Plan Overview

A Data Management Plan created using DMPonline

Title: Extending the Earcheck intervention: behavior change to prevent hearing damage, with a focus on young people with low social-economic status

Creator: Gjalt-Jorn Peters

Principal Investigator: Gjalt-Jorn Peters

Data Manager: Gjalt-Jorn Peters

Affiliation: Open University of the Netherlands

Funder: ZonMw (Nederlands)

Template: Data management ZonMw-template 2019

ORCID iD: 0000-0002-0336-9589

Project abstract:

This project aims to contribute to the prevention of hearing damage in young people with a low social economic status (SES). The existing Earcheck (“Oorcheck”), an online hearing test that detects hearing damage in people of 12-25 years will be extended from screening instrument to behavior change intervention using the Intervention Mapping protocol. Because this target population (i.e. young people of low SES) both engages in risk behaviors frequently and has a higher likelihood of exposure to loud sounds in future professions, this warrants a specific focus on this target population. Despite this urgent need for interventions for this target population as expressed by youth health care organisations (a type of organisations known as “JGZ” organisations in the Netherlands) and municipal health services (GGDs in Dutch), very few evidence based interventions are available for youth from 12 and up and none that target low SES groups. In addition, it is insufficiently clear what determinants such an intervention should focus on in this target population, let alone which behavior change principles can be expected to have an effect. This project consists of multiple studies to determine a) what measurement instruments (specifically, what questions to use in questionnaires) are suitable for young people with low SES, b) what determinants are relevant regarding hearing protection-related behaviros in young people with low SES, and c) how to optimize an intervention targeting those determinants with the most promising behavior change principles. In collaboration with the planning group, representing young people with low SES and other relevant stakeholders, this extension of the Earcheck will be developed. After extending the Earcheck, a randomized controlled trial will be employed to evaluate the intervention’s effectiveness on target population behavior, as well as process measures such as use, acceptability, and comprehension. The intervention will then be disseminated to the target population and implementers (JGZ organisations and GGDs). Part of this implementation will be obtaining accreditation from the RIVM CGL to facilitate further dissemination among health promotion professionals. Specifically, this project consists of six distinct studies. The first is a living review to create an overview of all determinant studies of hearing protection-related behavior. The second is a qualitative study to complement this with sub-determinants specific to Dutch young people (with special attention to young people

of low SES). In the third study, we will use cognitive interviews to adapt existing quantitative measurement instruments as typically used in determinant studies to work well in young people of low SES. In the fourth study, we will apply this measurement instrument to obtain current determinant structures enabling us to select the most relevant determinants for intervention. We will then develop this intervention, and pretest it in the fifth study. After using the results to optimize it, we will implement it and conduct the sixth study, a randomized controlled trial, to evaluate it. This project has four other characteristics that are important to point out. First, we will complement the trialled and trusted Intervention Mapping standard with newly developed innovations. As such, second, this project doubles as a knowledge translation effort: through this comprehensive intervention development process, these innovations will be implemented in VeiligheidNL, the prevention organisation in this project team. Third, we will adapt these instruments required for conducting determinant studies so that they become usable in young people of low SES, and we will make these instruments freely and publicly available. Fourth, we will collaborate with the Academy of Behavior Change (ABC), a Dutch foundation that was founded to promote the development of effective behavior change interventions in Dutch prevention practice, to disseminate the developed workflow and procedures to other Dutch prevention organisations. Therefore, when this project is concluded, conducting more representative determinant studies and leveraging recent innovations to support intervention development procedures has become more attainable for Dutch prevention efforts, not only at VeiligheidNL, but nationally.

ID: 54638

Last modified: 18-03-2020

Grant number / URL: 555002016

Copyright information:

Extending the Earcheck intervention: behavior change to prevent hearing damage, with a focus on young people with low social-economic status

1. General features of the project and data collection

1.1 Project leader contact details

Email: gjalt-jorn.peters@ou.nl
Phone: +31639629431
Twitter: https://twitter.com/matherion

1.2 I have composed my DMP with the assistance of a data stewardship (or management) expert. List his or her name, function, organisation/department, phone number and email address.

- The expert is connected to my department or institution (please explain his/hr expertise related to data stewardship)

The principal investigator themselves is one of the staff members in our institution with the most expertise in data management. Their first publication about the need for Open Data and Open Materials dates from 2012, is actively involved in the Open Science movement, and is very interested in the GDPR. As such, in his empirical work, he has obtained extensive experience with simultaneously working conform Open Science principles and adhering to the GDPR. In addition, he is involved in the development of several R packages to promote data sharing of qualitative and quantitative data.

1.3 In collecting data for my project, I will do the following:

- Generate new data

1.4 In my research, I will use:

- A combination of quantitative and qualitative data

1.5 I will be reusing or combining existing data, and I have the owner's permission for that.

- No, I will not be reusing or combining existing data

There is a small probability that we will decide to use data that one of the organisations involved in this project, VeiligheidNL, owns (Earcheck data). However, at this stage this is deemed unlikely.

