Plan Overview

A Data Management Plan created using DMPonline

Title: Outcomes Assessment of Sensory Education in Schools - Main Study

Creator: Nicholas Wilkinson

Affiliation: University of Leeds

Funder: European Commission

Template: Horizon 2020 DMP

ORCID iD: 0000-0002-4065-4397

Project abstract:

This DMP applies to the randomised control trial phase of the OASES project. This work will assess the impact of a programme of sensory food education ("Flavour School"), on willingness-to-taste and behavioural engagement in tasting activities in children in Reception and Key Stage 1 of primary school (ages 4-7). The programme will be taught by primary school teachers, for one hour per week, as part of the normal school week.

ID: 41113

Start date: 01-09-2021

End date: 30-09-2022

Last modified: 12-07-2022

Grant number / URL: AMD-799965-2

Outcomes Assessment of Sensory Education in Schools - Main Study - Detailed DMP

1. Data summary

State the purpose of the data collection/generation

Personal data will be collected to enable contact and consent documentation with participants and their parents/caregivers (project administration), and analysis of any effects of gender and participant background (Pupil Premium eligibility).

Experimental data gathered will be used to assess the outcomes of an educational programme, 'Flavour School'. The primary outcomes measured will be 'willingness-to-taste' and 'behavioural engagement' in a food exploration and tasting activity. *We also plan to collect data on user experience and usability of the Flavour School programme via a survey. We hope to extend to further, qualitative research questions by collaborating with students on dissertation projects. Details TBC.*

Explain the relation to the objectives of the project

The project aims to assess the effectiveness of a programme of sensory food education to increase children's (aged 4-7 years) openess to trying new foods (especially vegetables and fruit), and their enjoyment in exploring such foods. The survey, and other potential qualitative research, will help us to understand the enablers and blockers to uptake, and the user experience of teachers and schools delivering the programme. This evidence will help to inform decision making at the school and policy level, regarding take up and support for sensory food education as one aspect of improving children's dietary health and wellbeing.

Specify the types and formats of data generated/collected

Personal data (collected as part of the informed consent process):
Parents/caregivers: name, email address
Participants (children aged 4-7): name, date of birth, gender, food allergy/intolerance information, eligibility for Pupil Premium (state support for disadvantaged children, paid to schools). Eating behaviour survey, filled by parents on behalf of their child.

Research data (collected online vis Bristol Online Surveys):

Eating behaviour survey, filled by parents on behalf of their child.

Research data (collected onsite in primary schools, from observation and recording of children engaged in a group food exploration and tasting activity):

Supervised self-report of items tasted during a tasting activity session.
Video recordings of individual children engaging in tasting activities. Approx 200 hours of footage.

Quantification of 'behavioural engagement' from video recordings.

Survey data - Scale/rank questionnaire with free text comments

Specify if existing data is being re-used (if any)

None

Specify the origin of the data

Personal data will be provided by parents/caregivers and schools, as part of the process for obtaining caregiver consent. Consent process will be managed via an online survey run through onlinesurveys.ac.uk. Paper versions of the consent form will also be used, administered by teachers in participating classes, to engage caregivers who don't use the online route.

Experimental data will be recorded onsite in primary schools, during a 'Flavour Explorers' activity in which 3 or 4 children choose items to taste from a selection. Number of items tasted will be via supervised self-report, recorded on a 'My Tasting Card'. Behavioural data will be recorded on video, with a separate video camera for each child doing the activity.

Participating teachers and headteachers will be asked to complete the user experience surveys, via onlinesurveys.ac.uk.

State the expected size of the data (if known)

Approximately 800 self-report 'My tasting cards'. These will also be recorded as photographs. The children will be able to keep the hard copy.

Approx 200 hours of video data.

Approx. 20 user experience surveys.

Outline the data utility: to whom will it be useful

The raw data will be useful to the research team. Processed, published data and results will be relevant to teachers, school leaders and education policy makers.

