Make Your Research Reproducible: Practical Guide for Researchers

Author:

Özgün Ünver, Data Steward, Research Department, Vrije Universiteit Brussel (VUB)

Contributor to section 4:

Perseverence Savieri, Consultant for Quantitative Research, SQUARE, Vrije Universiteit Brussel (VUB)

Acknowledgements:

In alphabetical order, we thank our colleagues for their feedback: Pieter De Bruyn, Thijs Devriendt, Elisa Maes, Jone Paesmans

Table of Contents

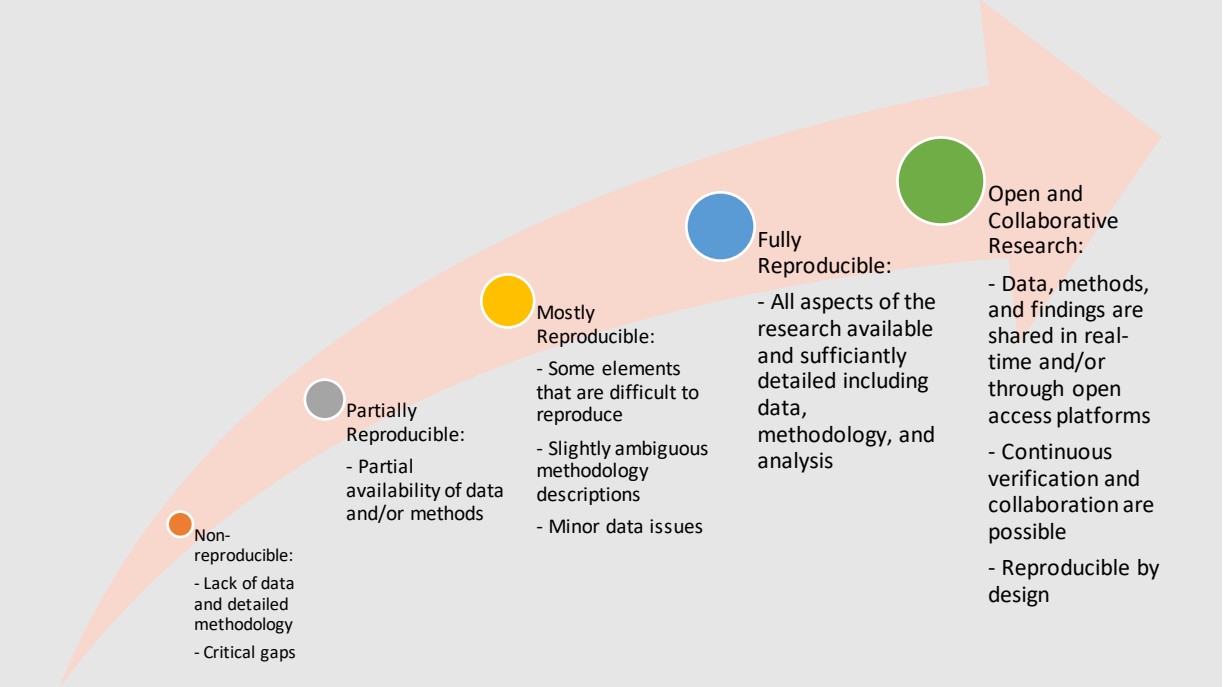
1. Introduction to Reproducibility

Research is reproducible when others can reproduce the results of a scientific study given only the original data, code, or documentation. Reproducibility is vital for the credibility of scientific research. It allows other researchers to verify findings, build upon them, and apply them in practical contexts. When a study is not reproducible, it raises questions about the credibility of scientific findings and can erode the trust in researchers publishing these findings. Therefore, ensuring reproducibility is a critical goal in the research community, involving transparent and detailed reporting of methodologies, open sharing of data and code, and rigorous peer review processes. Unsurprisingly, this concept is a cornerstone of scientific integrity and responsible research practices ensuring that research findings are reliable and can be trusted.

1.1. The Reproducibility Continuum

There are many ways and levels to make your research reproducible depending on whether you can share your original data, code, and/or (detailed and complete) documentation. We call this the "Continuum of Reproducibility". We speak of a continuum, as reproducibility is not a binary state (reproducible or not) but rather a spectrum where different levels of reproducibility and validation can be achieved. Each stage on this continuum reflects a higher degree of transparency, detail, and accessibility in research reporting. Moving research towards the more reproducible end of the spectrum is a goal for many scientific fields, as it enhances the reliability and trustworthiness of scientific findings. Here are the general stages along this continuum:

1. **Non-Reproducible**: At the lowest end of the continuum, the research cannot be reproduced at all. This may be due to missing data, lack of detailed methodology, or other critical gaps that prevent others from repeating the study.
2. **Partially Reproducible**: Here, some aspects of the research can be reproduced, but not entirely. This might be due to partial availability of data or methods, or because only some results can be reproduced with the provided information.
3. **Mostly Reproducible**: At this stage, most of the research can be reproduced by independent parties, but there might still be some minor elements that are difficult to reproduce. This could be due to slightly ambiguous methodology descriptions or minor data issues (e.g. outliers removed but not described).
4. **Fully Reproducible**: This represents the ideal situation where all aspects of the research, including data, methodology, and analysis, are available and sufficiently detailed so that independent researchers can fully reproduce the study and its results.
5. **Open and Collaborative**: Beyond mere reproducibility, this stage represents an approach where research is conducted in an open and collaborative manner from the start. It involves sharing data, methods, and findings in real-time or through open access platforms, facilitating continuous verification and collaboration. In other words, such research is reproducible by design.



