA	NA	NA	random	Pupil
SD (Intercept)	0.223	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	2.199	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. SDQ = Strengths and Difficulties Questionnaire, intern = internalising, FSM = Free School Meals, SEN = special educational needs, LA = local authority.

Appendix C Table 29: Interaction by ARTIC On the Job Behaviour Model Summary.

Parameter	Estimate	SE	CI	CI low	CI high	T	df error	p	Effects	Group
Intercept	5.044	1.510	0.950	2.084	8.005	3.340	6891.000	0.001	fixed	
Group (Intervention)	-0.691	1.942	0.950	-4.498	3.115	-0.356	6891.000	0.722	fixed	
SDQ Extern (Baseline)	0.586	0.010	0.950	0.567	0.606	59.381	6891.000	0.000	fixed	
Mean ARTIC (on the job behaviour)	-0.366	0.269	0.950	-0.893	0.162	-1.359	6891.000	0.174	fixed	
Cohort (Cohort 2)	-0.503	0.288	0.950	-1.068	0.062	-1.744	6891.000	0.081	fixed	
SDQ Intern (Baseline)	0.030	0.011	0.950	0.008	0.053	2.679	6891.000	0.007	fixed	
Gender (Female)	-0.014	0.082	0.950	-0.173	0.146	-0.167	6891.000	0.867	fixed	
Gender (Other)	-0.144	0.280	0.950	-0.694	0.405	-0.514	6891.000	0.607	fixed	
Ethnicity (Asian)	-0.432	0.113	0.950	-0.654	-0.210	-3.820	6891.000	0.000	fixed	
Ethnicity (Black)	-0.451	0.162	0.950	-0.767	-0.134	-2.791	6891.000	0.005	fixed	
Ethnicity (Mixed)	-0.250	0.177	0.950	-0.597	0.096	-1.417	6891.000	0.156	fixed	
Ethnicity (Other)	-0.288	0.196	0.950	-0.673	0.096	-1.470	6891.000	0.141	fixed	
LA (1)	0.031	0.396	0.950	-0.746	0.807	0.077	6891.000	0.939	fixed	
LA (2)	0.253	0.448	0.950	-0.625	1.131	0.566	6891.000	0.572	fixed	
LA (3)	-0.226	0.239	0.950	-0.694	0.242	-0.947	6891.000	0.344	fixed	
LA (4)	-0.701	0.458	0.950	-1.600	0.197	-1.530	6891.000	0.126	fixed	
LA (5)	-0.518	0.386	0.950	-1.274	0.238	-1.342	6891.000	0.180	fixed	
LA (6)	-0.481	0.201	0.950	-0.875	-0.086	-2.389	6891.000	0.017	fixed	
LA (7)	0.023	0.317	0.950	-0.597	0.644	0.074	6891.000	0.941	fixed	
LA (8)	-0.152	0.250	0.950	-0.641	0.338	-0.607	6891.000	0.544	fixed	
LA (9)	-0.457	0.320	0.950	-1.085	0.170	-1.428	6891.000	0.153	fixed	
LA (11)	-0.401	0.225	0.950	-0.843	0.040	-1.782	6891.000	0.075	fixed	
LA (12)	-0.707	0.342	0.950	-1.378	-0.036	-2.064	6891.000	0.039	fixed	
Group (Intervention) x Mean ARTIC (on the job behaviour)	0.155	0.361	0.950	-0.552	0.862	0.429	6891.000	0.668	fixed	
SD (Intercept)	0.328	NA	0.950	NA	NA	NA	NA	NA	random	School
SD (Observations)	2.471	NA	0.950	NA	NA	NA	NA	NA	random	Residual

Note. ARTIC = Attitudes Related to Trauma, SDQ = Strengths and Difficulties Questionnaire, intern = internalising, LA = local authority.

