Plan Overview

A Data Management Plan created using DMPonline

Title: An Evaluation of the Training Workstream in the UK Reproducibility Network’s Open Research Programme

Creator: Danny Smith

Principal Investigator: Neil Jacobs

Data Manager: Danny Smith, Joseph Corneli

Affiliation: Royal Veterinary College

Funder: UKRI Future Leaders Fellowships

Template: UKRI Future Leaders Fellowships Template

ORCID iD: 0000-0002-8050-8175

Project abstract:

The UK Reproducibility Network (UKRN) aims to promote rigorous and transparent research practices through collaboration between researchers, institutions, funders, publishers, and others. Its flagship Open Research Programme (ORP), running from 2021 to 2027 and supported by the Research England Development Fund, aims to accelerate the uptake of high-quality open research practices in UK universities. Workstream 1 of the ORP is focussed on building institutional capacity via the development of a "train-the-trainer" model for delivering training on open research practices.

This sub-project aims to understand the impact of the ORP’s training activities, specifically, how trainers and trainees within participating institutions have engaged with the Train-the-Trainer courses and how this has impacted their open research practices and/or ability to train others in those practices.

To achieve this we will run focus groups with the Trainers’ Community of Practice (CoP) and those trained by UKRN accredited trainers, with questions structured around Normalisation Process Theory. We will also analyse data sourced from COM-B (Capability Opportunity Motivation) questionnaires, completed by both trainers and trainees.

Other administrative data will be collected, managed, and processed as part of the ORP training programme’s day to day operations, for example, details of those attending training. This data is out of scope of this more specific sub-project and is therefore not covered in this DMP.

ID: 185667

Start date: 01-09-2021

End date: 31-08-2027

Last modified: 04-11-2025

Grant number / URL: Research England Development Fund (102338-101 86440)

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

An Evaluation of the Training Workstream in the UK Reproducibility Network’s Open Research Programme - Initial DMP

Management and sharing data

How will you manage and share data collected or acquired through the proposed research?

Question not answered.

An Evaluation of the Training Workstream in the UK Reproducibility Network’s Open Research Programme - Detailed DMP

0. Proposal name

0. Enter the proposal name

An Evaluation of the Training Workstream in the UK Reproducibility Network’s Open Research Programme

1. Description of the data

1.1 Type of study

This is a Qualitative Study (with both quantitative and qualitative data) using surveys and focus groups structured around Normalisation Process Theory (NPT) and psychological models of behaviour change (COM-B) to determine the impact of the UK Reproducibility Network's Open Research Programme and its Train-the-Trainer activities.

1.2 Types of data

- Quantitative tabular data from surveys and focus groups, stored in spreadsheets
- Qualitative textual data sourced from focus groups, stored in both spreadsheets and text documents
- Research materials, such as copies of surveys, questionnaires, data coding protocols, etc.
- The above may be analysed or combined to create additional data types such as graphs or other visualisations

1.3 Format and scale of the data

During the acquisiton and analysis of data the format(s) of the quantitative tabular data will likely be .xlsx and/or .csv. Qualitative data will likely be in .docx and/or .pdf file formats. Software to be utilised will include Microsoft Office (Excel, Word, Forms, etc.) and Google Workspace (Drive, Sheets, Docs etc.).

These programmes and file formats have been chosen as they are robust, available to all members of

the project team, and are either open file formats/programmes or formats that can easily be converted to open file formats that are interoperable with numerous non-proprietary software applications. Any data arising from this project to be made openly available will be converted to open formats as necessary prior to publication. This includes any additional data (e.g. graphs) derived from our core data and/or research materials (e.g. survey templates/questionnaires).

We expect that all data/materials in all formats will require minimal storage space (less than 0.5 GB).

2. Data collection / generation

2.1 Methodologies for data collection / generation

As this project looks at a specific training programme (the UKRN ORP Train-theTrainer programme) it is not possible to use other existing data related to Open Research training or training more generally.

Survey data will be acquired via Microsoft Forms with requests for completion sent via email. A Participant Information Sheet which includes a request for confirmation of consent to participate in the study and, where relevant, for data to be shared, is included within the Microsoft Form.

The survey has been designed according to the principles of Normalisation Process Theory (NPT) and also uses the Capability-Opportunity-Motivation (COM-B) model. Responders are asked to state the extent to which they agree with six statements based on the widely used COM-B model of behaviour change, ranging from 'Strongly Disagree' to 'Strongly Agree'. They are then asked to repeat for a further 16 statements based on the Normalisation Process Theory framework, again with a six-point scale, this time ranging from 'Not at all' to 'Completely'.

A target of N=180 attendees was set. The required sample size has been calculated to be 127 to detect with a 95% confidence level and a standard deviation of 8.5. A descriptive analysis (Means, standard deviations) will be undertaken for the individual items of the COM-B questionnaire and the closed-form NPT questionnaire (per NPT construct).