1.2. The Importance of Reproducibility

"Basic science is not always about chasing the new, it's not always about chasing something groundbreaking. It's about building a house. And if the foundations of the house are rotten, if from the beginning a discipline was built on shoddy foundations, the entire enterprise can fall."
- Michael Inzlicht

Think of scientific knowledge as a house. There are endless ways of building a house, endless combinations of materials to use, and endless ways of designing and decorating it. However, no matter how beautiful a house looks from the outside, if the foundation is not robust and reliable, the house will never be a safe shelter. In science, that foundation is reproducibility.

Scientific knowledge relies on building upon each other's work and reproducibility is essential to build reliable scientific knowledge and foster trust in research outcomes. Reproducibility of their research results adds to scientists' credibility within their research community.

Research that is not reproducible is essentially a waste of resources, be it time, money, or expertise. Moreover, this waste of resources grows as other researchers build their work upon the results of an unreproducible research. Even worse is the potential for unreliable research outcomes having detrimental effect on humans and other living beings, as well as the environment. For example:

In the field of medicine, unreliable research can lead to incorrect treatments, unnecessary procedures, or missed opportunities for effective care. For example, a drug might be considered effective based on flawed studies, only to cause harm or provide no real benefit to patients. Conversely, effective treatments might be overlooked due to unreliable negative results.

In environmental science, unreliable research can lead to inadequate or misdirected regulations. This can have direct consequences on human health and safety, such as failing to recognize or appropriately address pollutants or environmental hazards.

When high-profile research is later disproven or retracted, it can erode public trust in science and experts. This scepticism can lead to resistance to important health measures (like vaccines or climate change initiatives) and contribute to the spread of misinformation.

It is difficult to talk about scientific reproducibility without referring to the **reproducibility crisis** which led to the increased awareness on this topic. The reproducibility crisis in science refers to a growing concern over the inability to reproduce the results of numerous scientific studies. This crisis originates from the observation that a significant number of scientific studies, in multiple disciplines, even those published in reputable journals, are difficult or impossible to replicate or reproduce by independent researchers. For more information on the reproducibility crisis, see Baker (2016).

1.3. Types of Reproducibility

When it comes to reproducibility, we often come across various terms that talk about the **same, similar or interrelated concepts.** This section is intended to give an overview of these concepts so that you understand what is meant when you see one of them.

1.3.1. Methods, Results, and Inferential Reproducibility

One of the distinctions you will come across is the one between *methods*, *results*, and *inferential* reproducibility proposed by Goodman et al. (2016), defined as follows:

- *Methods reproducibility* is the ability for a **different** team to obtain the same results by using the **same** data and **same** methods, tools and processes.
- *Results reproducibility* is the ability for a different team to obtain the same results by using **different** data but the **same** experimental methods, tools and procedures.
- *Inferential reproducibility* is the ability for a different team to draw the same results by using **different** data and **different** methods, tools, and processes. This one is especially interesting as not everyone would draw the same conclusions from the same results.

1.3.2. Repeatability, Replicability, and Reproducibility

The Association for Computing Machinery (ACM) has introduced another widely used distinction, this time, between the so-called *repeatability*, *replicability*, and *reproducibility* (ACM, 2020) and defined the three terms as obtaining the **same results** in the following scenarios:

- *Repeatability*: **Same** team, **same** experimental setup, **same** data, **same** measurement procedure, **same** measuring system, under the **same** operating conditions, in the **same** location on multiple trials. (For computational experiments, this means that a researcher can reliably repeat their own computation.)
- *Reproducibility*: **Different** team, **same** experimental setup, **same** data, **same** measurement procedure, **same** measuring system, under the **same** operating conditions, in the **same or a different** location on multiple trials. (For computational experiments, this means that an independent group can obtain the same result using the author's own artifacts.)
- *Replicability*: **Different** team, **different** experimental setup, **different** data, **different** measuring system, in a **different** location on multiple trials. (For computational experiments,

this means that an independent group can obtain the same result using artefacts which they develop completely independently.)

1.3.3. Reproducibility, Replicability, Robustness, and Generalisability

The Turing Way (2022) contributed to the discussion by differentiating between reproducibility, replicability, robustness, and generalisability with slightly different definitions than what was proposed before. Accordingly:

- *Reproducibility* happens when the **same** analysis steps performed on the **same** dataset consistently produces the **same** answer.
- *Replicability* happens when the **same** analysis performed on **different** datasets produces qualitatively **similar** answers.
- *Robustness* is the case when the **same** dataset is subjected to **different** analysis workflows to answer the same research question (e.g. by using another statistical software) and a qualitatively **similar or identical** answer is produced. In other words, results are robust if they are not dependent on the specificities of a programming language or analysis workflow.
- *Generalisability* is the combination of replicability and robustness, where the results are independent of a particular dataset or a particular analysis method, tool, or software.

1.3.4. Back to the Continuum

Below, you can see a table matching the different definitions from Goodman et al. (2016), ACM (2020), and the Turing Way (2022) with each other:

	Classification 1: Goodman et al. (2016)	**Classification 2: ACM (2020)**	**Classification 3: The Turing Way (2022)**
Same Team, Same Data, Same Methods		Repeatability	
Different Team, **Same Data, Same Methods**	Methods Reproducibility	Reproducibility	Reproducibility
Different Team, Different Data, **Same Methods**	Results Reproducibility	Replicability	Replicability
Different Team, **Same Data,** *Different Methods*			Robustness
Different Team, Different Data, Different Methods	Inferential Reproducibility		Generalisability

Many research fields aim for **generalisability**, which is, theoretically, the greatest form of reproducibility. In other words, the results obtained in one study should also be found by other researchers, when analysing other but similar data, and using different methods. Achieving this level is also the hardest as it is very demanding of scientific research (and researchers doing the research). At the other end of the spectrum, **repeatability** is the easiest and least demanding form of reproducibility. It is the minimum that is expected that the study results are exactly the same when the same team processes the same data using the same methods.

