LEARNING TO BE FAIR:

A PERSONAL JOURNEY IN FAIR PRACTICES AND A WORKSHOP ON PRE-REGISTRATION

Zefan Zheng

Max Planck Institute for Empirical Aesthetics

Open Science Ambassadors

SECTION 1:
A PERSONAL JOURNAL IN FAIR PRACTICES

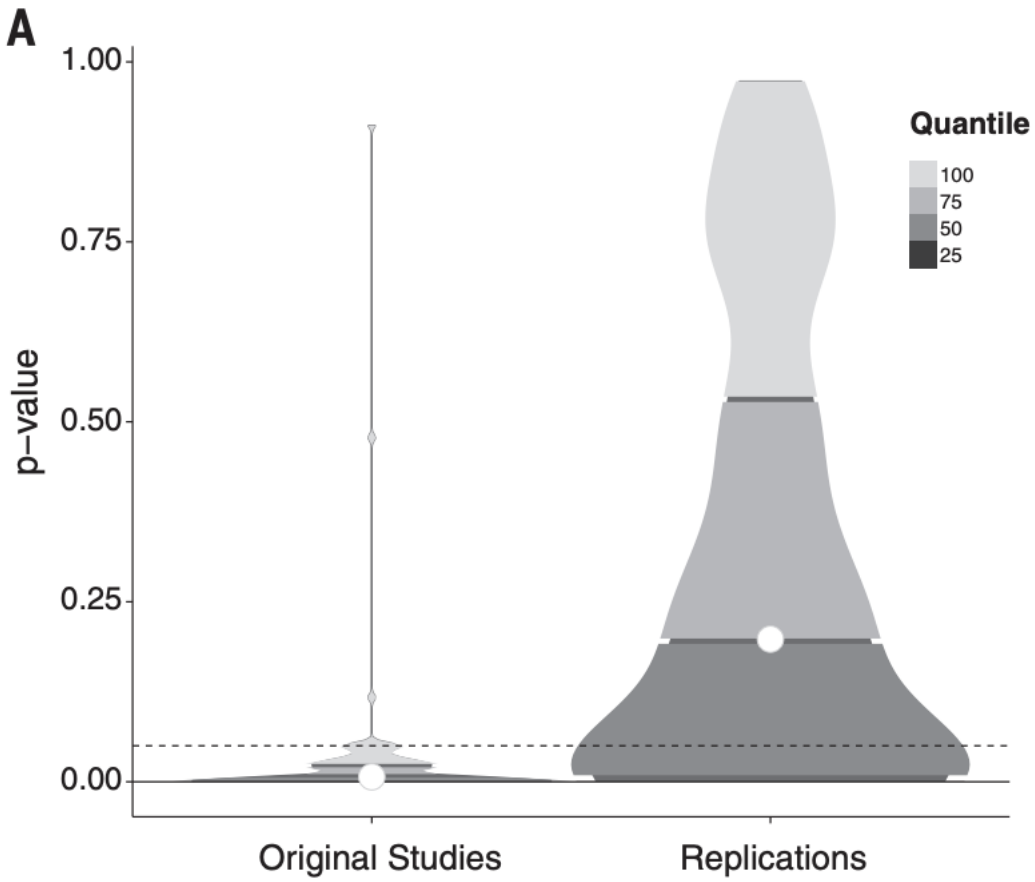
RESEARCH ARTICLE

Estimating the reproducibility of psychological science

OPEN SCIENCE COLLABORATION Authors Info & Affiliations

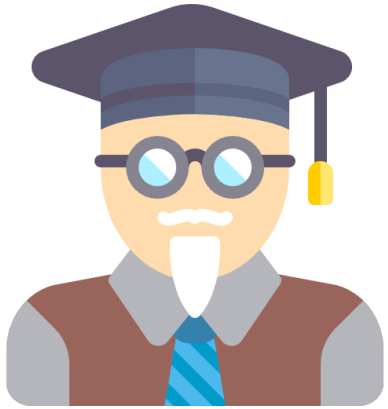
SCIENCE • 28 Aug 2015 • Vol 349, Issue 6251 • DOI: 10.1126/science.aac4716

*"**39%** of effects were subjectively rated to have replicated the original result."*



A personal story

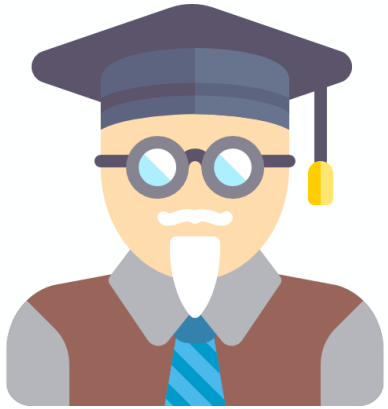
There is this effect. Run this experiment and show the effect!