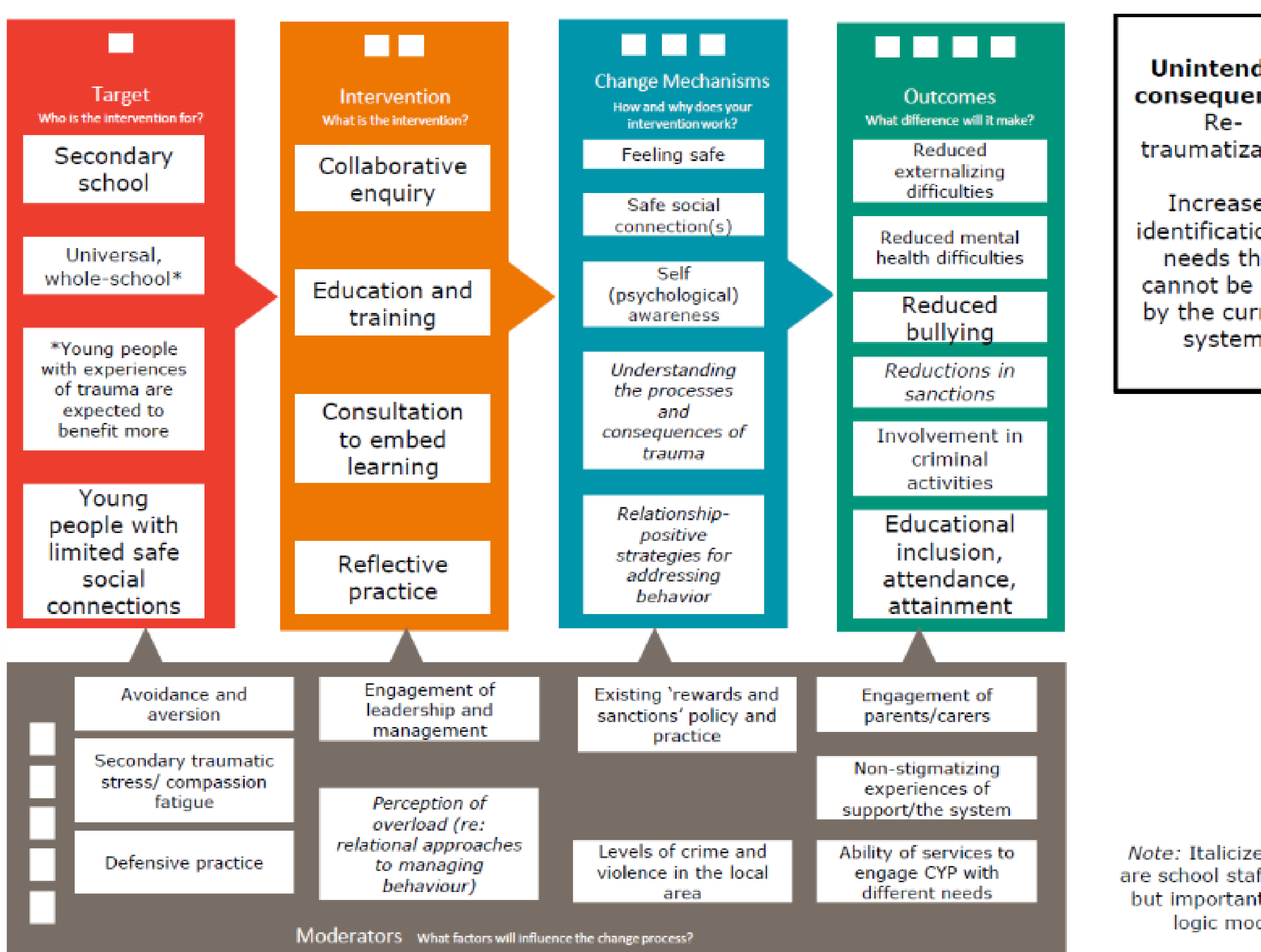
Appendix C Table 30: Longitudinal Mediation Analysis of SEI TSR.

To	From	Estimate	SE	CI low	CI high	z	p	Component
SEI TSR (Midline)	group	0.062	0.454	-0.828	0.952	0.137	0.891	Regression
SDQ extern (Endline)	SEI TSR (Midline)	-0.242	0.014	-0.270	-0.215	-17.168	0.000	Regression
SDQ extern (Endline)	group	-0.156	0.245	-0.636	0.323	-0.639	0.523	Regression
Indirect	a*b	-0.015	0.110	-0.230	0.200	-0.137	0.891	Defined
Total	c+(a*b)	-0.171	0.295	-0.749	0.407	-0.581	0.561	Defined