Note that the confusion around the definitions of these terms is still ongoing. Since we are far from a consensus regarding the definition of reproducibility (and other related terms), the easiest way to

cover all these various definitions is to simply consider **reproducibility as a continuum.** In order to avoid any confusion, we encourage you to use the term “reproducibility” and define shortly what you mean by it, every time you use it (i.e. whether you speak of the same or different scientists, data, and methods).

2. Responsible Conduct of Research

Responsible Conduct of Research (RCR) refers to conducting research while simultaneously

(1) Observing the shared values of outlines in the VUB Charter for Researchers: professionalism, responsibility, conscientiousness, accountability, and openness;
(2) Avoiding research misconduct and questionable research practices;
(3) Adhering to the rules of research ethics and scientific integrity.

As such, merely not committing to research misconduct is not sufficient to conducting research responsibly.

Reproducibility and RCR are closely interconnected, as both rely on transparency, accountability, integrity, and ethical behaviour during the entire research process. Adherence to the RCR principles enhances the likelihood of scientific reproducibility. Similarly, if your research is reproducible to the greatest degree, then you are likely to have adhered to the principles of Responsible Conduct of Research.

In the following sections, the concepts of research misconduct, questionable research practices, research ethics, and scientific integrity will be shortly discussed in relation to the reproducibility of science.

2.1. Research Misconduct

Research misconduct is traditionally defined as the presence of fabrication, falsification, or plagiarism (also known as FFP) during the research process. The ALLEA Code (2023) defines FFPs as follows:

- **Fabrication** is making up results and recording them as if they were real.

- **Falsification** is manipulating research materials, equipment or processes or changing, omitting or suppressing data or results without justification.

- **Plagiarism** is using other people's work, ideas, or words without proper attribution or without giving proper credit to the original source, thus violating the rights of the original author(s) to their intellectual output. Note that *self-plagiarism* is also a form of plagiarism where the author does not recognise the role of previously published work in their current work.

There are other forms of research misconduct than FFPs, which are called Other Unacceptable Practices (OUPs). According to the ALLEA Code (2023), OUPs are “violations of good research practices that distort the research or damage the integrity of the research process or of researchers”. There are a great number of examples to OUPs, including misrepresentation of authorship, data manipulation, dual publication, conflict of interest violations (*cf.* ALLEA Code, 2023).

2.2. Questionable Research Practices

> *"The first principle is that you must not fool yourself – and you are the easiest person to fool. So you have to be very careful about that. After you've not fooled yourself, it's easy not to fool other scientists. You just have to be honest in a conventional way after that."* — Richard P. Feynman

Questionable research practices (QRPs) are activities that are not transparent, ethical, or fair and thus threaten scientific integrity and contribute to irreproducibility in science. QRPs are, in essence, decisions made during the research process that raise questions regarding the rigor and precision of one's scientific work (AJE, 2022). QRPs are hard to identify and define – and hence, sometimes easy to get away with – making them "questionable" rather than technically illegal or explicitly prohibited.

In 2019, Dorothy Bishop has published a short but powerful opinion piece in Nature summarising the so-called "four horsemen of irreproducibility", which are HARKing, p-hacking, low power, and publication bias.

2.2.1. HARKing

What is it: Also known as "**h**ypothesizing **a**fter the **r**esults are **k**nown", *HARKing* is a common QRP. One presents a hypothesis as though this had been predicted from the start, while the truth is otherwise.

Why is it a problem (AJE, 2022):

1. Blurs the line between exploratory and confirmatory research (exploratory research is used to generate hypotheses, while confirmatory research is used to test hypotheses);
2. Leads to hypotheses being always proven (and never falsified), which prevents the research community from finding out why some hypotheses are wrong;
3. Causes ungeneralisability because hypotheses are always tailored to a specific sample;
4. Limits valuable information on contrary evidence or things that do not work.

How to counter it: Pre-registration of your research, peer-reviewed or not, is one of the best ways to prevent HARKing from happening.

2.2.2. P-hacking

What is it: Also known as "p-value hacking", *p-hacking* refers to keeping on doing (slightly) different analyses until one gets statistically significant results, when the initial analysis does not yield statistically significant results.

Why is it a problem (AJE, 2022):

1. Makes it look like effects are present, when in fact, they are not;
2. Distorts meta-analytic findings;
3. Leads to selective reporting;
4. Gives more power to artificially established significance thresholds (e.g. in Social Sciences).

How to counter it: Pre-registration of your research and deliberately explaining how you will use each of the variables is one of the best ways to prevent p-hacking from happening. Another thing to do is to be transparent and report all the results one gets, significant or not. Note that there are other statistical methods (e.g. Bayesian statistics) that might circumvent this issue (but come with their own pitfalls).

2.2.3. Low Power

What is it: This one refers to carrying out *underpowered* statistical analyses and treating the effect size as if it is sound and robust. Albeit statistically significant, *low power* in statistical research can never present correct results. Low power is usually a result of small sample size or incorrect sampling.

Why is it a problem:

1. Wastes resources (time, money, effort);
2. Makes it look like effects are present, when in fact, they are not;
3. Cannot detect the true effect, hence the results are not valid or real.

How to counter it: If you do not have the resources to design a powerful enough study, do not go ahead and work with a small sample size just to have done the work. Reusing the data previously collected by others with correct method and adequate power will help you counter this QRP.

2.2.4. Publication Bias

What is it: There are two types of *publication bias*.

1. Scientific journals are less inclined to publish the so-called 'failed' experiments or studies that show 'no (significant) effect'.
2. Researchers are less inclined to write up the so-called 'failed' experiments or studies that show 'no (significant) effect' (also called "selective reporting").

While the previous QRPs are related mostly to research employing quantitative methods, publication bias affects research employing qualitative methods as well.

Why is it a problem (AJE, 2022):

1. Fails to report, justify, and explain all conditions, experiments, and results;
2. Hides important information as to why something did not work;
3. Misleads other researchers to waste resources while exploring the areas that were already explored but not reported.

