Plan Overview

A Data Management Plan created using DMPonline

Title: Passport to Success

Creator: Joao Santos

Principal Investigator: Neil Humphrey, Pamela Qualter

Data Manager: Joao Santos

Project Administrator: Joao Santos

Contributor: Margarita Panayiotou, Ola Demkowicz, Alexander Thompson, Jan Boehnke, Jennifer Lau, Liz Neill

Affiliation: University of Manchester

Template: University of Manchester Generic Template

ORCID iD: 0000-0002-8148-9500

ORCID iD: 0000-0001-6114-3820

Project abstract:

Research shows that a considerable proportion of 5-16 year-olds experience mental health difficulties. These difficulties, defined as enduring, maladaptive changes in thoughts, feelings and/or behaviour, impair quality of life and are concurrently and prospectively associated with academic attainment and other salient outcomes. The transition from childhood to adolescence (with adolescence comprising the period between 10–24 years of age appears to be a crucial period, given the major physical, psychological and social changes that occur. In light of these circumstances this project aims to examine the utility of a universal social and emotional learning (SEL) intervention (“Passport to Success”) in promoting emotion regulation to reduce internalising symptoms, loneliness, and bullying during the critical developmental transition from childhood to adolescence. The principal objective is to significantly enhance SEL intervention research, in relation to 'Passport to Success', by examining:

- Main (‘intent to treat’) effects;
- Maintenance/sleeper effects;
- Implementation moderator effects;
- Subgroup moderator effects;
- Cost-effectiveness;
- Mechanisms underpinning changes in outcomes.

To accomplish the objectives detailed above, we will implement a robust trial that will follow a two-group (intervention vs. control) parallel cluster RCT design, with schools as the unit of randomisation. Schools allocated to the intervention arm will implement Passport during the academic year 23/24, while schools in the control arm will continue with their usual practice

during this period. Outcomes will be assessed annually at baseline/pre-randomisation (T0), post-intervention (T1), and 12-month post-intervention follow-up (T2).

Our trial setting is primary schools in England. We will recruit our trial sample from mainstream primary schools in the 10 Local Authorities (LAs) of the Greater Manchester (GM) city-region because they provide great diversity in their urbanicity, socio-economic status, ethnicity, and other factors that will help to ensure that our research setting reflects the heterogeneity across England. Our target population is children aged 8-9 years (Year 4) at T0 attending English primary schools.

ID: 99673

Start date: 01-10-2022

End date: 31-12-2025

Last modified: 01-02-2024

Grant number / URL: Kavli2021-0000000019

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Passport to Success

Manchester Data Management Outline

1. Will this project be reviewed by any of the following bodies (please select all that apply)?

- Ethics
- Funder

2. Is The University of Manchester collaborating with other institutions on this project?

- Yes - Part of a collaboration and owning or handling data

This project involves a collaboration between the University of Manchester, the Common Room, the University of Dundee, the Queen Mary University of London, and the Kavli Trust (the funder of this project).

3. What data will you use in this project (please select all that apply)?

- Acquire new data
- Re-use existing data (please list below)

This project will entail the use of new data and the re-use of existing data.
Regarding **existing data**, the participants' socio demographic data will be obtained from up to participating across Greater Manchester. The data to be obtained includes:

- **Unique Pupil Number**
- **Forename and Surname**
- **School Name**
- **Sex**
- **Year group**
- **Class and class teacher**
- **Ethnicity** (Any Other Ethnic Group, Asian, Black, Chinese, Mixed, Unclassified, White)
- **Free school meal eligibility** (yes/no)
- **Special educational needs** (no/SEN support/SEN EHCP)

Data sharing/transfer agreements will be set up with participating schools. The purpose of these agreements is to allow the schools to provide these data directly to the research team.
The aforementioned information will be used to:

- ensure accurate 1:1 matching (see below regarding National Pupil Database)
- generate pupil password lists for the completion of the pupil online surveys (thus ensuring that each pupil has an individual project password that allows access to the online surveys)
- Perform statistical analyses in service of project objectives
- Generate aggregated feedback reports for schools

We will use further existing data:

- **Attainment data**. These data (Key Stage 2 academic end data) will be acquired from the National Pupil Database (NPD) via the Office for National Statistics (ONS).

Participant surveys will allow the generation of **new data** involving multiple mental health and wellbeing constructs, as follows:

- **Internalizing symptoms**
- **Emotional regulation**
- **Wellbeing**
- **Loneliness**
- **Bullying**
- **Peer support**
- **Health-related quality of life (HRQoL)**.