Professor

PhD

When I was a research intern…



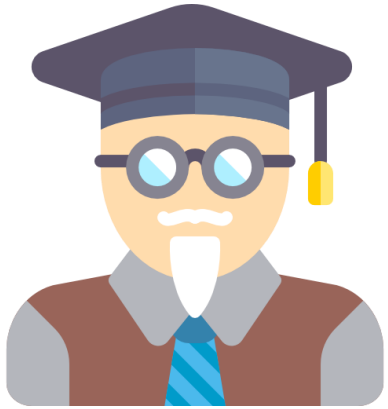
Professor

But the whole field says this effect does not exist…

PhD

When I was a research intern…

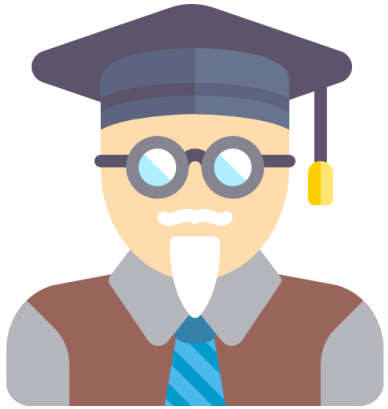
That is exactly why we should try to show it does exist!



Professor

PhD

When I was a research intern…



Professor

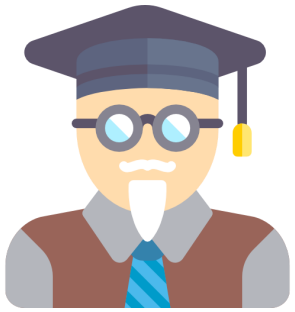
Ok…

PhD

When I was a research intern…

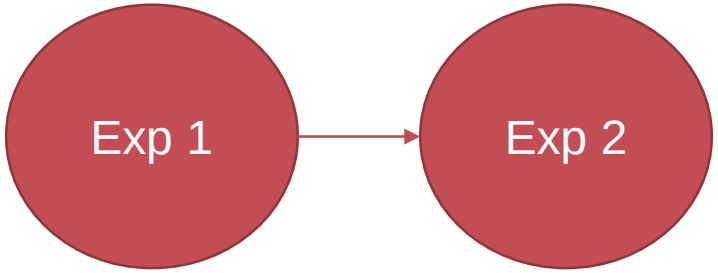
Exp 1

p > 0.05



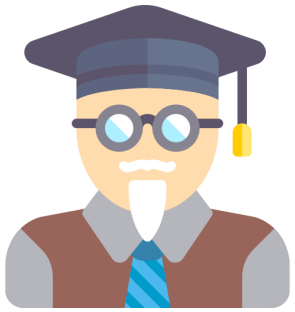
“*You failed experiment!*
Run it again”

When I was a research intern…



p > 0.05

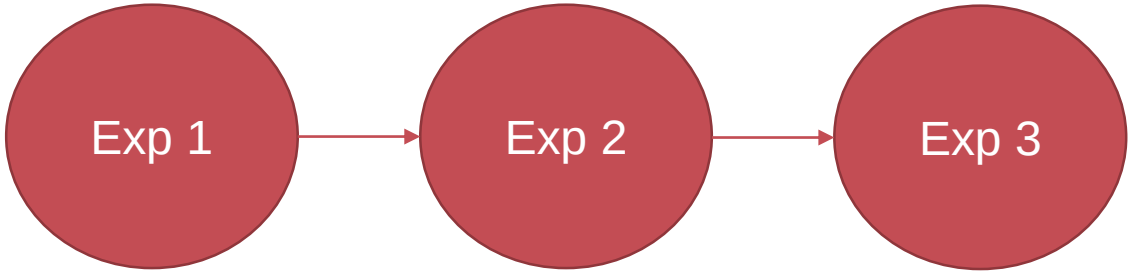
p > 0.05



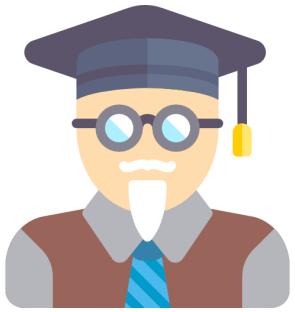
“*You failed experiment! Run it again*”