How to counter it: Pre-registration of your research and deliberately explaining how you will use each of the variables is one of the best ways to prevent publication bias. In the case of registered reports (peer-reviewed pre-registration), the journal that accepted your registered report also publishes your research results, regardless of significance. Even if you had not pre-registered your research, do try to get your results published in a peer-reviewed journal, or at least as a preprint, so that others can access your research results.

2.3. Research Ethics

Research ethics refers to the responsibility of researchers to protect individuals, animals, plants, the environment and society from possible negative aspects of research. Not harming or deceiving research participants and third parties are important to ethical behaviour of researchers. Ethical committees of the university work towards helping researchers regarding any ethical issues they may be facing in their research.

2.4. Research Integrity

Research integrity refers to the commitment of the researcher to conduct their research in a transparent, professional, reliable, and ethical manner. If “research ethics” refers to the responsibility of the researcher toward research participants and third parties, “research integrity” refers to the responsibility of the researcher toward other researchers, the scientific community, and science in general.

3. Open Science and Research Data Management

3.1. Open Science and Reproducibility

Open Science is an umbrella term used to refer to the concepts of openness, transparency, rigor, reproducibility, and accumulation of knowledge, all of which are the fundamental features of the scientific endeavour (Crüwell, et al., 2019).

Open Science is an umbrella term for many activities including (but not limited to):

- Open Access publishing
- Open Data
- Open Peer Review
- Open Notebooks
- Open Methodologies
- Open Code (a.k.a. Open Source)
- Citizen Science
- Open Educational Resources

Observing the principles of Open Science inevitably leads to improved reproducibility. Among these, Open Data, Open Methodologies, Open Notebooks, and Open Access publishing are especially important to ensure reproducibility.

Openness is useless if the data are flawed, poorly documented, not findable, or unusable. For Open Data to fulfil their mission, they should be **FAIR** (**F**indable, **A**ccessible, **I**nteroperable and **R**eusable).

3.1.1. Pre-registration & Registered Reports

Pre-registration is the publication of the rationale, hypotheses, and methodology of a study, before the study is conducted. It is typically done in a time-stamped, non-editable file, which is then deposited in a secure online archive. Pre-registration can be unreviewed or reviewed (the latter is called a “registered report”).

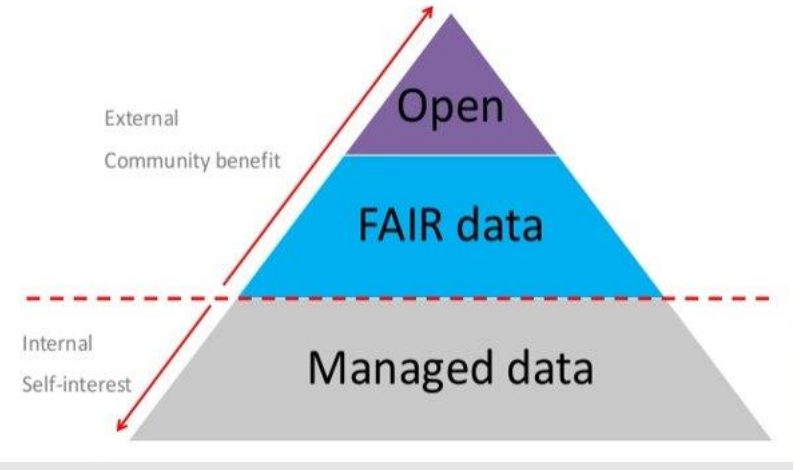
Pre-registration is one of the best ways to prevent Questionable Research Practices such as HARKing, p-hacking, publication bias, and selective reporting. Moreover, in the case of registered reports, the journal that accepts to publish the registered report also commits to publishing the research article, regardless of the outcome of the study. Hence, pre-registration tackles publication bias from all angles.

One critical question as to the suitability of pre-registration is the nature of research being exploratory versus confirmatory. In principle, one can pre-register any type of research; however, registered reports are primarily suited for confirmatory research with clear hypotheses. Regardless, pre-registration holds great potential in ensuring reproducibility.

Practical guidance to employ pre-registration is provided in the “Tools and Resources” section of this course.

3.2. Research Data Management and Reproducibility

Research Data Management (RDM) encompasses all practices and actions you perform to ensure that your research data are properly and securely collected, organized, described, stored, preserved and shared. RDM is the foundation upon which you can build FAIRness and Openness (see the image below by Jones (2019)).



Even if you do not plan to share your data openly, managing your data well will benefit YOU the most. Good RDM practices help you minimise data loss and improve the integrity and quality of your data. By employing good RDM practices, you can increase the reproducibility of your research, as well as your compliance with legislation and funder requirements.

3.2.1. Documentation & Metadata

Documentation consists of the information and guidelines you write for someone else to be able to understand your data and methods, and reproduce your study. This can be in the form of loose text or text/data fields filled in in a structured fashion.

Metadata is a specific form of documentation that are in a format that is both

1. Structured and standardised
2. Human- and machine-readable

Accurate, systematic, complete, transparent and secure documentation is key to the reproducibility of any scientific research. This needs to happen from the start to the finish of a project, including conceptualisation, planning, data collection/generation, analysis, reporting, publication, and preservation of materials.

For best results, go above and beyond the expected information and:

- Document the specifics of experimental procedures, methodologies, and protocols allowing other researchers to understand exactly why and how an experiment was conducted.
- Keep track of all the modifications and manipulations made on the data.
- Enrich your human- and machine-readable metadata by generating and filling in as many fields as possible.
- Use a metadata standard that is highly recognised in your field.

- Always attach a README.txt file to your data to make sure everything is clear.
- Employ controlled vocabularies so that everybody understands the exact same thing from the terms that you use.
- Pay attention to versioning and version control.

