

MUNI
ARTS

Good Research Practices

by A.

Outline

- Before...
- Replication crisis
- Science and degrees of freedom
- ...After
 - Replications
 - Transparency
 - Open science
 - Effect sizes
 - Pre-registration of studies
- Your decision!
- Building brands



Born: June 10, 1938, in
Denver, Colorado

PhD: in social psychology
from the University of
Michigan, 1964

Status: started his
academic career as a
respected mainstream social
psychologist

Before...

- Publication bias, "file-drawer effect"
- "Publish or perish"
- Novelty
- Impact

...and doubts



Born: June 10, 1938, in
Denver, Colorado

PhD: in social psychology
from the University of
Michigan, 1964

Status: started his
academic career as a
respected mainstream social
psychologist

Which Article Should You Write?

There are two possible articles you can write: (a) the article you planned to write when you designed your study or (b) the article that makes the most sense now that you have seen the results. They are rarely the same, and the correct answer is (b).

[D.Bem \(2003\) - Writing the Empirical Journal Article](#)

Daryl Bem

- Self-perception theory
- Exotic becomes erotic theory
- Precognition: Feeling the future



Feeling the Future: Experimental Evidence for Anomalous Retroactive Influences on Cognition and Affect (2011)

- Uni: Cornell University
- Journal: Journal of Personality and Social Psychology
- 8-year study
- 9 experiments
- Over 1,000 participants
- Finding: people can predict future





Born: October 19, 1966, in Oegstgeest, Netherlands

PhD: in social psychology from the University of Amsterdam in 1997

Status: a prolific and talented scientist but one of the most significant cases of academic fraud in modern psychology

Diederik Stapel

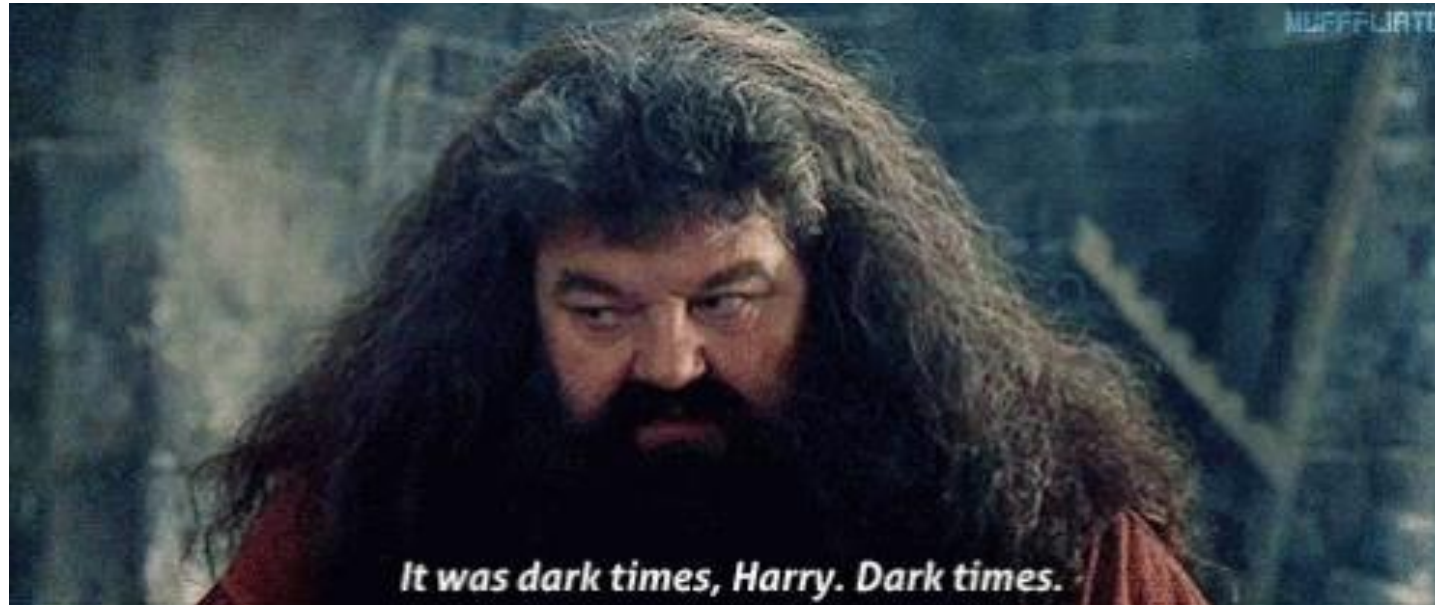
- fabricated data for at least 55 papers
- exposed by a group of three junior researchers in 2011



Diederik Stapel



Replication crisis



Replication crisis

- The Marshmallow Test (Walter Mischel, 1972)
 - much weaker relationship
- The Stanford Prison Experiment (Philip Zimbardo, 1971)
 - methodological flaws
- Power Posing (Amy Cuddy, 2010)
 - non-replicable results regarding hormonal changes
- Ego Depletion (Roy Baumeister, 1998)
 - overstated effect
- Social Priming (John Bargh, 1996)
 - non-replicable
- The Facial Feedback Hypothesis (Strack, 1988)
 - non-replicable
- Priming Intelligence with Stereotypes (Claude Steele, 1995)
 - mixed evidence



Replication crisis in science

Science and degrees of freedom

Statistics

- Data
 - 1, 2, 3, 4, 5
- Mean = parameter
 - $(1+2+3+4+5)/5 = 3$
- Df = N – 1
 - $5-1 = 4$
- Why?
 - 4 numbers can vary independently, but...
 - there is just one option for the 5th number to get the same mean value
 - $2 + 2 + 3 + 3 + 5$...mean = 3
 - $3 + 3 + 3 + 3 + 3$...mean = 3

Research

- Research goal
- Multiple ways of reaching it

Science and degrees of freedom



Good research practices (GRPs)



Questionable research practices (QRPs)



Bad research practices

Science and degrees of freedom

Preregistrations
Appropriate N
Open science
Replications
Transparency

P-hacking
HARKing
Cherry-picking
Data-fishing
Salami publication

Fabrication of data
Falsification of data
Plagiarism
Ghostwriting
Conflict of interest

Good research
practices (GRPs)

Questionable research
practices (QRPs)