Data will also be acquired via open-ended NPT questions discussed in 4-6 60-minute virtual focus groups of c.8 participants. Data obtained via the Focus Groups will be subject to a qualitative analysis based on the NPT framework and its previous implementation by Ross et al*.

*(Ross, J., Stevenson, F. A., Dack, C., Pal, K., May, C. R., Michie, S., Yardley, L., & Murray, E. (2019). Health care professionals’ views towards self-management and self-management education for people with type 2 diabetes. In BMJ Open (Vol. 9, Issue 7, p. e029961). BMJ. https://doi.org/10.1136/bmjopen-2019-029961).

2.2 Data quality and standards

We are using the COM-B framework to structure our surveys and response data, and the Normalisation Process Theory framework for our focus group questionnaires. For the focus group data, a documented process for Data Coding (the assignation of numerical values to textual inputs) will be agreed and shared across the Project Team to accompany the datasets, with Data Coding carried out by two or

more team members and subsequently peer-reviewed by other team members.

3. Data management, documentation and curation

3.1 Managing, storing and curating data

Survey response data will automatically be saved to spreadsheets stored within the University of Bristol's Microsoft cloud storage (on the Onedrive of the creator of the forms) from where it will be backed up via manual transfer to the University of Bristol Research Data Storage Facility (RDSF), which provides secure, long-term storage for research data. Copies of the survey forms will also be stored on the RDSF.

Focus Group questionnaires and any data generated during the focus groups (such as transcripts, meeting chat exports, facilitator notes, etc.) will also be added to the RDSF.

The RDSF is a set of disks and servers housed in two separate data centres on the University of Bristol campus, with regular back ups to tape storage both on and off site, managed by Bristol’s Advanced Computing Research Centre (ACRC) which has a dedicated steering group and a rigorous data storage policy (https://www.acrc.bris.ac.uk/acrc/RDSF_policy.pdf). The RDSF upholds and reinforces Bristol’s wider Information Security Policy (www.bris.ac.uk/infosec/policies/docs/isp-01.pdf). Data stored on the RDSF is only available to authorised and registered users.

3.2 Metadata standards and data documentation

Adequate consideration will be given to data curation, metadata, and documentation, with folder structure, file structure, version control, and naming conventions to be discussed and agreed upon by the project team. A document outlining these agreed conventions will be stored on the RDSF along with all other project documentation.

Survey data spreadsheets will contain a ReadMe sheet within the workbook to provide additional context and metadata (e.g. creation date, creator, source/method of collection/generation, etc.) so as to facilitate reuse. The same will apply to data collected/generated from the focus groups.

Persistent identifiers will be used wherever possible (e.g. ORCiD, ROR, DOI etc.), for both working data and data published at the end of the project.

3.3 Data preservation strategy and standards

All original research data and materials will be stored on the University of Bristol's Research Data Storage Facility (RDSF) for 10 years from the publication of findings or project end date (whichever is later).

Research data and materials will also be preserved via deposit to University of Bristol's data repository Data.Bris (https://data.bris.ac.uk/data/); see section 5. "Data sharing and access" for further details. Bristol's Research Data Service (data.bris.ac.uk) is responsible for the repository and is committed to maintaining published datasets for a minimum of twenty years.

4. Data security and confidentiality of potentially disclosive information

4.1 Formal information/data security standards

The survey has been designed in such a way that all responses are anonymous from the outset. Focus group participants' personal information (e.g. name, institution) will not be recorded. Focus group data (e.g. transcripts, comments, facilitator notes) will be screened by the Data Managers and/or Bristol's RDS Team and redacted to ensure anonymity of participants.

4.2 Main risks to data security

Data/materials may be lost or corrupted during transfer from one of the Data Manager's University of Bristol Microsoft OneDrive storage to the Research Data Storage Facility (RDSF), but the risk is minimal. Once data/materials have been transferred to the RDSF the risk of data being lost or corrupted is further diminished; RDSF is a set of disks and servers housed in two separate data centres on the University of Bristol campus, with regular back ups to tape storage both on and off site, managed by Bristol’s Advanced Computing Research Centre (ACRC) which has a dedicated steering group and a rigorous data storage policy (https://www.acrc.bris.ac.uk/acrc/RDSF_policy.pdf). Data stored on the RDSF is only available to authorised and registered users.

5. Data sharing and access

5.1 Suitability for sharing

Survey data is anonymous at inception and suitable for sharing, along with research materials such as the survey questionnaire itself. However, the wording in the Participant Information Sheets (PIS) for the surveys preclude us from publishing the results as an open dataset. Instead, we will deposit this data to the University of Bristol's data repository Data.Bris (https://data.bris.ac.uk/data/) as 'Restricted' data, available to approved researchers on request, subject to satisfactory ethical approval. This process will be managed by Bristol's Research Data Service Team.