Additional new data will be generated via online surveys designed to assess **implementation** and **process evaluation**. These

surveys include the following:

- A ‘usual SEL practice’ survey (administered to class teachers at T0 and T1) to establish the counterfactual to Passport (in control schools) and the level of programme differentiation (in intervention schools).
- An implementation survey (administered to class teachers in intervention schools at T1) to gather data on aspects of implementation including dosage (how much of the intervention was delivered) and fidelity and adaptation (whether the intervention was delivered as intended and any adjustments made).
- Brief teacher/classroom level outcome measures (administered at T0), including teachers’ perception of SEL culture, teacher burnout, and classroom management, to improve compliance prediction.
- Qualitative data generation (semi-structured interviews with teachers and other staff, such as members of senior leadership, and focus groups with children) in 5 purposively sampled intervention schools, to expand on aspects of implementation covered in the survey and explore more nuanced considerations, such as adaptations and reasoning for these, acceptability (whether the intervention is perceived as helpful and appropriate for staff and children), and reach and responsiveness (how children respond to the intervention), while also highlighting factors affecting implementation.

Once the project is completed, all new data that are generated by the project will be included in a public data set in an anonymised format. Data obtained from participating schools, and from the National Pupil Database will not be included in the project dataset that is made available to the public.

4. Where will the data be stored and backed-up during the project lifetime?

- University of Manchester Research Data Storage Service (Isilon)

Data will be kept in an encrypted Veracrypt container/folder (AES-256 algorithm), in a protected folder on the University's Research Data Storage system, known as Isilon. Accessing the data requires dual factor authentication to log into the university's network. The folder with the Veracrypt file can only be accessed by named users, and the container is itself protected with a long password (30+ characters).

The creation of encrypted drives and their storage in the University folder is a process that has been used by the research team for the several years, with no data breaches. The nature of the project means that data has to be accessed frequently and elements exported on demand on a daily basis by the data manager and/or relevant team members. The aforementioned system allows for this while maintaining a very high level of security.

The encrypted container noted above is stored on an access restricted data share on the University’s network storage infrastructure which is the recommended location for storing sensitive or critical University data. The storage infrastructure is hosted across two data centres (approx. 4KM apart) for resilience and disaster recovery purposes. Physical access to the data centres is strictly limited to data centre staff and a limited number of authorised IT Services staff. The data centres are protected by physical and electronic access security systems, swipe card access in and out of the data centres and CCTV coverage. The data centres are locked down out of hours and access is discouraged, but can be arranged by prior agreement with the data centre manager.
The University’s IT Services utilises Legato Networker Backup domains. Supporting infrastructure comprises disk libraries and both physical and virtual tape libraries. Cross data centre backup is performed, so services hosted within data centre 1 (Joule House) are backed up to data centre 2 (Reynold House) and vice versa.
Backup/recovery plans are documented as part of the service install process during the commissioning of a specific service. Each Service is responsible for its business continuity and disaster recovery plans, to which IT Services feed in its technical recovery plans ITSD operates change and release management processes. All proposed changes to infrastructure hosted, maintained and administered by IT Services are recorded via the Request for Change (RFC) process with changes being reviewed and authorised by a Change Advisory Board (CAB).
Research data storage is not backed up to tape in the traditional fashion. However, resilience is obtained through replication and use of snapshots means that files deleted by accident or corrupted can be recovered. Hourly backups can be accessed within 24 hours and daily backups can be accessed within 35 days.

5. If you will be using Research Data Storage, how much storage will you require?

- < 1 TB

6. Are you going to be receiving data from, or sharing data with an external third party?

- Yes

The participating schools will provide participant socio-demographic data as detailed in item 3. The research team will provide anonymised reports at the end of the project to the participating schools. These reports will describe the online survey results in the form of summary statistics (both in tabular and data visualisation formats).

Data sharing agreements, approved by the information governance leads at respective parties, will be implemented.
The research team will apply for access to datasets of the National Pupil Database. The objective will be to obtain attainment data for the pupils participating in the project. Some of the sociodemographic variables (Unique Pupil Number, Forename, Surname, School Name) will be provided to the NPD in order to ensure matching with attainment data. Said matching will take place within the Office for National Statistics secure web environment.