Where specific consent has been obtained, video data may also be useful for communications and promotions activities by Flavour School (a registered charity in England and Wales, which has developed the study intervention).

Methods and equipment blueprints may be of use to other researchers wishing to replicate our study, or conduct similar research.

2.1 Making data findable, including provisions for metadata [FAIR data]

Outline the discoverability of data (metadata provision)

Where data supports a publication, that data will be assigned a DOI. Data will be bi-directionally linked to publications to make it findable, and both will also be associated with the ORCID of the lead researcher (and perhaps other co-authors). Non-experimental data/tools (e.g. CAD drawings of the apparatus or software developed) will be shared via a dedicated project domain on github under Creative Commons, and linked to both the data stored with Uni. Leeds Library, and to any relevant publications. This will facilitate replication and extension of our work.

Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers?

Where data supports a publication, that data will be assigned a DOI. Data will be bi-directionally linked to publications to make it findable, and both will also be associated with the ORCID of the lead researcher (and perhaps other co-authors). Non-experimental data/tools (e.g. CAD drawings of the apparatus or software developed) will be shared via a dedicated project domain on github under Creative Commons, and linked to both the data stored with Uni. Leeds Library, and to any relevant publications. This will facilitate replication and extension of our work.

Outline naming conventions used

Naming conventions: The project file structure on the n: drive will be the standard used by the Nutritional Epidemiology Group. Beyond this, we will use AAPS Record Management Project guidelines for naming:

1. Keep file and folder names short, but meaningful.
2. Avoid unnecessary repetition and redundant words in file names and file paths.
3. Use capital letters to delimit words, not spaces.
4. When including a number in a file name always give it as a two-digit number rather than one, i.e. 01, 02 … 99, unless it is a year or another number with more than two digits.
5. If using a date in the file name always state the date ‘back to front’, and use four digit years, two digit months and two digit days: YYYYMMDD or YYYYMM or YYYY or YYYY-YYYY.
6. When including a personal name in a file name give the family name first followed by the initials.
7. Avoid using common words such as ‘draft’ or ‘letter’ at the start of file names, unless doing so will make it easier to retrieve the record.
8. Order the elements in a file name in the most appropriate way to retrieve the record.
9. The file names of records relating to recurring events should include the date and a description of the event, except where the inclusion of any of either of these elements would be incompatible with rule 2.
10. The file names of correspondence should include the name of the correspondent, an indication of the subject, the date of the correspondence and whether it is incoming or outgoing correspondence, except where the inclusion of any of these elements would be incompatible with rule 2.

11. The file name of an email attachment should include the name of the correspondent, an indication of the subject, the date of the correspondence, ‘attch’, and an indication of the number of attachments sent with the covering email, except where the inclusion of any of these elements would be incompatible with rule 2.
12. The version number of a record should be indicated in its file name by the inclusion of ‘d’ followed by the version number and, where applicable, ‘d’ indicating ‘draft version’.
13. Avoid using non-alphanumeric characters in file names.

Outline the approach towards search keyword

Keyword definition guidelines:
1. Think from the point of view of the reader. What keywords would the reader search for that would help retrieve your article?2. Keywords should ideally be phrases of 2-4 words; single word keywords are acceptable, but they may lead to many false matches.
3. Keywords should contain words and phrases that suggest what the topic is about. Also include words and phrases that are closely related to your topic.
4. Also use variants terms or phrases that readers are likely to use.
5. The full forms of shortened words or acronyms and abbreviations should be included as well.

Outline the approach for clear versioning

Use MS Word versioning tool to save versions together in one document, with a clear comment to specify what purpose the version has been saved for.
Use separate, clearly named folders for final versions of documents in a group.

Specify standards for metadata creation (if any). If there are no standards in your discipline describe what metadata will be created and how

Metadata creation: We will use the Uni. Leeds Research Data Management service standards for metadata creation.