“*You failed again! Change these parameters, and run it again! ”*

When I was a research intern…



p > 0.05 **p > 0.05** **p > 0.05**



"You failed experiment! Run it again"

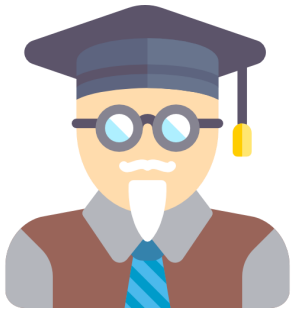
"You failed again! Change these parameters, and run it again! "

"I think you can do a better job. Change these parameters, and run it again! "

When I was a research intern…

Exp 1 → Exp 2 → Exp 3 → ..... Exp 10

p > 0.05 | p > 0.05 | p > 0.05 | p < 0.05



"You failed experiment! Run it again"

"You failed again! Change these parameters, and run it again! "

"I think you can do a better job. Change these parameters, and run it again! "

"Publish this! You can graduate now! "

False Positive

When I was a research intern…

I am done with academia! I will give up my PhD and graduate as a master…

When I was a research intern…

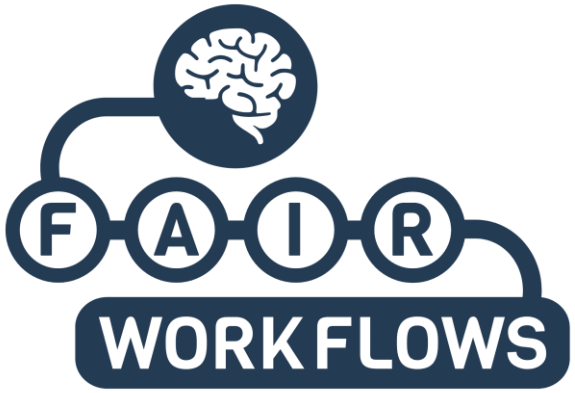
No wonder there is replication crisis in psychology.
I swear I won’t do my PhD in this way!
This is unFAIR for science!

Me

When I was a research intern…

Max-Planck-Institut
für empirische Ästhetik

University of California
CDL
California Digital Library



ChronosHub

DRYAD

Stanford
University

DataCite
FIND, ACCESS, AND REUSE DATA

FAIR WORKFLOWS

DMP

Pre-registration | Code | Pre-registration | Code | Pre-registration | Code | Pre-registration | Code

Exp 1 → Exp 2 → Exp 3 → Exp 4

Data | Report | Data | Report | Data | Report | Data | Paper

Openly shared

Now I am a PhD

My workload for one paper

Pre-registration	Data	Code	Report
Pre-registration	Data	Code	Report
Pre-registration	Data	Code	Report
Pre-registration	Data	Code	Paper

DMP

Other people's workload for one paper

Paper

Now I am a PhD

Damn it!
What I am doing
is
FAIR for science
unFAIR for me!!!

What my work potential employer cares

Paper

What other work potential employer cares

Paper

Now I am a PhD

This won’t work!
If I work for equal amount of time as others, I will publish much less papers than the others.
In the end, I will fail myself in the system!

Now I am a PhD

Wait, first let me figure out how much extra time has it costed to be FAIR for science in the past 2 years.