University of Bristol's Research Data Service will review Participant Infomation Sheets/consent forms for the focus groups to ensure that these are sufficiently clear to understand and include robust data

sharing clauses.Focus group data will then be anonymised by the Data Managers, reviewed by the University of Bristol's Research Data Service, and, subject to their approval, made openly available, along with relevant research materials (e.g. questionnaires) and additional resources needed to ensure reproducibility (e.g. Data Coding guidelines). Prior to being made openly available all focus group data will be subject to a University of Bristol disclosure risk assessment (https://www.bristol.ac.uk/staff/researchers/data/publishing-research-data/disclosure-risk/), and where risks are identified measures will be implemented to mitigate them, such as aggregation of the data or changing the access level (e.g. to Restricted or Controlled rather then Open).

5.2 Discovery by potential users of the research/innovation data

Data and materials will be deposited to a University of Bristol's dedicated research data repository Data.Bris. Each data.bris deposit is accompanied by appropriate metadata and is assigned a unique Digital Object Identifier (DOI) via the DataCite scheme. All DataCite DOIs are searchable, internationally. Three access levels are available: Open, Restricted, and Controlled. The University Research Data Service (data.bris.ac.uk) is responsible for the repository and is committed to maintaining published datasets over the long-term and for a minimum of twenty years.

Any publications arising from this research will make use of Data Access Statements and the repository record for the data/materials will be subsequently updated with the PIDs of any publications or related research outputs.

5.3 Governance of access

Focus group data/materials will be made openly available in Data.Bris and licensed with a canonical and machine-readable open licence (such as athe CC-BY licence), facilitating accessibility and reuse without manual intervention.

As the survey data cannot be made openly available as the Participant Information Sheets preclude this, we will deposit this to Data.Bris with Restricted or Controlled access. Data in Data.Bris that is not openly available will be subject to the University of Bristol's access request process (https://www.bristol.ac.uk/staff/researchers/data/accessing-research-data/), which utilises an application in the repository with requests for access handled by the institution's Research Data Service (if data is set to Restricted access), or escalated to the Data Access Committee (DAC) (if data is set to Controlled access). Where a request is approved, data transfer will be managed by the Research Data Service Team.

5.4 The study team’s exclusive use of the data

Data and materials will be published no later than the date of first publication of the project's findings.

5.5 Restrictions or delays to sharing, with planned actions to limit such restrictions

Survey data will be deposited to Data.Bris with access restrictions, limiting access to authenticated researchers who have successfully applied to access the data subject to ethical approval.

Focus group data will be made openly available, albeit in a redacted form to preserve the anonymity of participants.

5.6 Regulation of responsibilities of users

Responsibility for the data will be held by the Data Managers. External users of open data (focus group data) deposited to Data.Bris will be bound by the requirements of the licence applied to the data and materials (e.g. attribution if a CC BY licence is used). External users of data in Data.Bris subject to access restrictions (survey data) are regulated via the University's data access agreeements (https://www.bristol.ac.uk/staff/researchers/data/accessing-research-data/). For restricted datasets, an institution-to-institution data access agreement will be signed, which outlines the terms and conditions of use. All users of Data.Bris are bound by the Data.Bris licence (https://data.bris.ac.uk/data/licence).

6. Responsibilities

6. Responsibilities

The Data Managers for this project are:

Danny Smith (Royal Veterinary College - https://orcid.org/0000-0002-0724-2374)

Joseph Corneli (Oxford Brookes University - https://orcid.org/0000-0003-1330-4698)

Responsibility for the University of Bristol Research Data Storage Facility (RDSF) lies with Bristol’s Advanced Computing Research Centre (ACRC).

Support and advice will also be provided by the University of Bristol's Research Data Service (data-bris@bristol.ac.uk).

7. Relevant policies

7. Relevant institutional, departmental or study policies on data sharing and data security

Policy	URL or Reference
University of Bristol Open Research Policy	https://www.bristol.ac.uk/staff/researchers/open-research/open-research-policy/
University of Bristol Policy for the use of the Research Data Storage Facility	https://www.acrc.bris.ac.uk/acrc/RDSF_policy.pdf
University of Bristol Information Security Policy (ISP-01)	https://www.bristol.ac.uk/media-library/sites/infosec/documents/ISP-01v1.2.pdf
Data.Bris Licence	https://data.bris.ac.uk/data/licence

8. Author and contact details

8. Author of this Data Management Plan (Name) and, if different to that of the Principal Investigator, their telephone & email contact details

Danny Smith
https://orcid.org/0000-0002-0724-2374
dasmith@rvc.ac.uk