7. How long do you intend to keep your data for after the end of your project (in years)?

- 11 - 20 years

Guidance for questions 8 to 13

Highly restricted information defined in the Information security classification, ownership and secure information handling SOP is information that requires enhanced security as unauthorised disclosure could cause significant harm to individuals or to the University and its ambitions in respect of its purpose, vision and values. This could be: information that is subject to export controls; valuable intellectual property; security sensitive material or research in key industrial fields at particular risk of being targeted by foreign states. See more examples of highly restricted information.

Personal information, also known as personal data, relates to identifiable living individuals. Personal data is classed as special category personal data if it includes any of the following types of information about an identifiable living individual: racial or ethnic origin; political opinions; religious or similar philosophical beliefs; trade union membership; genetic data; biometric data; health data; sexual life; sexual orientation.

Please note that in line with data protection law (the UK General Data Protection Regulation and Data Protection Act 2018), personal information should only be stored in an identifiable form for as long as is necessary for the project; it should be pseudonymised (partially de-identified) and/or anonymised (completely de—identified) as soon as practically possible. You must obtain the appropriate ethical approval in order to use identifiable personal data.

8. What type of information will you be processing (please select all that apply)?

- Personal information, including signed consent forms
- Special category personal data, or criminal offence data
- Anonymised personal data
- Pseudonymised personal data
- Audio and/or video recordings

- **Personal information, including signed consent forms** The project will collect the pupils' unique pupil numbers, forename, surname, residential post-code and date of birth. All this information will be stored in accordance to the guidelines provided in item 4. The project also entails opt-out parental/carer consent, meaning that opt-out forms will be received at various points in the project lifespan. These forms contain personal information (e.g. forename, surname) and will be kept in a locked drawer in the Ellen Wilkinson Building. This same storage location will be used for the audio recorders that the research team utilises during the semi structured interviews.
- **Special categories**. Ethnicity and health (mental health and wellbeing measures discriminated in item 3).
- **Pseudonymised personal data**. Statistical modelling will be implemented using pseudonymised personal data, to ensure accurate assessment of the intervention's effectiveness.
- **Anonymised personal data**. The 'public' version of the project dataset will be a completely anonymised version which will not allow re-identification of participants (both pupils and teachers).
- **Aggregate data**. Aggregate pupil anonymised data will be provided via reports to schools in tabular and data visualisation formats. The data will be provided in a way that prevents the re-identification of participating pupils. Measures are built into the reports to prevent disclosing information to schools with fewer than 10 participating pupils, per ONS disclosure guidance.
- **Audio recordings**. Qualitative data will be generated as disclosed in item 3. Audio recordings will be obtained either in person, using audio recorders, or via Zoom. Although video will be enabled during the interview, Zoom will be set up so that only audio is recorded. The audio recording will then be stored in an encrypted Veracrypt container/folder (AES-256 algorithm), in a protected folder on the University's Research Data Storage system, known as Ipsilon (as disclosed in item 4). The research team will abide by the Standard Operating Procedures of the University of Manchester regarding recordings of participants for research projects.
- **Attainment data** will be obtained from the National Pupil Database, via the Office for National Statistics. Access to this data requires abiding by the security measures implemented by both the NPD and the ONS.

9. How do you plan to store, protect and ensure confidentiality of any highly restricted data or personal data (please select all that apply)?

- Store data in buildings, rooms or filing cabinets with controlled access
- Store data in encrypted files, folders, computers or devices
- Pseudonymise data and apply secure key management procedures
- Anonymise data
- Impose suitable data sharing and collaboration agreements
- Where needed, follow University of Manchester guidelines for disposing of personal data
- Access data hosted by the University of Manchester via its secure Virtual Private Network (VPN)
- Store data on University of Manchester approved and securely backed up servers or computers
- Access data hosted by a third-party data provider via their secure facilities (e.g. the UK Data Service Secure Lab)