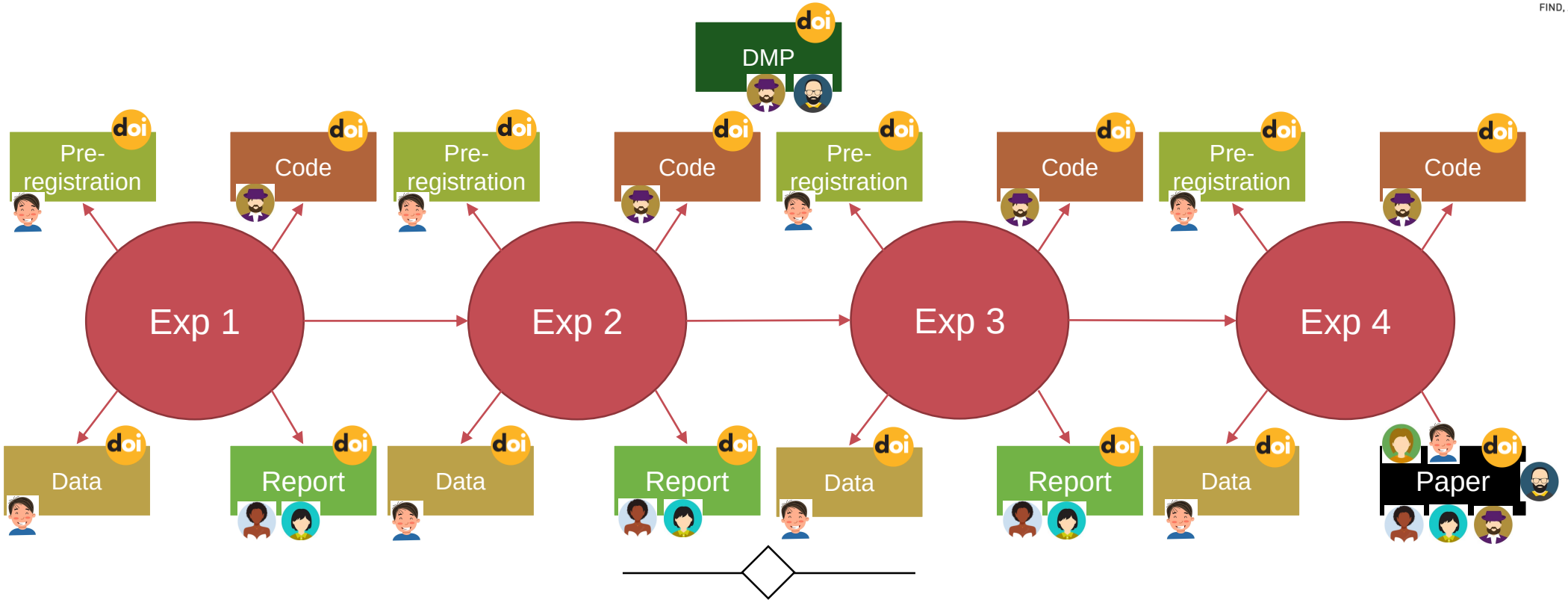
I am recording the **extra time costs** on depositing these outputs on the weekly basis

Pre-registration	Data	Code	Report
Pre-registration	Data	Code	Report
Pre-registration	Data	Code	Report
Pre-registration	Data	Code	

$$\text{Extra work load} = \frac{\text{Extra time cost}}{\text{Total working time}} = \sim 10\%$$

Now I am a PhD

It is all about credits.
What if we make these extra outputs also creditable?