For more information on the basics of metadata see the following document: Understanding Metadata: A Beginners' Guide for Researchers & Data Stewards (zenodo.org)

3.2.2. Data Preservation & Sharing

Even if you document your research process perfectly, if you do not preserve your data after your research has ended, your research cannot be reproduced in its entirety. The data underlying research outputs are as valuable as the outputs themselves. Therefore, it is important to think carefully about the long-term preservation of your data, already at the start of your research.

To ensure the reproducibility of your research, consider publishing and sharing your data openly. Open Data (along with other Open Science principles) is paramount to the reproducibility of science. Trustworthy data repositories play an important role in your effort to observe Open Data. However, while preserving research data upon the completion of a project for a certain period of time is compulsory, publishing these data in a data repository is often optional.

Publishing your data with rich metadata is highly recommended, as it would allow you to get more visibility, credibility, and potential collaborators as a researcher. Moreover, it helps you get more citations both for your publications (since your results are openly backed up by data) and your data (when they are reused by other researchers). Some funders even make open data publishing mandatory (following the principle of "as open as possible, as closed as necessary"). Therefore, please explore data repositories where you can deposit your data for public access.

There are other ways of sharing data during and after research than publishing data in a repository. For instance, you can give access to the data storage location, send certain parts of data via a secure file transfer medium, or write a data paper. Note that not all data can be openly shared. If you are working with the following types of data, you need legal and ethical clearance before you share research data in any way:

- Personal data
- Third party data
- Medical/health data
- Military/dual use data
- Valorisable/commercial data
- Intellectual property data (e.g. copyright, database right)

> *"We will have achieved reproducibility when we no longer debate it."* (Leonelli & Lewandowsky, 2022)

4. Reproducibility in QUANtitative Research

Quantitative research involves collecting and analysing numerical data to understand phenomena or answer research questions. It uses statistical methods to collect and analyse data, and to draw conclusions based on the results. The data can be collected through various methods such as surveys, questionnaires, experiments, and observation. To make quantitative research reproducible, one can follow the steps outlined below at each stage of a research project.

4.1. Before Research

- Clearly define the **research question and hypothesis**: A clearly defined research question and hypothesis are essential for ensuring that the study is focused and reproducible.
- Select appropriate **research methods and data collection techniques**: The choice of research methods and data collection techniques should be based on the research question and should be appropriate for the type of data being collected.
- Develop a **detailed and rigorous study design**: A well-designed study includes a clear plan for data collection, analysis, and interpretation, and takes into account potential sources of bias and confounding factors.
- **Pre-register your research**, especially if you are conducting confirmatory research.

4.2. During Data Collection

- Collect **high-quality data**: High-quality data is accurate, reliable, and *representative* of the population being studied. Think about correct sampling method, calibrated equipment, pilot testing, training interviewers, data cleaning, etc.
- Use **standardized data collection methods**: Standardized methods, such as predefined survey questions or laboratory protocols, ensure consistency and comparability of data across studies.
- **Document all data collection procedures and decisions** as outlined in Section 3: Detailed documentation of data collection procedures and decisions made during data collection helps other researchers to replicate the study.
- If you **reuse data** that were collected in the context of a previous research project, make sure that those data are of high quality. At least be aware of any (potential) shortcomings. For more information, check out our e-learning course called "Discover and Reuse Research Data".

4.3. During Data Analysis

4.3.1. Transparency in Data Analysis

- Use appropriate **statistical methods**: The choice of statistical methods should be based on the research question and the type of data being analysed.

- Provide transparency in the **data analysis process** by detailing the statistical tests employed. Some examples are any assumptions made and adjustments for covariates (Alston & Rick, 2021).
- Clearly **outline** the tools, software, dependencies, and algorithms used, along with their versions, settings, and parameters.
- Whenever feasible, conduct data manipulation and analysis using **coding scripts** instead of the point-and-click system built in some of these tools (such as in SPSS) to provide a record of all analytical activities and enable the repeatability of each step.
- Include detailed **comments and documentation** in your code to explain each step of the analysis. This helps other researchers to understand your methodology.
- Consider using **containerisation** tools to encapsulate your entire analysis environment, including dependencies and software versions. This is explained in detail in Section 6.6.

4.3.2. Sensitivity Analyses

Sensitivity analysis is a technique used to assess how sensitive the results of a study are to changes in the assumptions, data, or methods used in the analysis (Saltelli, 2002). It helps to identify the most critical components of the analysis and to understand the impact of uncertainties on the results. Sensitivity analysis is an essential aspect of reproducible quantitative research, as it contributes to the reliability and validity of the findings.

Steps a researcher can take to perform a sensitivity analysis:

1. Identify the key variables or assumptions that could impact the results of the analysis.
2. Vary each of these variables or assumptions within a reasonable range.
3. Conduct the analysis for each variation and compare the results.
4. Identify which variables or assumptions have the greatest impact on the results and assess the implications for the interpretation of the findings.

4.4. During Reporting

- Follow **reporting guidelines**: Reporting guidelines, such as CONSORT (randomised controlled trials), STROBE (epidemiological studies), and PRISMA (systematic reviews and meta-analyses), help to ensure that all relevant information is included in the report.
- Provide a **detailed and transparent description** of the study: A detailed and transparent description of the study, including the research question, study design, data collection and analysis methods, and results, helps other researchers to understand and replicate the study.
- Include a **discussion of limitations**: A discussion of limitations, such as potential sources of bias or limitations of the study design, helps other researchers to assess the validity of the results and to design future studies.
- To obtain reproducible reports (in output formats such as HTML, PDF, or Word), use tools like R Markdown, Quarto or Jupyter Notebooks that **combine your code, analysis, and results**. This ensures that the entire workflow is documented in one file facilitating

collaboration between researchers and allowing other researchers to verify and replicate the analysis.

5. Reproducibility in QUALitative Research

This section elaborates on reproducibility for research that relies on qualitative research methodologies such as ethnographic studies, case studies, phenomenological studies, grounded theory studies, etc.