- **Store data in buildings, rooms or filing cabinets with controlled access** Room B4.10 in the Ellen Wilkinson building will contain the locked filing cabinets where opt out consent forms and audio recorders are stored (although audio recorders will have their data uploaded to a secure setting and subsequently deleted as soon as possible). This room can only be accessed using a key.
- **Store data on University of Manchester approved and securely backed up servers or computers** All data will be stored on servers that are UoM approved. These are securely backed up.
- **Store data in encrypted files, folders, computers or devices** All electronic and audio data are to be kept in an encrypted Veracrypt container/folder (AES-256 algorithm), in a protected folder on Ipsilon.
- **Anonymise data**. The 'public' version of the project data set will consist of a completely anonymised data set which will not allow re-identification of participants. Furthermore, the aggregated data provided to schools via reports containing tabular and data visualisation formats, will be completely anonymised and will not allow re-identification of participants. A failsafe is built into the school feedback reports such that scores are not calculated for any output representing fewer than 10 participants, per ONS disclosure guidance. The existing data provided by participating schools and the National Pupil Database will not be part of the publicly accessible anonymised data set.
- **Pseudonymise data and apply secure key management procedures** Identifiable personal data (e.g. forename, surname) will only be kept for as long as is necessary to meet project objectives, after which it will be destroyed. The statistical modelling required to answer the project's questions will entail pseudonymised data. Once all relevant analyses are completed, anonymisation will ensue. The required keys to transform pseudonymised data into identifiable data will be stored within the encrypted VeraCrypt drives, within UoM shared folders that require access clearance and dual factor authentication.
- **Impose suitable data sharing and collaboration agreements** Data sharing/transfer agreements (DSAs/DTAs) with relevant parties (e.g. participating schools) will be set up. The research team will seek the support of the UoM contracts team and IGO office to ensure that the DSAs meet the required standard. Contracts outlying the responsibilities of non-University of Manchester team members will be implemented between the University of Manchester and the team members' respective institutions.
- **Where needed, follow University of Manchester guidelines for disposing of personal data** The research team will seek the support of UoM IT services to ensure that destruction/disposal of personal data meets the required standard.
- **Access data hosted by the University of Manchester via its secure Virtual Private Network (VPN)** All participant data stored in the UoM folders will require VPN access.
- **Access data hosted by a third-party data provider via their secure facilities** NPD data on attainment obtained via the ONS will be processed and disposed of following the requirements of the respective data providers.

10. If you are storing personal information (including contact details) will you need to keep it beyond the end of the project?

- Yes – Other

There are two reasons why the research team will be storing personal information beyond the end of the project (December 2025):

1. The research team hopes to obtain attainment data from the NPD in the Autumn of 2025. However, there are common delays to this process and it is possible that the NPD data will only be made available to the team in 2026;
2. Pending trial results, the research team may aim to procure additional funding with the purpose of establishing a longitudinal assessment of this cohort of participants;

Based on the two points mentioned above, the research team estimates keeping personal data for at least two years after the project, to enable statistical analyses involving the data described in point 1. Pending point 2, data may be kept for longer.

11. Will the participants' information (personal and/or sensitive) be shared with or accessed by anyone outside of the University of Manchester?

- Yes - Public institutions with contractual arrangements (e.g. NHS research sites)

The trial statistician, Jan Boehnke, from the University of Dundee, will conduct statistical analyses on pseudonymised data. This data will be accessed from Isilon.
Additionally, as proponents of open science and open data the research team intends to do the following:

- Make available to the general public a completely anonymised dataset which contains no identifiable personal data and for which re-identification would not be possible. The current plan is for said dataset to be housed with the Open Science Framework;
- Provide participating schools with a feedback report containing aggregated data in tabular and data visualisation formats. This will be completely anonymised and will not allow re-identification of participations. A failsafe will be built into the reports such that summary statistics are not calculated for any output representing fewer than 10 participants, per ONS disclosure guidelines.

12. If you will be sharing personal information outside of the University of Manchester will the individual or organisation you are sharing with be outside the EEA?

- No

13. Are you planning to use the personal information for future purposes such as research?

- Yes

We plan to undertake data linkage work that can maximise the potential for further insights resulting from the research. Accordingly, the data linkage strategy is clearly outlined in relevant documentation (e.g. participant information and consent sheets).

14. Will this project use innovative technologies to collect or process data?

- No

15. Who will act as the data custodian for this study, and so be responsible for the information involved?

Professor Neil Humphrey and Professor Pamela Qualter

16. Please provide the date on which this plan was last reviewed (dd/mm/yyyy).

2022-09-29

Project details

What is the purpose of your research project?

Research has shown that the teaching of social and emotional skills (social and emotional learning; SEL) improves mental health (Taylor et al., 2017) and reduces loneliness (Eccles & Qualter, 2021; Hennessey, Qualter, & Humphrey, 2021). However, most of that work is based on interventions implemented in the initial years of primary school, and there is a need to determine effects for older school aged children. 'Passport' is an SEL intervention for schools and we will determine, via a randomised controlled trial (RCT), whether its use with 9 to 11 year olds in England improves mental health and reduces loneliness. This work will advance our understanding of how SEL can be used to support children’s mental health and relationships in school settings.