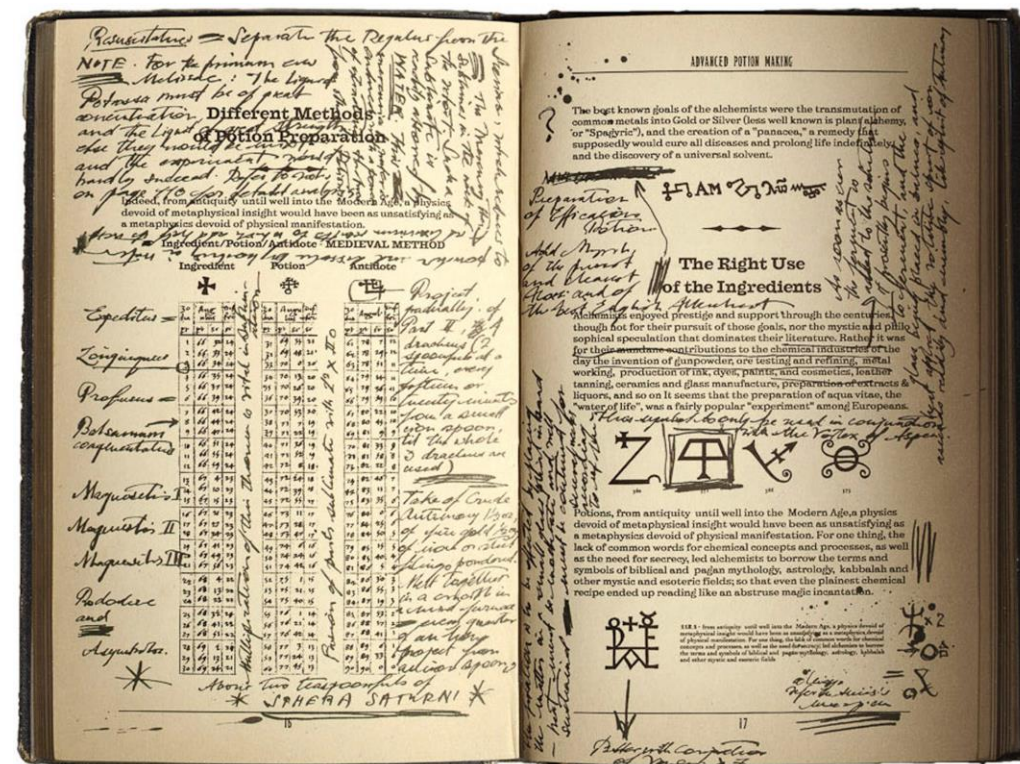
Bad research
practices

After...



Replications

- Direct / exact replication
- Systematic replication / replication-plus-extension
- Conceptual replication



Transparency

- What was done?
- How was it done?
- Is that all that has been done?
- Has anything been modified? Preregistration?
- What are the limits?
- Were there any problems?
- Has anything been done additionally?

(a) Participants

During two waves of data collection, we recruited a total of 2228 participants from 15 societies (1126 females; M age = 37.0, $s.d.$ = 14.8). Specifically, during Wave I, we recruited 591 participants who played the DISTANT RAGs and reported the results of this data collection in several publications [8,22]; however, 208 of those participants were contacted again during Wave II to collect the OUTGROUP RAGs. For Wave II, 1637 new participants were recruited, playing either the DISTANT and OUTGROUP RAGs or the DISTANT and OUTGROUP DGs (153 participants played both RAGs and DGs). Here, we collapsed both Wave I and Wave II samples to provide robust tests of our hypotheses.

We excluded all participants from our analyses whose allocations did not sum to 30 for a particular RAG or 10 for a particular DG. Specifically, we excluded 30 participants from at least one RAG and 33 from at least one DG. Furthermore, we excluded 22 participants who misunderstood the procedure or did not correctly follow procedural steps. **At one site, two research assistants counterfeited data, thus all the RAG and DG data collected by these assistants were removed (72 participants).** The number of participants in each analysis is displayed under specific models. While tables in the main text report only full models (these are missing three sites due to missing some of the covariates), reduced models including all sites can be found in the electronic supplementary material, section S3. Our protocols were approved by the University of British Columbia's Behavioural Research Ethics Board (BREB) and by the equivalent at each individual researcher's home university. All subjects provided an informed verbal consent for participation before the experiment.

Open science

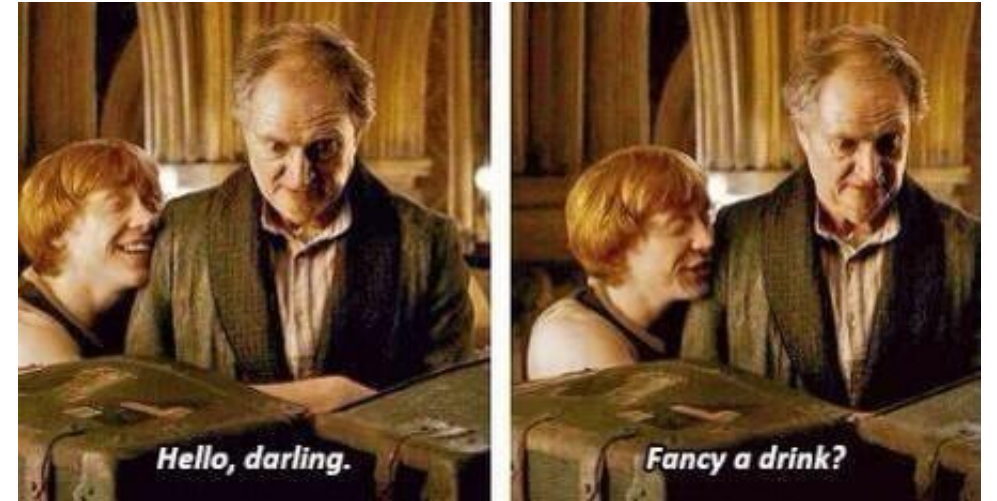
- Open access (paid)
- Sharing data, code, materials...
- Transparency
- Free software (R, JASP, jamovi)

- Knowledge should be free and available to everyone
 - Sci-Hub, LibGen



Effect sizes

- Sufficient sample size
- Statistical vs. practical significance



Effect sizes

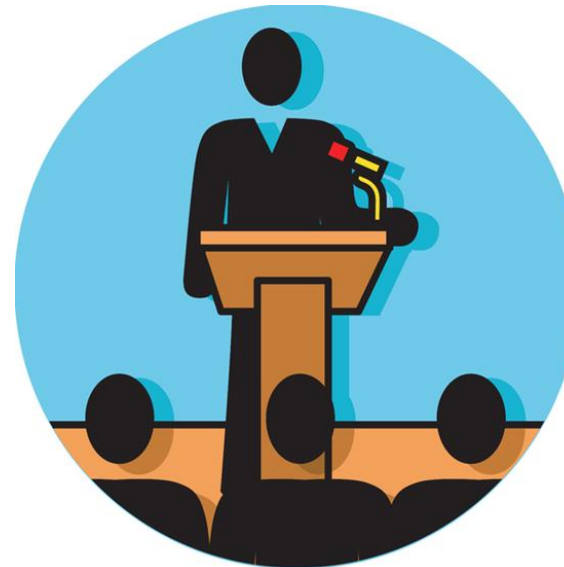
- Sufficient sample size
- Statistical vs. practical significance

Does treatment reduce social anxiety?