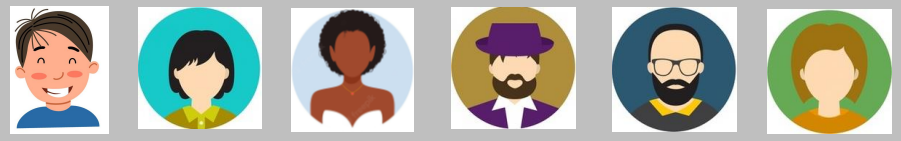


Now I am a PhD

Transforming the Ecology of Academic Crediting System

Creditable and Reusable Research Outputs Now

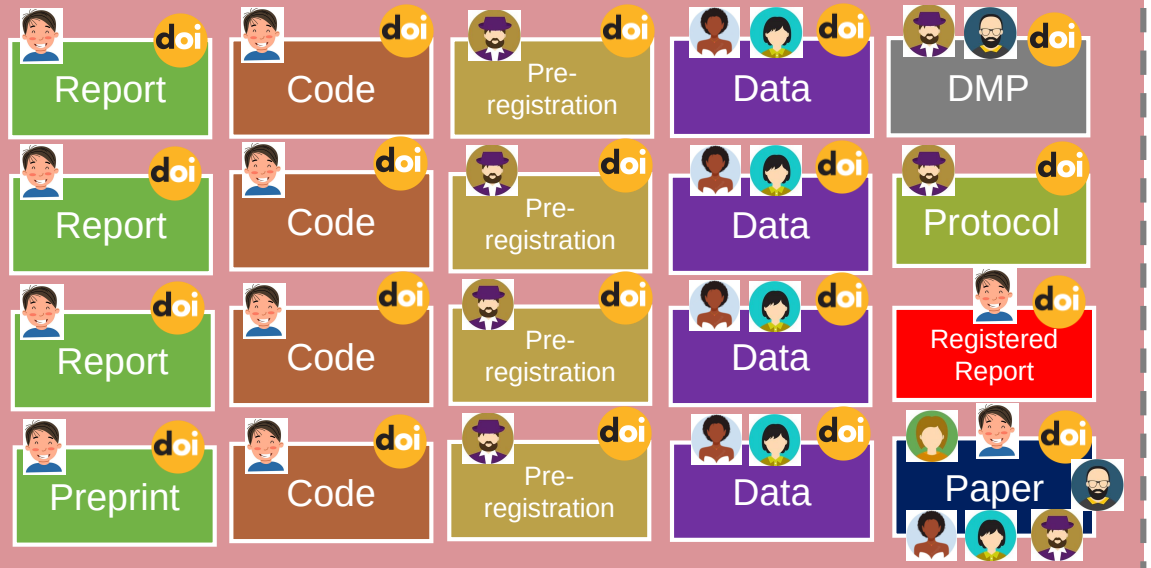
first author



senior author

Paper

Creditable and Reusable Research Outputs With the FAIR Workflow



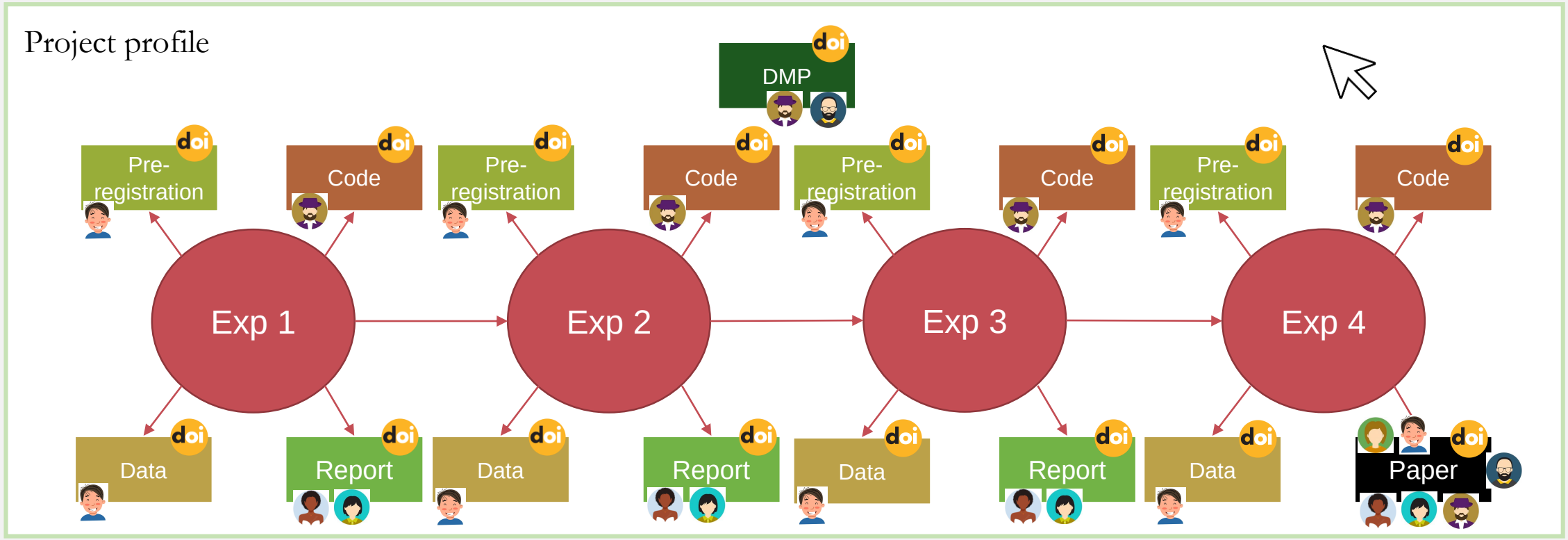
Now I am a PhD

http://www.fairworkflowdashboard.com

FAIR Workflow Dashboard

Project: *Isolating the True Neural Correlates of Consciousness without Cognitive Confounds*

DataCite
FIND, ACCESS, AND REUSE DATA



Now I am a PhD

http://www.fairworkflowdashboard.com

FAIR Workflow Dashboard

Project: *Isolating the True Neural Correlates of Consciousness without Cognitive Confounds*