1.6 In collecting new data, I will be collaborating with other parties.

- Yes, I will collect the new data in conjunction with other researchers or research groups
- Yes, the new data will be (partly) provided by a project partner or supplier
- Yes, we have reached agreements on the user rights of the data used in the project

We will collaborate with multiple partners in the course of this project. However, we will diligently safeguard the Open Science nature of this project, and all data will become public (anonymized where appropriate) unless there are severely prohibiting reasons.

1.7 I am a member of a consortium of 2 or more partners. Clear arrangements have been made regarding data management and intellectual property. (also consider the possible effect of changes within the consortium on issues of data management and intellectual property)

- Yes, clear arrangements have been made regarding data management and intellectual property through a consortium agreement

All data will become public (see 1.6).

1.8 I can give an estimate of the size of the data collection; specifically, the number of participants or subjects (“n=”) in the collection and its size in GB/TB

- Yes (please specify)

- Study 1: systematic review; secondary data, will all become public; no sample size known yet; full project size will be less than 100 megabyte;
- Study 2: qualitative study: primary data; raw audio data will not become public, anonymized transcripts will become public; ~60 participants; project size will be less than 100 megabyte when not counting the audio files;
- Study 3: qualitative study: primary data; most likely, only notes will be taken, and these will become public; in case audio recordings are made, these will not become public, but anonymized transcripts may become public; ~40 participants; project size will be less than 100 megabyte when not counting the audio files;
- Study 4: quantitative study: only anonymous data will be collected, and these will be made public; ~ 1600 participants; full project size will be less than 100 megabyte when not counting the audio files;
- Study 5: either a quantitative or a qualitative study: data will be treated similarly to studies 2 and 4; either ~20 or anything upwards from 150 participants (see the project proposal); full project size will in any case be less than 100 megabyte (not counting potential audio files);
- Study 6: quantitative study: some personal data will likely be collected (depending on how exactly we will design this study; it only takes place four years from now, and its design depends on discussions with partner organisations and outcomes of earlier studies) - if so, these personal data will be removed and the anonymize datasets will be made public; ~ 2222 participants; full project size will be between 100 megabytes and 1 gigabyte.

1.9 The following end products I will make available for further research and verification (please elaborate briefly)

- Audiovisual material/ Images
- Raw data
- (Several versions of) processed data
- Data documentation
- Documentation of the research process, including documentation of all participants
- Syntaxes
- Software

We will make all produced product publicly available conform the data management description in section 4.1 of the project proposal.

(Note that 'syntaxes' is not a valid term. You probably mean 'analysis scripts' - for some unfathomably reason, SPSS script files are called 'syntax files'.)

1.10 During the project, I will have access to sufficient storage capacity and sites, and a backup of my data will be available. (please elaborate briefly)

- Yes, I will make use of an external provider's services for storage and backup of my data

We will primarily use Git, GitLab, and OSF, selectively committing files using the .gitignore file to secure our participants' privacy. Secondarily, we will rely on a zero-knowledge cloud service, and tertiarily, data repo's such as e.g. DANS EASY.

2. Legislation (including privacy)

2.1 I will be doing research involving human subjects, and I am aware of and compliant with laws and regulations concerning privacy sensitive data.

- Yes, I will involve human subjects in my research. I will comply with the Algemene Verordening Gegevensbescherming (AVG)

2.2 I will be doing research involving human subjects, and I have (a form of) informed consent from the participants for collecting their data.

- Yes (please describe the form this consent takes)
- Yes, and this informed consent allows for the reuse of data (note that in the Code of Conduct for Medical Research, ‘reuse’ is also referred to as ‘further use’)

For studies involving data collection in analog settings, we will use physical informed consent forms that will be scanned, after which the scanned PDF files will be encrypted using AES-256 encryption and

the originals will be destroyed. For studies collecting data using FLOSS (specifically, LimeSurvey), we will embed the informed consent in the questionnaire, such that data provision guarantees consent provision.

2.3 I will be doing research involving human subjects, and I will protect my data against misuse.

- Yes, the data will be anonymised. I realise that this will limit the options for re-use of my data. (explain)

2.4 I will stick to the privacy regulations of my organisation

- Yes

3. Making data findable

3.1 The data collection of my project will be findable for subsequent research. E.g., on a catalogue, a web portal, or through the search enginge of the repository (note: this is key item 3, which you should report to ZonMw at the end of your project).

- Yes, it can be found through the search engine of the archive or repository in which it is stored (please specify)
- Yes, it can be found through an online (metadata) catalogue or web portal (please specify)

We will primarily use OSF to make our data findable. OSF is the primary resource used to facilitate Open Science in the Social Sciences.

3.2 I will use a metadata scheme for the description of my data collection (note: this is key item 7, which you should report to ZonMw at the end of your project).

- Yes, The metadata scheme includes metadata elements about the way my data are collected, preserved, etc (provenance) (please specify)

We will likely apply the psychDS metadata scheme, but depending on the progress of two other projects (the DCT project in Heerlen/Maastricht and the ontology project in London), we might strive to describe metadata using links to existing ontologies.