Treatment 1



Treatment 2



Effect sizes

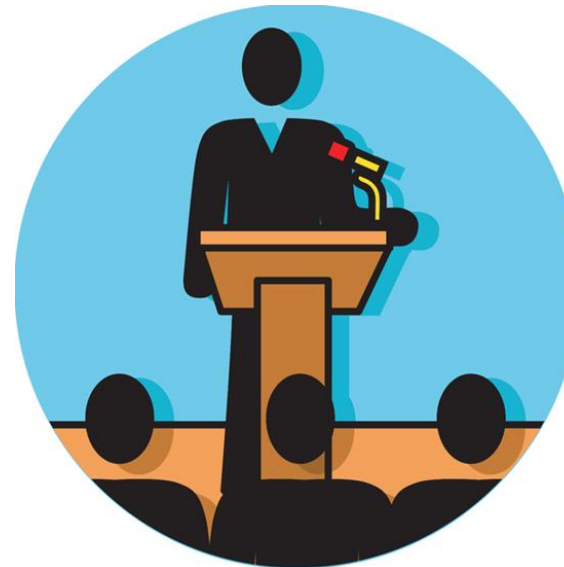
- Sufficient sample size
- Statistical vs. practical significance

Does treatment reduce social anxiety?

Treatment 1



Treatment 2

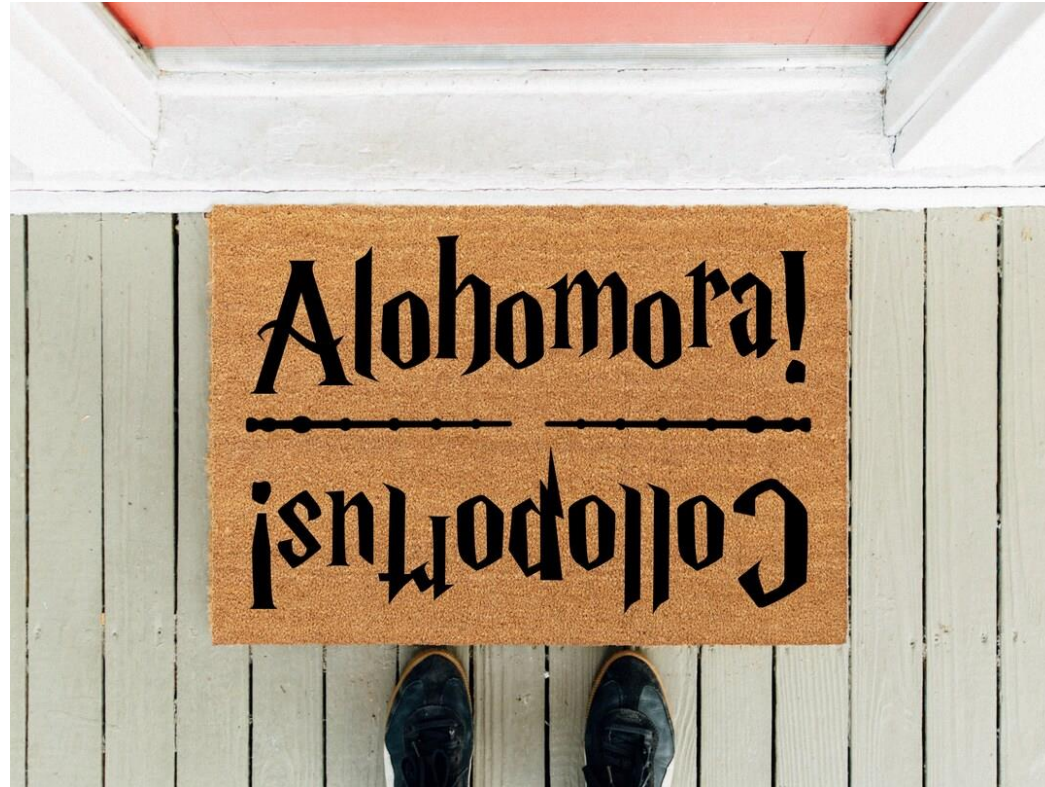


Effect sizes

- Sufficient sample size
- Statistical vs. practical significance
- Power analysis (G*power)



Pre-registration of studies



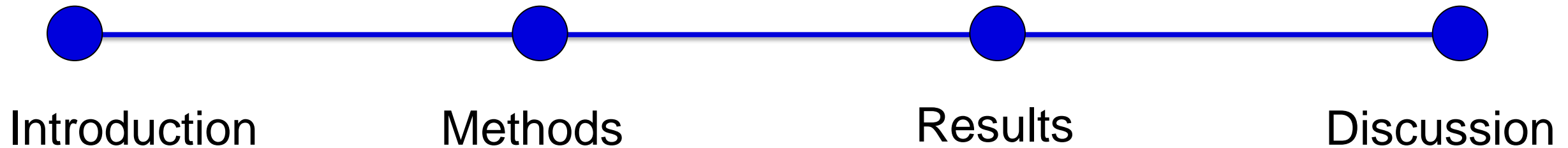
Preregistration

- practice of **registering** the hypotheses, methods, and/or analyses of a scientific study before it is conducted

Registered Report

- format of empirical article where a study proposal is **reviewed** before the research is undertaken

Research paper (IMRAD)