DataCite
FIND, ACCESS, AND REUSE DATA

Researcher profile

Pre-registration

Project	Citation
XJSIDH	43
ZJCDHSJ	1
ASKDJS	12
SDJCN	0

Data

Project	Citation
CSFCVF	12
SXSXC	1
WEFW	10
CSXCSC	0
SCDSCF	2

Code

Project	Citation
XCVXVF	99
XCVSV	43
FNGBR	48
VBDFJVF	0
DBJGVJB	5
NDJFBVD	65

Protocol

Project	Citation
SDCSVS	0
VBDFJVF	0
DBJGVJB	5
NDJFBVD	65

Registered Report

Project	Citation
XJSIDH	43
ZJCDHSJ	1
ASKDJS	12
SDJCN	0

Report

Project	Citation
CNVJSD	10
SVVNJD	35
OGHNB	67
FGOBFJ	97
FNGBR	48
VBDFJVF	0
DBJGVJB	5
NDJFBVD	65

Paper

Project	Citation
CNVJSD	10
SVVNJD	35

Now I am a PhD

http://www.fairworkflowdashboard.com

FAIR Workflow Dashboard

Project: *Isolating the True Neural Correlates of Consciousness without Cognitive Confounds*

DataCite
FIND, ACCESS, AND REUSE DATA

Researcher profile

Pre-registration

Project	Citation

Data

Project	Citation
CSFCVF	12

Code

Project	Citation
XCVXVF	99

Protocol

Project	Citation

Registered Report

Report

Project	Citation
CNVJSD	10
SVVNJD	35

Paper

Project	Citation
CNVJSD	10
SVVNJD	35
OGHNB	67
FGOBFJ	97
FNGBR	48
VBDFJVF	0
DBJGVJB	5
NDJFBVD	65

Now I am a PhD

A quantitative approach to measure and compare open science engagement

Pre-registration

Project	Citation
XJSIDH	43
ZJCDHSJ	1
ASKDJS	12
SDJCN	0

Data

Project	Citation
CSFCVF	12
SXSXC	1
WEFW	10
CSXCSC	0
SCDSCF	2

Code

Project	Citation
XCVXVF	99
XCVSV	43
FNGBR	48
VBDFJVF	0
DBJGVJB	5
NDJFBVD	65

Protocol

Project	Citation
SDCSVS	0
VBDFJVF	0
DBJGVJB	5
NDJFBVD	65

Registered Report

Project	Citation
XJSIDH	43
ZJCDHSJ	1
ASKDJS	12
SDJCN	0

Report

Project	Citation
CNVJSD	10
SVVNJD	35
OGHNB	67
FGOBFJ	97
FNGBR	48
VBDFJVF	0
DBJGVJB	5
NDJFBVD	65

Paper

Project	Citation
CNVJSD	10
SVVNJD	35

Pre-registration

Project	Citation

Data

Project	Citation
CSFCVF	12

Code

Project	Citation
XCVXVF	99

Protocol

Project	Citation

Registered Report

Report

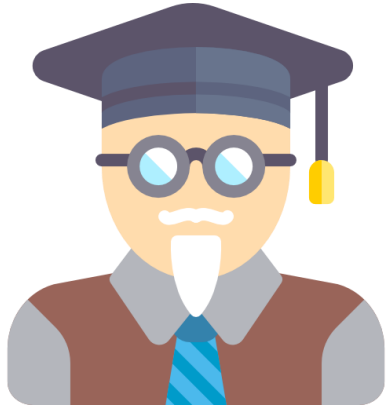
Project	Citation
CNVJSD	10
SVVNJD	35

Paper

Project	Citation
CNVJSD	10
SVVNJD	35
OGHNB	67
FGOBFJ	97
FNGBR	48
VBDFJVF	0
DBJGVJB	5
NDJFBVD	65

I have few papers but a lot of open science outputs!

Who should I hire?



I have a lot of cool papers but
I don't wanna open anything!