Note that, similar to QUANtitative research, the reproducibility of QUALitative research relies heavily on documentation. Good documentation for qualitative research entails maintaining comprehensive research notes, logs, and methodological decisions. In order to maintain transparency, everything should be meticulously recorded and transcribed; and each decision should be well-motivated.

Data sharing may not be straightforward when it comes to qualitative data since this type of research may include sensitive or confidential information. If legal, ethical and feasible, sharing (anonymised) qualitative data contributes greatly to reproducibility. However, anonymisation while keeping valuable data is close to impossible in qualitative research. Remember that, where data sharing is not possible, it could still be possible to include supplementary materials or a detailed documentation regarding the data and research. Note that these supplementary materials must be anonymous to be shared openly.

5.1. Transparency in Data Collection and Analysis

5.1.1. Coding

When employing content analysis, thematic analysis, or other coding strategies, provide a detailed account of how codes were derived, applied, and refined. Clearly document and communicate the coding process, including the decisions made and the rationale behind those decisions.

Here are some pointers:

- Create a codebook where each code is defined and described
- Describe how the codes were generated, applied, and revised (e.g., deductive, inductive)
- Document if/how your codebook evolves over time
- Document the rationale behind coding decisions, both in the beginning and if/when they evolve over time
- If one than one person does the coding, document if/how a consensus is reached with regard to coding decisions. Include details on how the coders are trained and who oversees the alignment between coders or researchers.

5.1.2. Analytical Decision-Making

It is paramount to document the rationale behind interpretive choices and analytical paths one took, including why certain data were emphasized over others.

Here are some pointers:

- Clearly describe the theoretical or methodological framework guiding your analysis (e.g., grounded theory, thematic analysis, narrative analysis, discourse analysis)

- Provide a detailed account of your data analysis process – including how you approached the data initially, how categories or themes were developed, and how these evolved over time
- Discuss any quality checks done to ensure the accuracy and reliability of your analysis (e.g., coding reliability checks, peer debriefing)
- Describe how you used triangulation (data, methodological, investigator, theory) to validate and enrich your findings
- Document key analytical decisions – including why certain data were interpreted in a certain way or why some themes were prioritised or merged
- Report on how your perspectives, biases, and interactions with the participants or data might have influenced analytical decisions (reflexivity during analysis). Include reflections on how alternative interpretations were considered or ruled out.

5.1.3. Use of Software

Just like every other process in a research study, the use of software as well as the decisions linked to it should be documented. If you use a qualitative data analysis (QDA) software, please pay attention to the following:

- Provide information on the version and, if applicable, settings
- Describe how and why the QDA software was employed (e.g., coding, query functions, plugins, additional tools)
- Clearly document any process that was performed outside of the QDA software – including how manual analyses were conducted and integrated with software-supported findings
- Consider employing containerisation, which means 'freezing' the software version that was used in time. This will ensure that the files used/produced by the software can still be opened even if the latest version of the same software does not support an older version that you once used.

5.2. Reflective Reporting of Findings

5.2.1. Contextualization

Sufficiently contextualizing your research is one of the most important things you can do to ensure transparency and reproducibility when you are reporting your findings. This requires both clarity and detail in reporting findings, including the use of rich descriptions, appendices, and supplementary materials.

Here are some pointers:

- Offer a comprehensive background of the research context, research setting, participants, socio-cultural factors, historical factors, environmental factors, data itself, etc.
- If applicable, offer a "rich-thick description" by both deepening and broadening your contextualisation
- Include representative data excerpts to illustrate how conclusions were draw, to allow readers to see the link between data, analysis, and findings
- Discuss the limitations of your analytical approach and the impact these may have on the findings. Acknowledge any aspects of the data that were not explored and why.

- Consider using relevant reporting guidelines to ensure that all relevant information is included in the report (e.g., COREQ and SRQR for qualitative studies).

5.2.2. Triangulation

Triangulation is a research strategy that involves the use of multiple datasets, methods, theories, and/or investigators to address a research question. It is mainly used in qualitative research but can also be applied in quantitative research. The purpose of triangulation is to enhance the validity and credibility of research findings by developing a comprehensive understanding of phenomena. Triangulation is important in qualitative research because qualitative research is inherently interpretive in nature, and by using multiple sources or methods, one can cross-check and validate their findings, reducing the impact of individual biases and increasing the rigor of the research.

Triangulation can take various forms, including data triangulation (using multiple data sources), methodological triangulation (using multiple methods), theoretical triangulation (using multiple theories), and investigator triangulation (involving multiple researchers). It aims to provide a more complete and nuanced perspective on the research topic, contributing to the reliability and validity of the research findings.

Here are some pointers:

- Gather data from multiple sources or at different times to compare and cross-validate findings (e.g., combine interviews with observations and document analysis) (Data Triangulation)
- Use multiple qualitative methods (e.g., interviews, focus groups, ethnography) or combining qualitative and quantitative methods to study the same phenomenon (Methodological Triangulation)
- Incorporate multiple perspectives, theories, or conceptual frameworks to interpret the data to see and present your data through various lenses and explore alternative explanations (Theoretical Triangulation)
- Engage multiple researchers/investigators in the data collection and analysis process to see if different researchers observe or interpret the data differently, which can help in identifying biases and ensuring a more objective analysis (Investigator Triangulation)
- Clearly describe how each method, source, theory, or investigator contributed to the research process, including the rationale for choosing specific triangulation strategies and how they complement each other
- When presenting your findings, explicitly show how triangulation has informed your analysis and conclusions (e.g., convergences and divergences that result from triangulation)
- Discuss the impact of triangulation on your research findings – including how different methods, sources, theories, or investigators corroborated or challenged each other, and how this affected the outcomes
- Critically assess the limitations of triangulation (e.g., were there complex contradictions?)