Furthermore, we can use the data generated through the project for secondary analysis, addressing important knowledge gaps in our understanding of child and adolescent mental health, including, for example, longitudinal relations between bullying, loneliness, and wellbeing.

References

Eccles, A. M., & Qualter, P. (2021). Review: Alleviating loneliness in young people - a meta-analysis of interventions. Child and adolescent mental health, 26, 17–33. https://doi.org/10.1111/camh.12389

Hennessey, A., Qualter, P., & Humphrey, N. (2021). The impact of Promoting Alternative Thinking Strategies (PATHS) on loneliness in school-children: Results from a randomised controlled trial in the UK. Frontiers in Education, 6.

Taylor, R.D., Oberle, E., Durlak, J.A., et al. (2017). Promoting positive youth development through school-based social and emotional learning interventions: a meta-analysis of follow-up effects. Child Development 88, 1156–71.

What policies and guidelines on data management, data sharing, and data security are relevant to your research project?

UoM Research Data Management Policy:
https://documents.manchester.ac.uk/DocuInfo.aspx?DocID=33802%20
UK Concordat on Open Research Data:
https://www.ukri.org/about-us/policiesstandards-and-data/good-research-resource-hub/open-research/
UoM Data Protection Policy:
https://documents.manchester.ac.uk/display.aspx?DocID=14914
UoM Records Management Policy:
https://documents.manchester.ac.uk/display.aspx?DocID=14916
UoM Information Security Policy:
https://documents.manchester.ac.uk/DocuInfo.aspx?DocID=6525
UoM Information Governance Accountability and Assurance Policy:
https://documents.manchester.ac.uk/DocuInfo.aspx?DocID=8039

Responsibilities and Resources

Who will be responsible for data management?

Joao Santos will act as Data Manager and Trial manager.

What resources will you require to deliver your plan?

For working remotely and at the office, team members will need a fast, stable internet connection, VPN access to the shared network, and a connection to the office phone. A University managed desktop/laptop will suffice.

Data Collection

What data will you collect or create?

This project includes the collection of different types of data.
Data to be collected from participating Local Authorities, independent schools and schools belonging to non-participating Local Authorities:

- **Unique Pupil Number**
- **Forename and Surname**
- **School Name**
- **Sex**
- **Class name and class teacher**
- **Year group**
- **Ethnicity**
- **Free school meal eligibility**
- **Special educational needs**

Data to be collected from the National Pupil Database:

- **Attainment data**

The project includes the collection of the following pupil survey data regarding mental health and wellbeing:

- **Internalizing symptoms**

- **Emotional regulation**
- **Wellbeing**
- **Loneliness**
- **Bullying**
- **Peer support**
- **Health-related quality of life (HRQoL)**

The project includes the collection of the following survey data designed to assess intervention implementation and process evaluation:

- A 'usual SEL practice' survey to establish the counterfactual to Passport (in control schools) and the level of programme differentiation (in intervention schools).
- An implementation survey to gather data on aspects of implementation including dosage, fidelity, and adaptation.
- Brief teacher/classroom level outcome measures, including teachers' perception of SEL culture, teacher burnout, and classroom management, to improve compliance prediction.
- Qualitative data generation in 5 purposively sampled intervention schools, to expand on aspects of implementation covered in the survey and explore more nuanced considerations, such as adaptations and reasoning for these, acceptability, and reach and responsiveness, while also highlighting factors affecting implementation.

How will the data be collected or created?

With regard to **data collection**:

Existing data from participating schools will be provided via Zendto, a service used by UoM for the transfer of information. All information transferred via Zendto will be encrypted prior to sending, and will only take place after the relevant data sharing/data transfer agreements are in place.

Data from the National Pupil Database (NPD) will be accessed through a secure online environment, as facilitated by the Office for National Statistics (ONS). The research team will aim for remote access to NPD data with the ONS approved secure room located in the Ellen Wilkinson as a backup for accessing said data. The necessary steps for requesting access to said data will be followed and the relevant data sharing/data transfer agreements will be set up. The research team will lean on the expertise of the Information Governance team at the UoM.

With regard to **data creation**:

All quantitative data will be created by means of online surveys, built using Qualtrics, a GDPR compliant platform. Qualitative data will be obtained either face to face, via semi-structured interviews and focus groups, using audio recorders, or using Zoom.