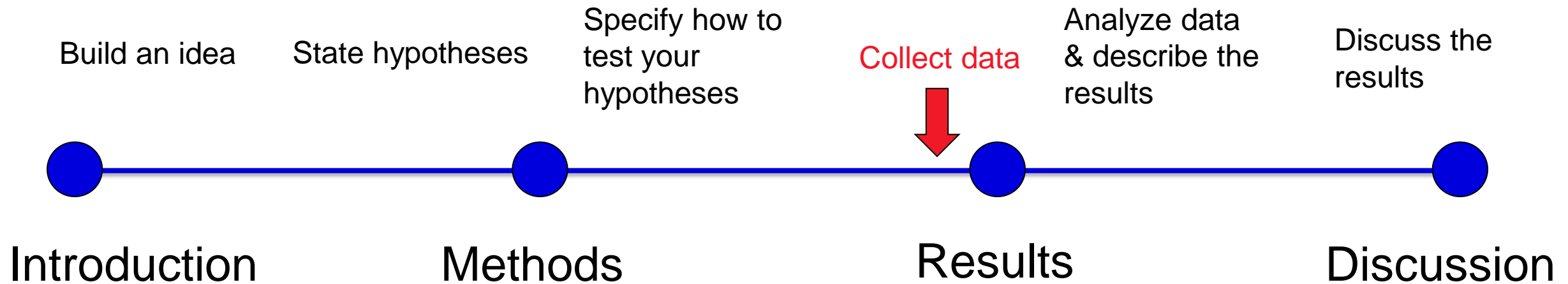
Preregistration

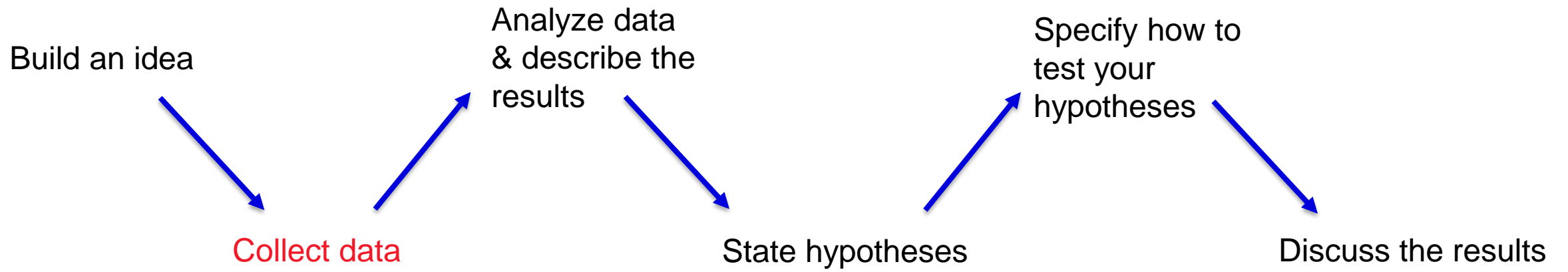
- practice of registering the hypotheses, methods, and/or analyses of a scientific study before it is conducted

Registered Report

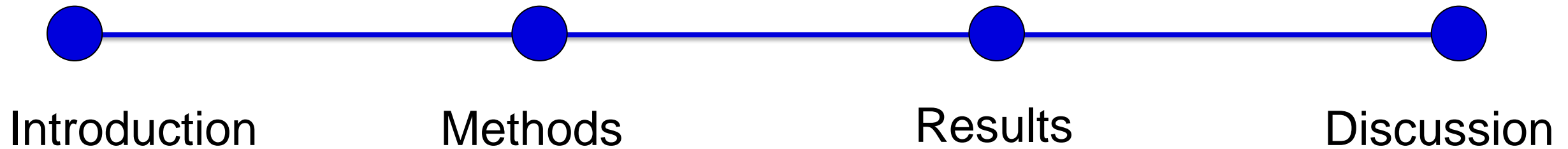
- format of empirical article where a study proposal is reviewed before the research is undertaken

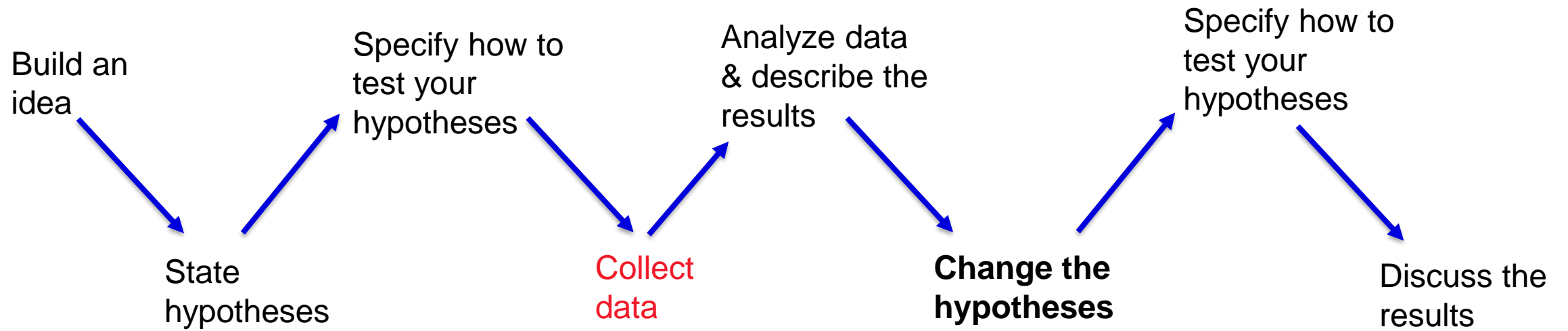
Research paper (IMRAD)