5.2.3. Reflexivity

Reflexivity in qualitative research refers to the process by which researchers critically examine and reflect upon their own contributions to the research process and outcomes. This introspection

encompasses the researcher's biases, values, beliefs, judgments, assumptions, and experiences, acknowledging how these elements influence all stages of the research, from the formulation of the research questions to data collection, analysis, and interpretation of findings. Reflexivity is vital for enhancing the credibility, rigour, validity, and authenticity of qualitative studies.

Reflexivity is not just a methodological tool but a fundamental aspect of conducting ethical and responsible research. By engaging in reflexive practices, researchers can strive for greater objectivity, acknowledging that complete objectivity is unattainable due to the inherently subjective nature of qualitative research. Through reflexivity, researchers aim to make the research process as transparent as possible, allowing readers to critically assess the influence of the researcher's positionality on the research findings.

Here are some pointers:

- Recognise your own perspectives, biases, and how your identity (e.g., cultural background, gender, profession) influences the research process
- Reflect on the power relations and dynamics between you (researcher) and participants, and how these dynamics affect the research process and data
- Continuously question and evaluate the choices made throughout the research, including methodological decisions and interpretations of data
- Be open and transparent about your position and how it influences all aspects of the research
- Employ several strategies to maintain reflexivity (e.g., reflective journaling, bracketing, seeking feedback, acknowledging researcher bias)

6. Tools and Resources

6.1. Data Management Plan

Well-managed data are the foundational bricks of reproducible research. Imagine doing everything right in your research, but not being able to show anything for it because the data underlying your results are a mess. Would your research still be considered reproducible? Of course, not!

Good data management begins even before you collect or acquire any data. Hence, your Data Management Plan (DMP) is your initial commitment to making your research reproducible. In order to write a DMP, you will be forced to think about what your data are, why you use these data, how you use these data, where the data would come from, when these data would be collected and processed, where the data are stored during/after the research, and if/how they will be shared.

6.2. Pre-registration

One of the best things you can do to make your research reproducible is to pre-register your work. This will help you keep yourself in check as well as prove to the outside world that you have a study that was designed in a certain way and kept within those lines. Even though you can also pre-register exploratory research, do no miss out on the chance to publish a peer-reviewed registered report especially if you are conducting confirmatory research. In some fields (such as clinical trials), publishing a pre-registration protocol is mandatory.

Typically, the following are covered in a pre-registration protocol or registered report:

- research question(s)
- hypotheses
- dependent and independent variables that will be collected
- study design
- planned methodology
- sample size and the method to determine the sample size
- method of data analysis
- outlier and data exclusion criteria

This plan is then timestamped and deposited into a registry. This does not necessarily mean that the plan must be publicly available or that there is no flexibility in the analysis of the data after these have been obtained. However, any deviations from the pre-registration should be made explicit and justified.

To find your way in where and how to pre-register your research, follow our guidelines: Open Science tools: Pre-registration, Registered Reports and Preprints (zenodo.org)

6.3. Electronic Lab Notebook

You can use an Electronic Lab Notebook (ELN) to document your experiments, research and procedures as well as share them between your colleagues during your research. Think of it like extra help with documenting your research. Using an ELN software means you no longer will have to

decipher the illegible scribbles from your colleague (or yourself), nor have to spend hours trying to find "that one analysis you did three years ago". Gone are the days when only lab scientists would use a lab notebook. Regardless of your discipline, you should definitely look into the possibility of documenting your research using the ELN.

6.4. README File

All types of research data can benefit from the addition of a README.txt file, where important details of the research can be documented in a free-form text format. A README file is a must-have documentation strategy because it can be fully tailored to your research and your data.

A README file can contain all kinds of information such as author details, important dates during the study, links to research output, list of data sources, relationship between datasets and/or files, recommended way to cite the data/ study, environmental/experimental conditions, software/hardware used, information on missing data, abbreviations, etc.

You can find several free README file generators online. Most of these are geared towards people who code. Here are several templates that you can adapt:

- Guide to writing "readme" style metadata – Cornell Data Services
- DataverseNO README File Template - General (zenodo.org)
- readme.so

6.5. Version Control

Version control refers to the systematic process of tracking, managing, and documenting changes to research data, code, or outputs over the course of the project. Ideally, a version control system should be agreed on and adopted by all collaborators at the beginning of the project. Version control can be carried out in several ways:

6.5.1. Document Versioning

Document versioning (also known as "functional versioning") is a simple version control can be done by simply keeping track of the changes across different versions in a version table and changing the file names as one goes along. Each version can be named and numbered (e.g. "filename_v1.3.txt", "filename_20200806_ou_v2.1.doc"). It is customary to use ordinal numbers (v1.0, v2.0, v3.0, …) in case of major revisions and decimals (v0.3, v1.1, v2.5, …) in case of minor changes.*

* If you know that you would create a version 10 or higher, it is advised to add a zero before the version number (e.g., v01.06, v14.11, …).

Status	Version	Author	Date	Changes
Draft	v0.1	Name Surname Email	20200115	First draft
Draft	V2.0	Name Surname Email	20200120	Inserted images and tables
Final	V1.0	Name Surname Email	20200130	Inserted references and typos correction
Approved	V1.0	Name Surname Email	20200205	Approved by XX

Source: Example of a version control table by Elixir Belgium (https://rdm.elixir-belgium.org/version_control)

6.5.2. Code and Data Versioning

Another option is to do advanced version control by using specific version control software such as Git and Subversion. These are used most in software development or in "data-heavy" research.

Both Git and Subversion help to:

- Keep a record of all changes made to the code and data.
- Manage collaboration between multiple researchers.
- Allow researchers to explore different analysis approaches without affecting the main analysis.
- Facilitate the replication of the analysis by other researchers.