3.3 I will be using a persistent identifier as a permanent link to my data collection (note: this is key item 1, which you should report to ZonMw at the end of your project).

- Yes, I will be using the DOI code

We will use DOIs generated through the OSF.

4. Making data accessible

4.1 Once the project has ended, my data will be accessible for further research and verification.

- Yes, immediately

We will strive to appoximate Born Open Data as much as possible.

4.2 Once the project has ended, my data collection will be publicly accessible, without any restrictions (open access).

- Yes, proceed to section 5 (Making data interoperable)

4.3 I have a set of terms of use available to me, which I will use to define the requirements of access to my data collection once the project has ended (please provide a link or persistent identifier; also note that this is a key item 4, which you should report to ZonMw at the conclusion of your project).

- Not yet, my institution will draft a set of terms of use with the help of a legal advisor

I will work with my institution to draft licenses where made necessary by collaborations with external parties. Note that this is only in exceptional cases; we aim to make data public as soon as possible.
Note that although we will not impose *access* restrictions, we may impose *use* restrictions, e.g. by applying a license that forces attribution (e.g. CC-BY) or a license that promotes a pay-it-forward model of Open Science (e.g. CC-BY-NC-SA).

4.4 In the terms of use restricting access to my data, I have included at least the following:

- The sharing of data for commercial purposes, taking into account the provisions of state aid law

See 4.3

5. Making data interoperable

5.1 I will select a data format, which will allow other researchers and their computers (machine actionable) to read my data collection (note: this is key item 5, which you should report to ZonMw at the end of your project).

- Yes (please specify)

We will use plain text files as much as possible; all qualitative data will be provided in the standards set out in the Reproducible Open Coding Kit (the ROCK), and all quantitative data will be provided in .csv format, possibly complemnented by .RData files. We will not use any proprietary formats.

5.2 I will select a terminology for recording my data (e.g., code, classification, ontology) that allows my dataset to be linked or integrated with other datasets (note: this is key item 6, which you should report to ZonMw at the conclusion of your project).

- Yes, metadata standard (please specify)

This depends on the state of concurrent developments. For the variables of interest in these studies, no useful ontologies exist at the moment, but there are two projects working towards providing that level of metadata specifications; the DCT project, a collaboration between the Open University and Maastricht University that the principal investigator is involved in, and the ontology project in London. Integration with ontologies, therefore, will depend on the progress in those projects.

5.3 I will be doing research involving human subjects, and I have taken into account the reuse of data and the potential combination with other data sets when taking privacy protection measurements.

- Yes, the participants have given their permission for reuse of the data, and the data have been pseudonymised

Data will be anonymized, not pseudonymized.

6. Making data reusable

6.1 I will ensure that the data and their documentation will be of sufficient quality to allow other researchers to interpret and reuse them (in a replication package).

- In addition, I will take further quality assurance measures (please specify)
- I will document the software used in the course of the project (please specify)
- I will perform quality checks on the data to ensure that they are complete, correct and consistent (please explain)
- I will document the research process (please explain)

We provide an extensive replication package, and will use reproducible manuscripts to optimize reproducibility. R will be used for all analyses, and R Markdown for the reproducible manuscripts. We will only use FLOSS packages, document those, and implement unit tests to very data integrity in our analysis scripts. We will transparently document the decisions we take and our justifications for those decisions, as enabled by the justifier R package and the linked functionality in the ROCK.

6.2 I have a number of selection criteria, which will allow me to determine which part of the data should be preserved once the project has ended. (see also question 1.9 and 6.1)

- Yes

All data will be shared unless this jeopordizes participants' anonymity.

6.3 Once the project has ended and the data have been selected, I can make an estimate of the size of the data collection (in GB/TB) to be preserved for long-term storage or archival.

- Yes (please specify)

It is exceedingly unlikely that we will collect more than 1 gigabyte of data.

6.4 I will select an archive or repository for (certified) long-term archiving of my data collection once the project has ended. (note: this is a key item, which you should report to ZonMw at the conclusion of your project)

- Yes, and this archive has a data seal of approval (please specify the archive)
- Yes, and this archive has a different form of certification (please specify the archive and certification)

Once all data collection has been finished, we will clone the data to DANS EASY, a Dataverse repo, or another repo with a seal of approval. We will use the OSF throughout the project to enable approximating the Born Open Data standard as closely as possible.

6.5 Once the project has ended, I will ensure that all data, software codes and research materials, published or unpublished, are managed and securely stored. Please specify the period of storage.

- Yes, in accordance with other guidelines (please explain, and specify the guidelines and the number of years)

We will use guidelines that surpass existing guidelines - specifically, we require that everything created during the project be made public as soon as possible. These products will be included in the GitLab and OSFs repo's as well as in other repo's the data will be deposited to upon project completion (as provenance).

6.6 Data management costs during the project and preparations for archival can be included in the project budget. These costs are:

- Amount ..... (please elaborate)

There are no costs; the research team possesses all required skills.

6.7 The costs of archiving the data set once the project has ended are covered.

- Yes (please elaborate)

There are no costs.