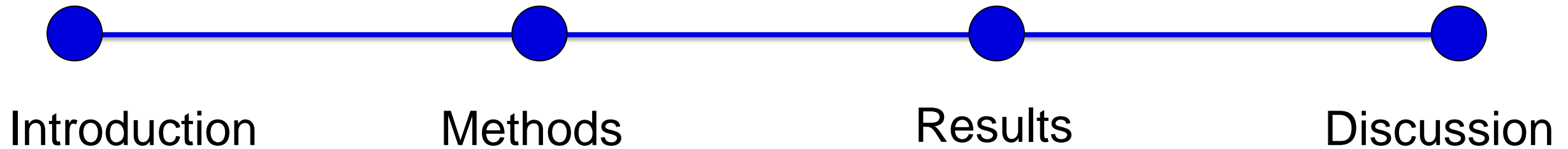


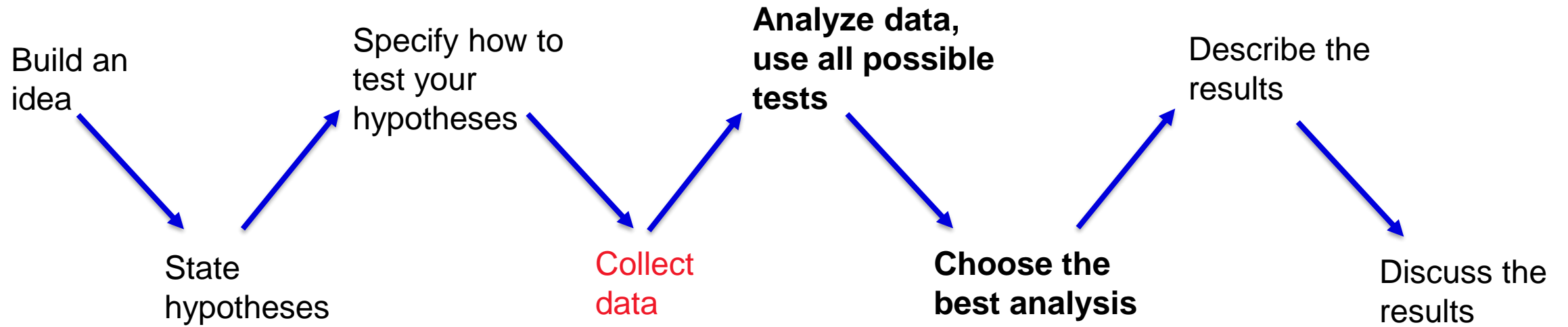
Research paper (IMRAD)





Research paper (IMRAD)

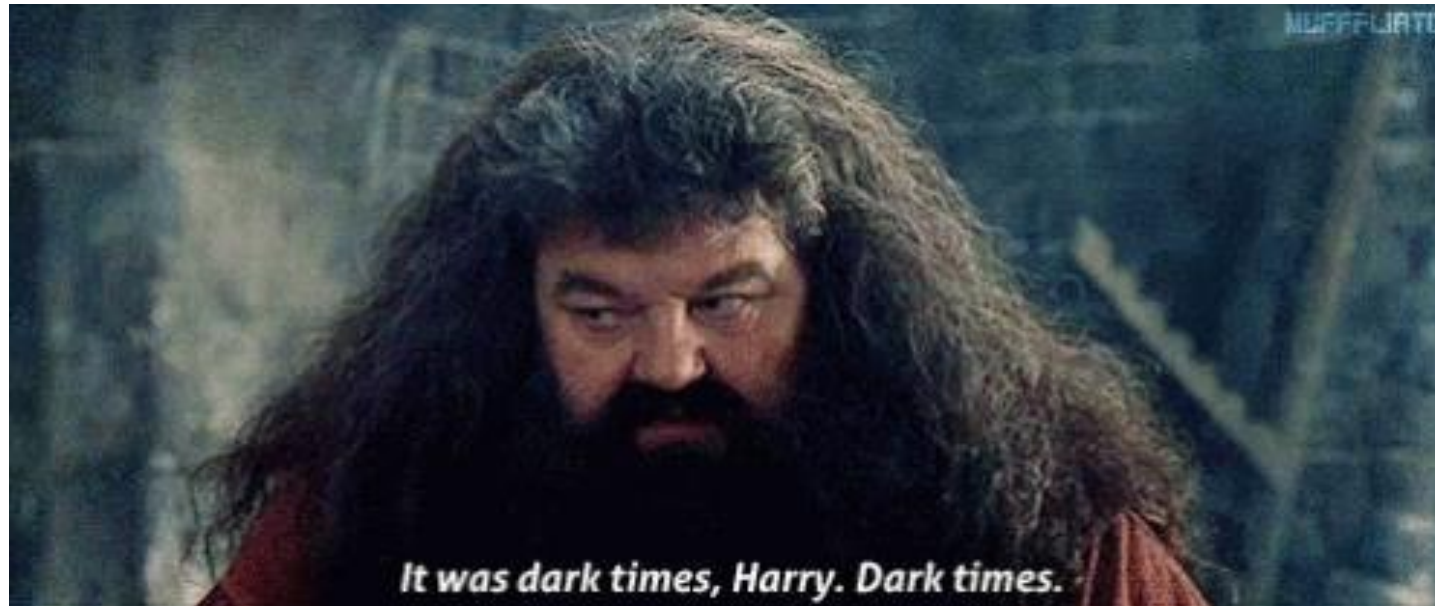




Research paper (IMRAD)



Replication crisis



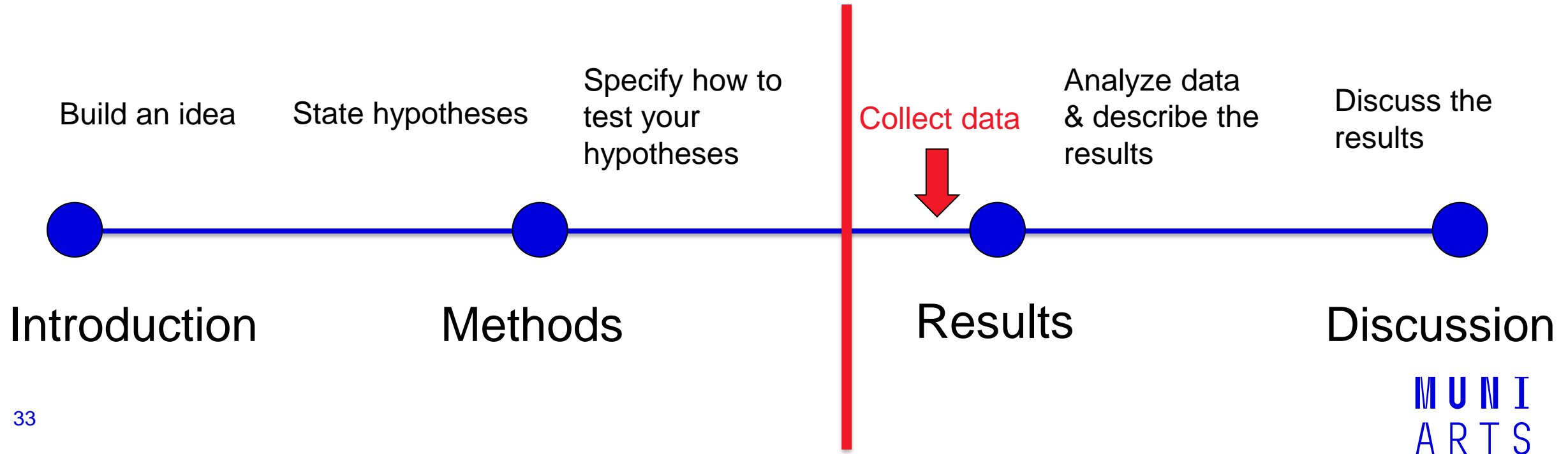
Preregistration

- practice of registering the hypotheses, methods, and/or analyses of a scientific study before it is conducted