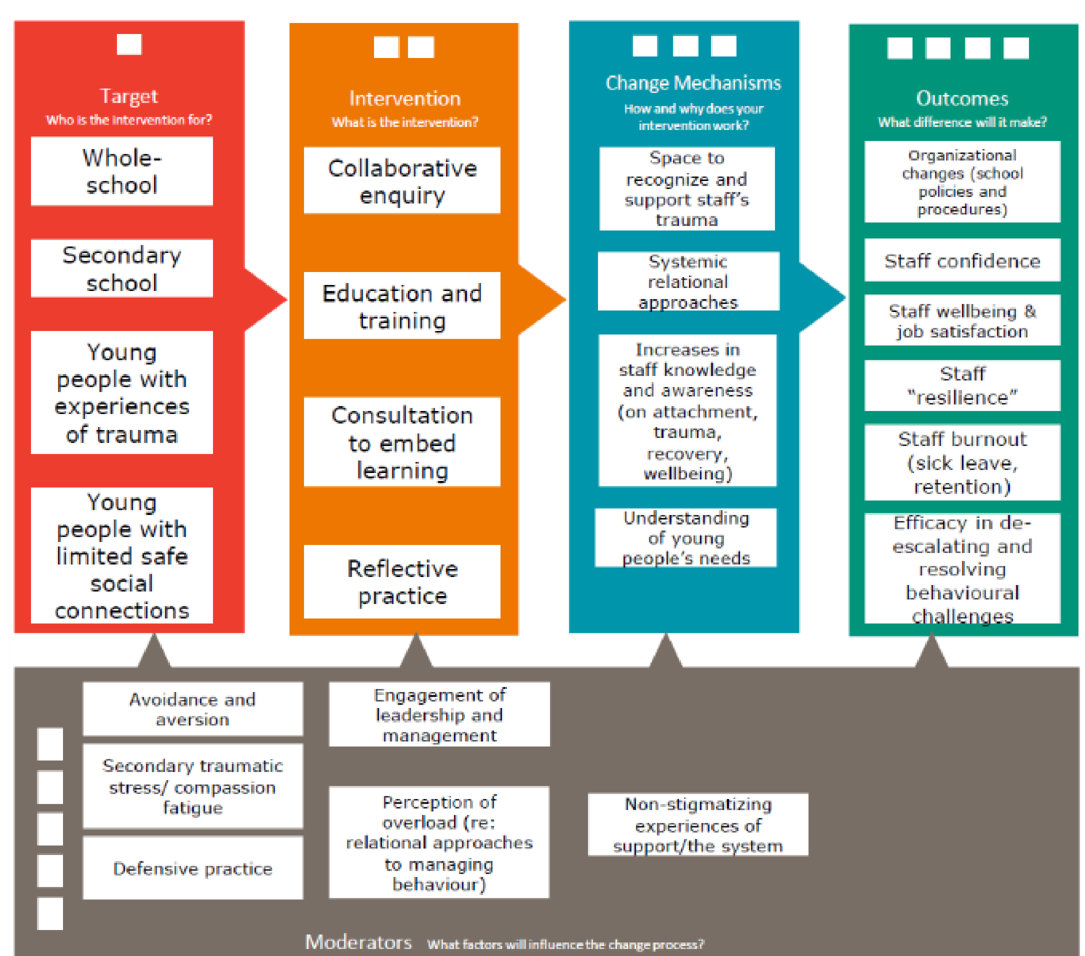
Note. SEI = Student Engagement Instrument, TSR = teacher-student relationships, SDQ = Strengths and Difficulties Questionnaire, extern = externalising

Appendix D: Logic model

Appendix D Figure 1: Evaluation logic model for pupils.



Appendix D Figure 2: Evaluation logic model for staff.



Target
Who is the intervention for?
Whole-school
Secondary school
Young people with experiences of trauma
Young people with limited safe social connections
Intervention
What is the intervention?
Collaborative enquiry
Education and training
Consultation to embed learning
Reflective practice
Change Mechanisms
How and why does your intervention work?
Space to recognize and support staff's trauma
Systemic relational approaches
Increases in staff knowledge and awareness (on attachment, trauma, recovery, wellbeing)
Understanding of young people's needs
Outcomes
What difference will it make?
Organizational changes (school policies and procedures)
Staff confidence
Staff wellbeing & job satisfaction
Staff "resilience"
Staff burnout (sick leave, retention)
Efficacy in de-escalating and resolving behavioural challenges
Unintended consequences
Re-traumatization
Increased identification of needs that cannot be met by the current system
Avoidance and aversion
Secondary traumatic stress/ compassion fatigue
Defensive practice
Engagement of leadership and management
Perception of overload (re: relational approaches to managing behaviour)
Non-stigmatizing experiences of support/the system
Moderators What factors will influence the change process?

Appendix E. YEF Security Rating

Rating	Design	MDES Outcome: Threshold*	Attrition	Initial score	Adjustments	Final score
5	Randomised design	Offending: <=0.1 SDQ tot: <= 0.3 Other: <= 0.2	0-10%	1		
4	Design for comparison that considers some type of selection on unobservable characteristics (e.g. RDD, Diff-in-Diffs, Matched Diff-in-Diffs)	Offending: 0.11 – 0.19 SDQ tot: 0.31 – 0.39 Other: 0.21 – 0.29	11-20%			
3	Design for comparison that considers selection on all relevant observable confounders (e.g. Matching or Regression Analysis with variables descriptive of the selection mechanism)	Offending: 0.2 – 0.29 SDQ tot: 0.4 – 0.49 Other: 0.3 – 0.39	21-30%			**1**
2	Design for comparison that considers selection only on some relevant confounders	Offending: 0.3 – 0.39 SDQ tot: 0.5 – 0.59 Other: 0.4 – 0.49	31-40%			
1	Design for comparison that does not consider selection on any relevant confounders	Offending: 0.4 – 0.49 SDQ tot: 0.6 – 0.69 Other: 0.5 – 0.59	41-50%			
0	No comparator	Offending: >= 0.5 SDQ tot: >= 0.7 Other: >= 0.6	>50%			

*MDES requirements vary by outcome measurement. Offending: Offending data collected through self-report or admin data; SDQ tot = SDQ total difficulties score; Other: all other outcomes, incl. SDQ externalising and internalising