6.6. Containerisation

If you are working with certain software, you know that they are periodically updated. A common challenge is for the code used "several updates ago" becoming unusable in the current version of the software. This happens, for instance, for some R packages where the package is not developed or maintained anymore where your code used that package becomes obsolete.* Earlier versions of some operating systems are another example.

This is where "containerisation" comes in. Basically, you encapsulate a certain computational environment, share or publish this so-called container (or 'a time capsule'), and another researcher (or your future self) can perfectly recreate the computational environment that you used in your research, regardless of the updates.

Some tools you can use to be able to do this are Docker, renv, liftr, and Holepunch.

* Note that a lot of previous versions of R packages can still be found online: The Comprehensive R Archive Network (r-project.org)

6.7. Reporting Guidelines

Reporting guidelines help to ensure that all relevant information is included in the report. You can find reporting guidelines for main study types on the Enhancing the QUAlity and Transparency Of health Research website. Below are some examples:

- CONSORT (randomised trials)
- STROBE (observational studies)
- PRISMA (systematic reviews)
- SPIRIT (study protocols)
- SRQR (qualitative research)
- COREQ (qualitative research)
- CHEERS (economic evaluations)
- ... (cf. equator)

6.8. Your Reproducibility Plan

The previous sections have already offered a plethora of tools and techniques for you to implement with a view to making your research reproducible. Now it is your turn to explore these tools and techniques better and decide which ones you will take on board throughout your research.

Regardless of the phase of your research – start, middle, or end – you can always do at least one more thing to make your research more reproducible.

Your challenge, if you choose to accept it, is to create a reproducibility plan with some (or all) of the following elements:

- Data Management Plan
- Pre-registration
- Triangulation
- Reflexivity
- Documentation and Metadata
    - Electronic Lab Notebook
    - README file
    - Metadata standard
    - Version control
    - Reporting guidelines
- Data Preservation
- Data Sharing
- Containerisation

7. Resources for Further Learning

- Belgian Reproducibility Network (RN-BE)
- FLAMES
- ReproducibiliTEA

8. References

Aarts, A.A., et al. (2015). Estimating the reproducibility of psychological science. Science. 349, (6251), 943-950. https://www.science.org/doi/10.1126/science.aac4716.

AJE. (2022). Questionable Research Practices (and research misbehaviors). Available online at: https://www.aje.com/arc/questionable-research-practices/. Accessed on December 6, 2023.

ALLEA Code. (2023). The European Code of Conduct for Research Integrity. Available online at: https://allea.org/code-of-conduct/. Accessed on December 7, 2023.

Alston, J.M. & Rick, J.A. (2021). A Beginner's Guide to Conducting Reproducible Research. The Bulletin of the Ecological Society, 102(2). https://doi.org/10.1002/bes2.1801

Association for Computing Machinery (2020). Artifact Review and Badging. Available online at: https://www.acm.org/publications/policies/artifact-review-and-badging-current. Accessed on November 8, 2023.

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. Nature 533, 452–454. https://doi.org/10.1038/533452a.

Bishop, D. (2019). Rein in the four horsemen of irreproducibility. Nature 568: 435. https://doi.org/10.1038/d41586-019-01307-2.

Charter for Researchers. (2019). Vrije Universiteit Brussel. Available online at: https://www.vub.be/sites/default/files/2023-12/2023_Reglementen_Centraal_Doctoraatsreglement_Bijlage_2_Charter_voor_onderzoeker_ENG.pdf. Accessed on December 20, 2023.

Crüwell, S., et al. (2019). Seven easy steps to open science: An annotated reading list. Zeitschrift für Psychologie, 227(4), 237–248. https://doi.org/10.1027/2151-2604/a000387

Gomes D.G.E., et al. (2022). Why don't we share data and code? Perceived barriers and benefits to public archiving practices. Proc. R. Soc. B289: 20221113. https://doi.org/10.1098/rspb.2022.1113.

Goodman, S.N., Fanelli, D., and Ioannidis, J.P.A. (2016). What does research reproducibility mean? Sci. Transl. Med. 8:341ps12. https://doi.org/10.1126/scitranslmed.aaf5027.

Jones, S. (2019). 'What it means to be FAIR'. Available online at: https://www.slideshare.net/sjdcc/what-it-means-to-be-fair. Accessed on December 8, 2023.

Leonelli, S., Lewandowsky, S. (2022). The reproducibility of research in Flanders: Fact finding and recommendations. KVAB Thinkers' report, KVAB Standpunt 81. Available online at: https://kvab.be/sites/default/rest/blobs/3679/tw_reproducability.pdf. Accessed on December 19, 2023.

Munafò, M., Nosek, B., Bishop, D., et al. (2017). A manifesto for reproducible science. Nat Hum Behav 1, 0021. https://doi.org/10.1038/s41562-016-0021.

Paesmans, J. (2023). Open Science tools: Pre-registration, Registered Reports and Preprints. Zenodo. https://doi.org/10.5281/zenodo.8088845.

Peikert, A., & Brandmaier, A.M. (2021). A Reproducible Data Analysis Workflow. Quantitative and Computational Methods in Behavioral Sciences, 1, Article e3763. https://doi.org/10.5964/qcmb.3763.

Saltelli, A. (2002). Sensitivity Analysis for Importance Assessment. Risk Analysis, 22(3). https://doi.org/10.1111/0272-4332.00040

The Turing Way Community. (2022). The Turing Way: A handbook for reproducible, ethical and collaborative research. Available online at: https://the-turing-way.netlify.app/reproducible-research/reproducible-research. Accessed on November 8, 2023.

Ünver, Ö. (2022). Understanding Metadata: A Beginners' Guide for Researchers & Data Stewards. Zenodo. https://doi.org/10.5281/zenodo.7124032.

Van Wettere, N. (2020). Reproducibility (course material). Zenodo. https://doi.org/10.5281/zenodo.4392900 .