Registered Report

- format of empirical article where a study proposal is reviewed before the research is undertaken

Research paper (IMRAD)



Preregistration



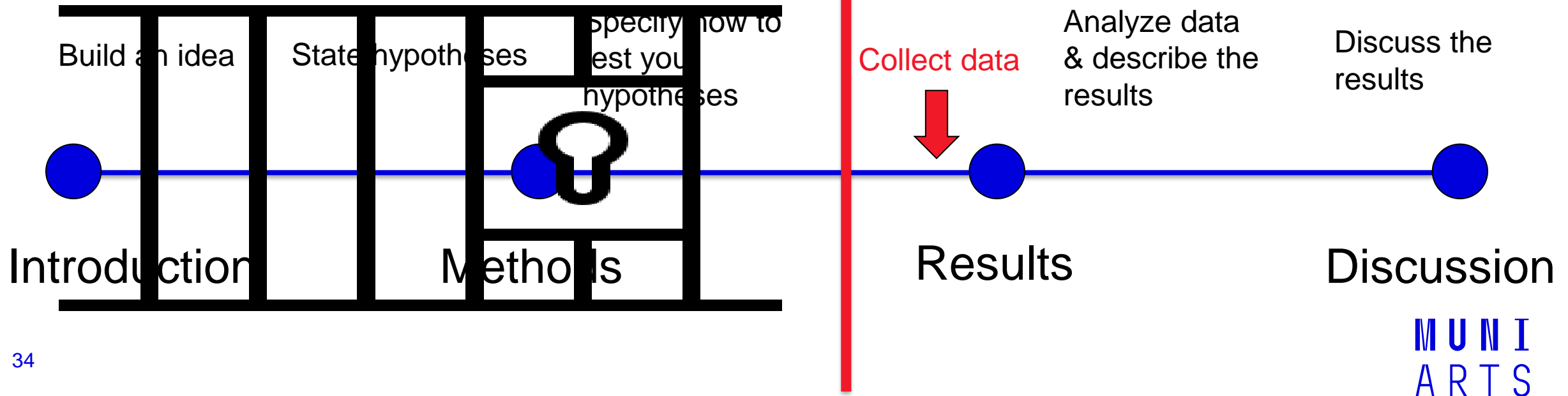
- practice of registering the hypotheses, methods, and/or analyses of a scientific study before it is conducted

Registered Report



- format of empirical article where a study proposal is reviewed before the research is undertaken

Research paper (IMRAD)



Preregistration



- practice of registering the **hypotheses**, **methods**, and/or **analyses** of a scientific study before it is conducted

Registered Report



- format of empirical article where a **study proposal** is reviewed before the research is undertaken

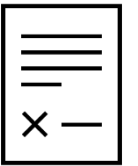


Public web
platform



Journal

Locking the first half of
the journal article before
data collection.



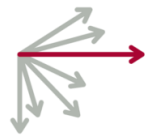
Preregistration



- practice of registering the **hypotheses**, **methods**, and/or **analyses** of a scientific study before it is conducted



Public web platform



AS PREDICTED

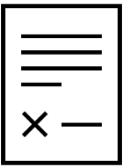
Registered Report



- format of empirical article where a **study proposal** is reviewed before the research is undertaken



Journal



[Registered Reports \(cos.io\)](https://www.cos.io)

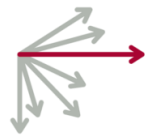
Preregistration



- practice of registering the **hypotheses, methods, and/or analyses** of a scientific study before it is conducted



Public web platform



AS PREDICTED

Title [✗](#)
Method of Loci: The capacity of individual places of the memory palace

Description [✗](#)
This year 2022 can be marked as the beginning of the research on Method of Loci by Mgr. Jan Ondřej in Brno. It is a project aiming at the revival of the ancient method by supporting it with psychological knowledge and research. This particular study is just a small part of the long journey since a lot of questions have arisen and keep on coming to this day. Curiosity is attracted by topics such as the relationship between this method and intelligence or imagination, even personality. There are, however, also smaller topics that are yet to be solved, one of them being the capacity of individual places in memory palaces, relationship between non/verbal memory and Method of Loci plus the possible limitations of using VR technology such as cybersickness. With this study, we aim to test these three phenomena in context of this method with an experiment in hope to learn more about the method, human memory and finding ways how to improve similar experiments in the future.

Contributors
[Remove me](#)
Soňa Kinterová, Jan Ondřej, and Petr Kveton

Category [✗](#)
[Project](#)

Affiliated institutions [✗](#)
No affiliated institutions

License [✗](#)
CC-BY Attribution-NonCommercial-NoDerivatives 4.0 International

Subjects [✗](#)
[Psychology](#) [Social and Behavioral Sciences](#)

Tags [✗](#)
[Memory](#) [Memory palace](#) [Memory palace capacity](#) [Method of Loci](#) [Virtual reality](#)

Study design
Between subjects design with 1 factor (number of words in association encoded into a place in memory palace) with 3 levels (4 words in association per place, 5 words in association per place, 6 words in association per place). The total number of places in the memory palace is 15 and the total number of to-be-remembered items is 60, which means, that the first group has to learn 15 places, second group has to learn only 12 places and the third group only 10. Participants are going to be tested only for the to-be-remembered items and not for the places since they are going to serve as a cue.
• [Tables.pdf](#)

Randomization
We will use simple randomisation. Each participant will be randomly assigned to each experimental group until the required number of participants is met. To check the equivalency of the groups we are going to measure individual memory of every participant.

Sampling Plan
Existing Data
Registration prior to creation of data
Explanation of existing data
No response
Data collection procedures
We will collect the data in laboratories of the Faculty of Arts of Masaryk University in Brno in course of a year and half, however we might use more laboratories in different towns in Czechia to fulfill sample requirements. Participants can register through the official (Czech) pages of the where they can fill in a form to register in the research and get further information (<https://www.janondrej.com/pages.html>). We are going to focus on a population of your middle adults (18-50 years), the reason for this is the possible effect of age on cognitive performance. Regarding sociodemographic factors, there are no limitations. However, if with neurological or psychiatric disorders, participants will be excluded from the sample could significantly affect their performance.