Steps will be undertaken to ensure accurate data versions and versions of all ethics documentation.

Documentation and Metadata

What documentation and metadata will accompany the data?

Drawing on appropriate guidance (e.g. https://www.icpsr.umich.edu/icpsrweb/content/shared/ICPSR/faqs/what-is-acodebook.html), a codebook will be generated that clearly outlines the contents, structure, and layout of the quantitative pupil dataset. In order to be self-explanatory and unambiguous, it will delineate as follows:

- Variable name (e.g. name or number assigned to each variable)
- Variable label (e.g. brief description to identify the variable for the user)
- Question text (e.g. exact wording from survey question)
- Values (e.g. actual coded values in the data for this variable, such as 1, 2, 3, 4, 5)
- Value labels (e.g. textual descriptions of the codes, such as Excellent, Very Good, Good, Fair, Poor)
- Summary statistics where appropriate (e.g. unweighted summary statistics for quick reference)
- Missing data (e.g. the values and labels of missing data)
- Universe skip patterns where applicable (e.g. information about the population to which the variable refers, as well as the preceding and following variables).
- Notes: Additional notes that contextualize the information conveyed in a variable or relay special instructions.

With regard to the qualitative data, a starting point code book will be provided as well as specific codes developed through the coding process. The former, given its conceptual natures, will allow for broader anticipated themes and areas to be coded into based on key areas from implementation theory and literature. The latter will allow for the actual coding and how this sits within and develops those anticipated themes.

Ethics and Legal Compliance

How will you manage any ethical issues?

This project has undergone full UREC review and been approved (2022-14050-24401). Currently, the following ethical issues are considered central:
Consent.

- Parental opt-out consent and assent of young people completing the survey will be sought. Parental opt-out consent will be happen at the beginning of the study, while pupil assent will be repeated prior to each annual data point. Participating schools will use their electronic parent/carer engagement system (e.g. ParentMail or equivalent) to distribute information and consent documents. This is by far the most reliable means to ensure that parents/carers are informed about the study and have the opportunity to make an informed decision about their child's participation; indeed, schools use these systems to communicate critical information to all parents/carers.
- Participating schools will also be instructed to brief children so that they have at least 24 hours to decide whether to take part in the research. Briefing information will be included in the standardised instructions sent to schools, and will include a brief Powerpoint presentation (see Additional Documents) and short video (to be produced, but see example from one of the research teams' previous projects:
https://www.youtube.com/watch?v=lGLbu1Uvx8w).
- Parental opt-in consent and assent of pupils from five of the 'Passport to Success' intervention schools taking part in the focus groups will be sought. There will be two focus groups with the same children on two occasions (one during implementation of the intervention and one at the end of the implementation). Consent from parents for both focus groups will be sought in advance of the first one. Pupil assent will be sought on both occasions. We will provide schools with packs to be sent home to parents at least two weeks before our visit, with an information sheet, a consent form, and a demographic form, as well as an information sheet for children to look through. Parents will need to return consent forms to school in advance of our visit and we will collect and check these on the day of our visit. This approach has worked well in previous projects. We will complete assent forms at the start of each focus
group with the pupils, making sure that they understand what involvement will entail and going through consent points together.
- The research team will also ask for consent from teachers before completing their surveys and the interviews, with enough time to consider taking part. Our communications asking schools to express interest in taking part in the qualitative strand clearly state that schools should initially discuss this with Year 5 teachers that they are comfortable with this in principle, to avoid coercion at a later stage if schools sign up without warning the teacher involved in delivering the intervention. We will then contact this teacher closer to the first interview time period to provide an information sheet and organise whether they would like us to interview them in person during our visit to talk with children, or whether they would prefer to participate via Zoom. All staff members will be made aware of the potential for internal re-identification by colleagues in their schools at the recruitment stage.

Confidentiality.

- Only the research team will access identifiable participant-level school and child data, and teacher data from the interviews. Survey data will be kept in an encrypted Veracrypt container/folder (AES-256 algorithm), in a protected folder on the University's Research Data Storage system, known as Ipsilon. Accessing to the data requires 2FA to log into the university's network. The folder with the Veracrypt file can only be accessed by named users, and the container is itself protected with a long password (30+ characters). Audio recordings conducted via Zoom will follow suggested UoM policy and the recordings will be saved into the aforementioned Veracrypt container.

Coercion.