2.2 Making data openly accessible [FAIR data]

Specify which data will be made openly available? If some data is kept closed provide rationale for doing so

All anonymised data will be open access.
All technical data (e.g. CAD drawing of equipment, software) will be freely available via github.
Video data will by default be kept private, with access only to the research team, to protect the privacy of the participants. Where specific consent has been obtained for wider sharing of video data, video data may be shared to help communicate the project outcomes and methods, and to promote take up of the Flavour School project.

Specify how the data will be made available

hosted with the Uni. Leeds Library Research Data Management service to enable open access online. Code and technical data will be hosted on a dedicated project folder on github, and shared under Creative Commons licensing (level tbc).

Specify what methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?

We wil make our analysis software/code openly available under Creative Commons, via github. The code is not yet developed.
We will provide an instruction sheet with any shared data. This will explain how to use the data and provide or link to any tools/code useful/needed to access the data.

Specify where the data and associated metadata, documentation and code are deposited

Uni. Leeds Library Research Data Management service. Github.

Specify how access will be provided in case there are any restrictions

Various levels of restricted access are provided by the Uni. Leeds Library Research Data Management service.

2.3 Making data interoperable [FAIR data]

Assess the interoperability of your data. Specify what data and metadata vocabularies, standards or methodologies you will follow to facilitate interoperability.

We will provide an instruction sheet with any shared data. This will explain how to use the data and provide or link to any tools/code useful/needed to access the data.
We'll use the DDI specifications for data types:
https://www.ddialliance.org/Specification/DDI-CV/DataType_1.0.html
Metadata should include:

- dataset title
- creator (s)
- dataset description (abstract)
- project title
- funder (s)
- grant number
- academic subject
- institutional division

The data deposit template comprises the following worksheets:

- **Dataset**: blank template to record metadata for your datasets
- **Files**: blank template to record metadata for your files
- **Dataset sample**
- **Files sample**
- **Info**: Information about how to deposit metadata
- **Controlled lists:** Lists of the controlled vocabulary for data fields
- **Subjects:** List of academic subjects

Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?

Metadata vocabularies will be the standards used by the Uni. Leeds Library Research Data Management service.

2.4 Increase data re-use (through clarifying licenses) [FAIR data]

Specify how the data will be licenced to permit the widest reuse possible

Creative Commons (level tbc)

Specify when the data will be made available for re-use. If applicable, specify why and for what period a data embargo is needed

Upon publication of associated reports/papers.

Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why

All anonymised data will be re-usable.
All methodological data will be re-usable.
Data in which participants are identifiable will only be shared with third parties where specific consent has been obtained.

Describe data quality assurance processes

We will follow the UK Data Service recommendations for ensuring data quality. As detailed here:
https://www.ukdataservice.ac.uk/manage-data/format/quality.aspx

Specify the length of time for which the data will remain re-usable

10 years

3. Allocation of resources

Estimate the costs for making your data FAIR. Describe how you intend to cover these costs

Person-hours of researcher time. Covered by the grant which funds the research.
Data hosting with Leeds University library data management service. Costs tbc. To be covered by the grant funding the research, and/or by open access publication fund.
Open access publishing may incur further costs - amounts will depend on publisher. Covered by open access publication fund.

Clearly identify responsibilities for data management in your project

Dr Nicholas Wilkinson, lead researcher

Describe costs and potential value of long term preservation

Value is primarily in replicability and assurance of result integrity. Re-use of data may also produce new knowledge.
Costs tbc.

4. Data security

Address data recovery as well as secure storage and transfer of sensitive data

Data will be stored securely in a limited access folder in the drive for the School of Food Science and Nutrition. Data will also be backed up on an external hard drive.
Some video and audio data will be shared with commercial companies who develop analysis software we plan to use. Data sharing agreements and protocols will be added to this DMP as they are completed.

5. Ethical aspects

To be covered in the context of the ethics review, ethics section of DoA and ethics deliverables. Include references and related technical aspects if not covered by the former

See ethics review board application form.

6. Other

Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)

We will follow the folder organisation standards for the Nutritional Epidemiology group for data storage and project management.