Study Information

Hypotheses
We have decided to state more hypotheses, some of them fall under a more general hypothesis, make them more specific while still supporting the main goal of the research. These hypotheses are:
H1: If the number of words stored in a single place of a memory palace affects recall, then mean recall of words is lower with higher number of words stored within one place of the memory palace.
H1a: Experimental group instructed to remember 4 words per place recalls more words than the group instructed to remember 5 words per place.
H1b: Experimental group instructed to remember 5 words per place recalls more words than the group instructed to remember 6 words per place.
H1c: Experimental group instructed to remember 6 words per place recalls the least words from the experimental groups.
H2: Individual memory abilities show a significant positive relationship with the number of recall words regardless of the group affiliation.
H2a: Verbal memory measured by I-S-T 2000 R shows a significant positive relationship with the number of recalled words during the experiment regardless of the group affiliation.
H2b: Nonverbal memory measured by I-S-T 2000 R shows a significant positive relationship with number of recalled words during the experiment regardless of the group affiliation.
H2c: The Total memory score measured by I-S-T 2000 R shows a significant positive relationship with the number of recalled words during the experiment regardless of the group affiliation.

Design Plan

Study type
Experiment - A researcher randomly assigns treatments to study subjects, this includes field or lab experiments. This is also known as an intervention experiment and includes randomized control trials.

Blinding
No blinding is involved in this study.
is there any additional blinding in this study?
No.

Analysis Plan

Statistical models
The main statistic used for confirming hypotheses is going to be one-way ANOVA, since we are going to compare total of 3 groups. This tool will be also used for testing other groups (for example if participants are divided into groups according to their education). For demographics with only two groups (such as sex) we are going to use t-test. The second tool to confirm hypotheses is correlation, for that we are going to use bivariate correlation (Pearson's r). With this correlation coefficients we are going to estimate the significance and size of the relationship between individual memory abilities - verbal and figural - and the number of recalled words. Data regarding memory are going to be analysed in more steps. We are going to analyse data encompassing memory abilities for 1. the whole experiment, 2. for individual groups, since we expect differences between groups to be present (however, this problem will cease to exist, if we are unable to confirm hypotheses).
No files selected

Transformations
There is no need for additional transformation than summing the number of words recalled.

Inference criteria
We will use the standard $p < .05$ criteria for determining if the ANOVA and the post hoc test yield significant results. We will compare the groups in terms of lenient and strict scoring.

Data exclusion
We are going to exclude incomplete data regarding number of recalled words or data gathered by an incorrect procedure (such as if the participant did not understand the task correctly). Outliers remain part of the analysis, we are going to exclude them in exploratory analysis. No awareness check is needed.

Missing data
Same as above.

Other
Other
The design has some specifics, that need to be mentioned. We are going to use virtual reality for presenting the memory palace. We have chosen the game Fallout4VR. There is also the fact, that we are going to compare our data with data collected by Mgr. Jan Ondřej, the supervisor of this preregistered study. Designs of these two experiments were adjusted, so that the experimental group has gone through the similar process and environment with the only difference being the use of the list of words. Different lists should not be, after all, an issue since the Method of Loci should be generalizable to different materials. This research is conducted based on previous research of Mgr. Jan Ondřej and the questions it raised. As this study is led by the Mgr. Jan Ondřej's team, we continue his work, which remains completely in Czech to this day.

Participants will be welcomed in the laboratory and guided by the administrator. The whole procedure can be summed up in the following steps:
1. Participant signs informed consent.
2. The administrator tests his memory with I-S-T 2000 R (memory module, Czech version).
3. The participant is instructed to watch an instructional video about the Method of Loci. The participant is also informed about the experiment's procedure through the video.
4. The participant is introduced to a VR environment. The administrator will guide him through all 15/12/10 locations. The participant will have 20 seconds to examine each location.
4. The participant exits the VR environment and will recall all 20 places.
5. The participant re-enters the VR environment and go through the locations again, without the 20-second time limit, to ensure that all locations are well remembered.
6. The participant then exits the VR environment and will recall again all 15/12/10 places.
7. The administrator asks the participant to mentally move to the first location. The participant has to name the location. Immediately after, the administrator plays a recording of the 4/5/6 words and the instructions on how to imagine them. The exact process is repeated for each location and each set of words. For further details we suggest looking into the attached files, as they contain tables showing the experimental design and the instructions participants will receive.
8. After the encoding period, when all 60 words are encoded, the participants is instructed to subtract the number 7 from 50.
9. The participants fill out a Google form with pictures of the 15/12/10 places serving as cues, instructed to recall as many words as they can. The time for filling out the form is limited by 10 minutes.
10. With debriefing done, the experiment ends.
The whole experiment will take approximately 60 minutes, with 15 minutes reserved for complications (e.g. cybersickness, late arrival of participants, etc.). This project has so far only one administrator: Soňa Kinterová, but we do not exclude the possibility of other administrators joining in during the collection of data.
No files selected

Sample size
Our target sample size is 159 participants - 53 participants for each group. We will attempt to recruit more, assuming that some participants might not complete all the measurements correctly, which would leave us with smaller sample than required.

Sample size rationale
With G*Power we have determined the sample size of 159 participants for $\alpha = 0.05$, $\beta = 0.20$, effect size = 0.25. This means that for each group there should be at least 53 participants.

Stopping rule
Data collection is limited by time (1.5 years) or by recruiting the required number of participants. The target sample is 159 participants, but we will attempt to recruit up to 180 participants, assuming that not all will complete the total task or some complication will occur.

Variables

Manipulated variables
We manipulated a list of 60 concrete words and divided them into groups of 4, 5 or 6 words associated together. These groups differ in the number of to-be-remembered items, the instruction associating them and the number of places, which participants will use as a cue. For further details we again suggest looking into the attached files. We tried to keep the to-be-remembered items at the length of 4-8 characters, while also ensuring the places contain approximately the same number of characters while considering the characteristics of particular experimental group. The same goes for the instruction. For example, the third and fourth place look like this in terms of characters and words in the instruction:
1. experimental group - 19 words and 96 characters for the third, 19 words and 96 characters for the fourth
2. experimental group - 20 words and 105 characters for the third, 20 words and 104 characters for the fourth
3. experimental group - 23 words and 111 characters for the third, 23 words and 109 characters for the fourth
• [Tables.pdf](#)