- School staff facilitating the child survey completion will be given a standardised set of instructions that reinforce each pupil's choice to participate. Those who opt not to do so (or who have been opted out by their parents/carers) will be given an alternative activity to complete, the nature of which will be decided by individual schools, but typically it is something of approximately the same estimated duration of survey completion (no more than 20 minutes for most children) that can be completed at a computer so as not to draw attention to individual pupils

Disclosures.

- The participant information sheet for children completing the survey will be shown on the first page of the survey; it will remind them that the data they provide will not be used for screening purposes (e.g. their individual responses will not be shared with school staff and/or trigger any kind of referral process). Because all data generated from the child survey are fixed-response in nature (e.g. no open field options), it will not be possible for participants to make disclosures through the survey system. However, it is possible (though highly unlikely) that children reveal information to the teacher following the completion of the survey, and school-staff will be reminded to follow their school disclosure protocol, and referral processes, in line with their safeguarding policy and procedures.
- The various teacher surveys (teacher outcomes and usual SEL practice surveys; and, for those in intervention schools, implementation surveys) also do not include any opportunities for disclosure (as above, they use fixed response categories) and the nature of the questions (which primarily focus on their social and emotional learning attitudes and practices; the only exception being a brief perceived stress measure) is such that they are highly unlikely to cause even minimal distress. However, as a safeguard, if the questions asked raise issues for the respondent, the final page of the survey suggests that

they raise that with their school's senior management team.
- The participant information sheet for pupils taking part in the focus group will be sent home with pupils two weeks before the focus group runs; researchers will also go through this with pupils at the beginning of the focus group. It is highly unlikely, but possible that children will reveal information to the person running the focus group and members of the research team will be reminded to follow the safeguarding protocol, which will include speaking with school staff and making a record for our notes. Our distress and safeguarding protocols details the steps to be taken here.

Safeguarding.

- The child survey does not include items that would reveal 'traditional' safeguarding issues (e.g. abuse, neglect, exploitation). It is also important to note that we will *not* include items relating to suicidal ideation. Inevitably though, the nature of the project means that some items will focus on thoughts, feelings, and behaviours of young people that may inadvertently be stressful or upsetting for some. However, this is very rare and likely to only cause temporary distress (Langhinrichsen-Rohling et al, 2006). Furthermore, our own research indicates that responding to mental health items can actually facilitate positive reflective processes among young people (Demkowicz et al, 2020). It is also worth noting that our previous research (including the HeadStart and BeeWell projects) has involved, cumulatively, over 100,000 children and young people being surveyed about similar issues (e.g. their mental health and wellbeing, and the factors that influence these) with no reported adverse effects/events. Nonetheless, we have a clear distress protocol as part of our standardised instructions for school staff facilitating survey completion. In addition, the survey will signpost young people to a range of sources of support that they can seek out (e.g. parent/guardian, Childline, member of school staff, see final pages of the survey) if any of the questions made them feel sad, worried, or upset. Finally, the order of in which the questionnaires are presented will be set such that items relating to mental health difficulties are 'sandwiched' by those measures that relate to more benign domains. Our distress protocol will also be used if needed for the focus group aspect of the project. Children who take part in the focus groups will also be signposted to the sources of support as they are upon completion of the survey.

Autonomy.

- Our standardised instructions to school staff and the survey assent page, and the information provided to pupils taking part in the focus groups, will both reinforce children's rights to make their own decisions about whether or not to participate, and their right to withdraw at any time. As noted above, an alternative activity will be available for those who choose not to take part in the survey completion (or who have been opted out by parents/carers). The right to withdraw will also be highlighted to the pupils in relation to the focus groups, and teachers/school staff regarding the interviews.

How will you manage copyright and Intellectual Property Rights (IPR) issues?

Some measures used in our surveys may be subject to copyright and licensing arrangements, and so the research team will ensure compliance with these.
Given the intention to make the project data publicly available, we will explore the prospect of licensing it, following relevant guidance as we do so (e.g., https://www.dcc.ac.uk/guidance/how-guides/license-research-data).

Copyright and IPR for the 'new' data generated by this project (e.g. participant and staff survey data) will rest with UoM. Copyright and IPR for the 'existing' data note rests with the Local Authority providing it.

Storage and backup

How will the data be stored and backed up?