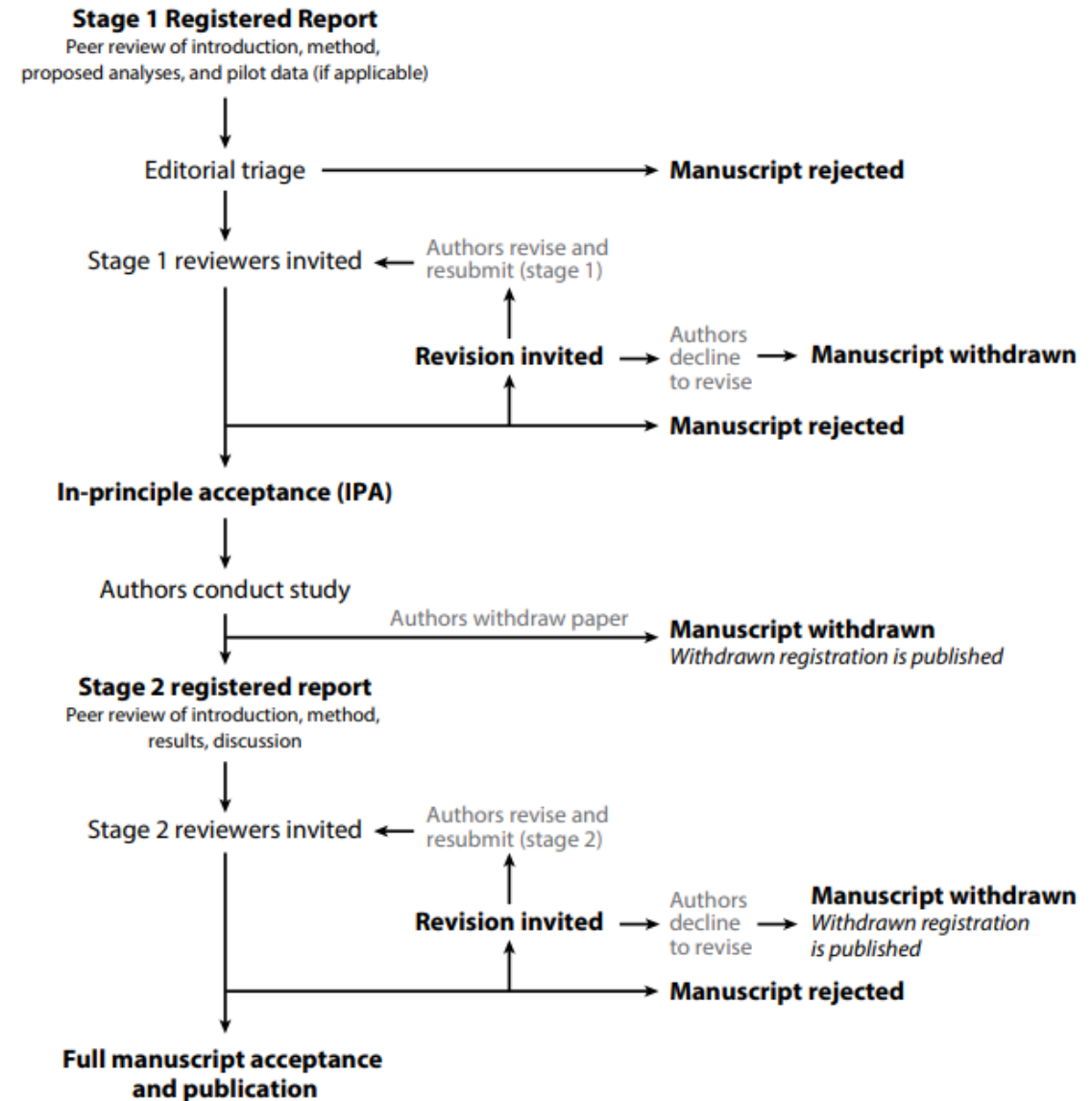
Measured variables
1. The total recalled items in each experimental group (0-60).
2. The total recalled items in each location in each experimental group (0-3).
3. The total correctly answered items in the Verbal Memory test measured by I-S-T 2000 R.
4. The total correctly answered items in the Nonverbal Memory test measured by I-S-T 2000 R.
5. The total correctly answered items in the Whole Memory test measured by I-S-T 2000 R.
6. Demographics (sex, education, age, nationality)
• [al memory abilities will be measured by I-S-T 2000 R \(official Czech version - Form A\)](#)
selected

to differentiate two types of scoring:
scoring = the number of words or synonyms recalled by participant
scoring = the number of words correctly recalled by participant
selected

Registered Report

- format of empirical article where a **study proposal** is reviewed before the research is undertaken

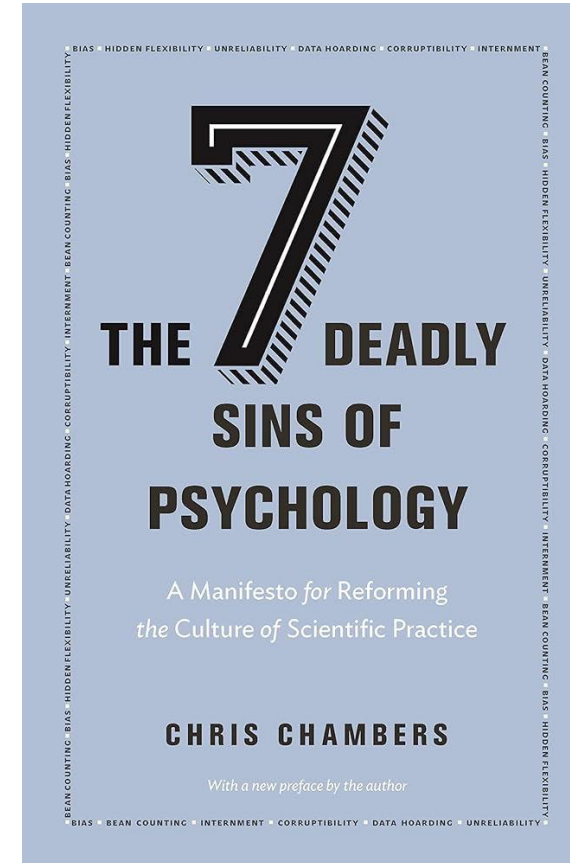
[Advertising cooperative phenotype through costly signals facilitates collective action | Royal Society Open Science \(royalsocietypublishing.org\)](https://royalsocietypublishing.org)





- PCI-RR community
- Popularizing GRPs

"Because the study is accepted in advance, the incentives for authors change from producing the most beautiful story to the most accurate one."



--Chris Chambers

Your decision!



Building brands



IQOS



Google



Apple



WIKIPEDIA
The Free Encyclopedia



HARVARD
UNIVERSITY

MUNI
ARTS

Building brands

- Researcher's limits & biases
- Resources and facilities
- Unexpected problems
- Workplace requirements



MUNI
ARTS

Thank you

...and use magic responsibly!

alexandra.ruzickova@mail.muni.cz