Data are to be kept in an encrypted Veracrypt container/folder (AES-256 algorithm), in a protected folder on the university's Research Data Storage system, known as Ipsilon. Accessing the data requires 2FA to log into the university's network. The folder with the Veracrypt file can only be accessed by named users, and the container is itself protected with a long password (30+ characters, changed annually). The exception to the aforementioned storage process concerns the fully anonymised teacher interview and pupil focus groups' transcripts. Although these will be stored in the encrypted container described above, during the qualitative analysis process the research team members responsible for analysis of the transcripts will store them in the Passport to Success TEAMS channel (which is restricted to the project team members and is equivalent to a sharepoint channel). Once the qualitative analyses are complete, the transcripts will be deleted from the TEAMS channel and kept only in the encrypted drive.
The separate storage process for the teacher interview and focus groups' transcripts is due to the fact that only one person can access the shared drive at one point in time, which would then preclude simultaneous analysis of said transcripts by the different post graduate researchers. Furthermore, access to an encrypted drive may require that some team members contact IT due to technical aspects. This separate storage process allows the research team to still uphold the anonymity of the data, while preventing data analysis delay.

The encrypted container noted above is stored on an access restricted data share on the University's network storage infrastructure which is the recommended location for storing sensitive or critical University data. The storage infrastructure is hosted across two data centres (approx. 4KM apart) for resilience and disaster recovery purposes. Physical access to the data centres is strictly limited to data centre staff and a limited number of authorised IT Services staff. The data centres are protected by physical and electronic access security systems, swipe card access in and out of the data centres and CCTV coverage. The data centres are locked down out of hours and access is discouraged, but can be arranged by prior agreement with the data centre manager.

The University's IT Services utilises Legato Networker Backup domains. Supporting infrastructure comprises disk libraries and both physical and virtual tape libraries. Cross data centre backup is performed, so services hosted within data centre 1 (Joule House) are backed up to data centre 2 (Reynold House) and vice versa. Backup/recovery plans are documented as part of the service install process during the commissioning of a specific service. Each Service is responsible for its business continuity and disaster recovery plans, to which IT Services feed in its technical recovery plans. ITSD operates change and release management processes. All proposed changes to infrastructure hosted, maintained and administered by IT Services are recorded via the Request for Change (RFC) process with changes being reviewed and authorised by a Change Advisory Board (CAB).

Research data storage is not backed up to tape in the traditional fashion. However, resilience is obtained through replication and use of snapshots means that files deleted by accident or corrupted can be recovered. Hourly backups can be accessed within 24 hours and daily backups can be accessed within 35 days.

How will you manage access and security?

Only named project staff will be able to access the project data. Accessing the data requires 2FA to log into the university's network. The folder with the Veracrypt file can only be accessed by named users, and the container is itself protected with a long password (30+ characters).

Survey data are collected via GDPR compliant Qualtrics, with results being exported directly into the above mentioned encrypted drive.

Qualitative data, in the form of audio recordings, will be collected following UoM standard operating procedures, with subsequent storage also taking place in the above mentioned encrypted drive.

Selection and Preservation

Which data should be retained, shared, and/or preserved?

The research team will retain a full version of the project dataset until such time as identifying personal information is no longer required (e.g. all data linkage opportunities which require such information to facilitate matching have been exhausted and there will be no additional analyses related to the possibility of a longitudinal follow-up of this cohort). Once personal identifiable data is no longer required the research team will dispose of said information securely, with the support of UoM IT Services.

In the longer term, for a period of 20 years, an anonymised version of the dataset will be retained to maximise the potential for further insights resulting from the research.

With regard to data accessed via the Office for National Statistics, the research team will abide by the access requirements of the data providers.

What is the long-term preservation plan for the dataset?

The intention of the research team is that the publicly available versions of the qualitative and qualitative datasets will be housed in the Open Science Framework, which has funding for 50+ years of data access hosting at the time of writing.

Data Sharing

How will you share the data?

As noted above, an anonymised, publicly available version of the dataset will be housed in the Open Science Framework.

Are any restrictions on data sharing required?

The anonymised personal data set - the 'public' version of the project dataset, will be a completely anonymised version which will not allow re-identification of participants. This data set will only contain data created by the project, not existing data collected from the aforementioned parties (e.g., schools, National Pupil Database).

Aggregated data will be made available in the form of feedback reports containing summary statistics in tabular and data visualisation formats. These feedback reports will be provided to participating schools, but this too will be completely anonymised and will not allow re-identification of participations. A failsafe is built into the school feedback system such that any scores for any output representing fewer than 10 participants are not calculated.