A quantitative approach to measure and compare open science engagement

Pre-registration

Project	Citation
XJSIDH	43
ZJCDHSJ	1
ASKDJS	12
SDJCN	0

Data

Project	Citation
CSFCVF	12
SXSXC	1
WEFW	10
CSXCSC	0
SCDSCF	2

Code

Project	Citation
XCVXVF	99
XCVSV	43
FNGBR	48
VBDFJVF	0
DBJGVJB	5
NDJFBVD	65

Protocol

Project	Citation
SDCSVS	0
VBDFJVF	0
DBJGVJB	5
NDJFBVD	65

Registered Report

Project	Citation
XJSIDH	43
ZJCDHSJ	1
ASKDJS	12
SDJCN	0

Report

Project	Citation
CNVJSD	10
SVVNJD	35
OGHNB	67
FGOBFJ	97
FNGBR	48
VBDFJVF	0
DBJGVJB	5
NDJFBVD	65

Paper

Project	Citation
CNVJSD	10
SVVNJD	35

Pre-registration

Project	Citation

Data

Project	Citation
CSFCVF	12

Code

Project	Citation
XCVXVF	99

Protocol

Project	Citation

Registered Report

Report

Project	Citation
CNVJSD	10
SVVNJD	35

Paper

Project	Citation
CNVJSD	10
SVVNJD	35
OGHNB	67
FGOBFJ	97
FNGBR	48
VBDFJVF	0
DBJGVJB	5
NDJFBVD	65

A quantitative approach to measure and compare open science engagement

Pre-registration

Project	Citation
XJSIDH	43
ZJCDHSJ	1
ASKDJS	12
SDJCN	0

Data

Project	Citation
CSFCVF	12
SXSXC	1
WEFW	10
CSXCSC	0
SCDSCF	2

Code

Project	Citation
XCVXVF	99
XCVSV	43
FNGBR	48
VBDFJVF	0
DBJGVJB	5
NDJFBVD	65

Protocol

Project	Citation
SDCSVS	0
VBDFJVF	0
DBJGVJB	5
NDJFBVD	65

Registered Report

Project	Citation
XJSIDH	43
ZJCDHSJ	1
ASKDJS	12
SDJCN	0

Report

Project	Citation
CNVJSD	10
SVVNJD	35
OGHNB	67
FGOBFJ	97
FNGBR	48
VBDFJVF	0
DBJGVJB	5
NDJFBVD	65

Paper

Project	Citation
CNVJSD	10
SVVNJD	35

Pre-registration

Project	Citation

Data

Project	Citation
CSFCVF	12

Code

Project	Citation
XCVXVF	99

Protocol

Project	Citation

Registered Report

Report

Project	Citation
CNVJSD	10
SVVNJD	35

Paper

Project	Citation
CNVJSD	10
SVVNJD	35
OGHNB	67
FGOBFJ	97
FNGBR	48
VBDFJVF	0
DBJGVJB	5
NDJFBVD	65

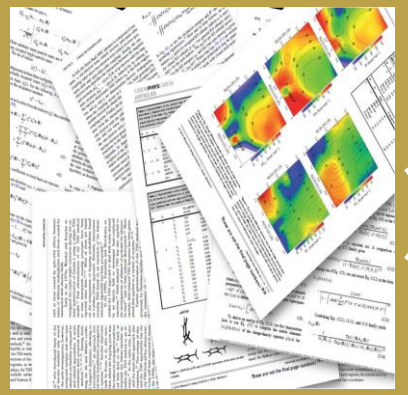
This won’t work!
Now it is true that there will be more data available out there.
But the more data there is out there, the less discoverable they will be!
This is unFAIR for the data!

Now I am a PhD

Current Practice

Hypothesis
"Neural Correlates of consciousness without report confound"

Concept: **Consciousness**
Condition: **Conscious and unconscious**
Task: **No report**
Modality: **MEG, fMRI**



Data not available…

Ideal Practice

Hypothesis
"Neural Correlates of consciousness without report confound"

Concept: **Consciousness**
Condition: **Conscious and unconscious**
Task: **No report**
Modality: **MEG, fMRI**

Now I am a PhD

A metadata template for human cognitive neuroscience

Dataset

Data Modality, Data Acquisition Technique

Preprocessing

Preprocessing Pipeline,
Preprocessing Script

Data Standard

Data Standard,
Reference to Data Standard

Source Dataset

Identifier
Contact info of data manager
ORCID of data manager

Crediting and Acknowledgment

Contributor

First Name, Last Name,
ORCID, Contributor Role,
Career Stage

Related Project

Name,
Project Identifier

Experiment

Area of Research, Clinical Trial, Protocol, Pre-registration

Ethic Approval

Ethic Approval Number
Start/End dates Approval
Authority Approval
Document

Paradigm

Experimental Paradigm,
Publication, Sensory Modality

Subject

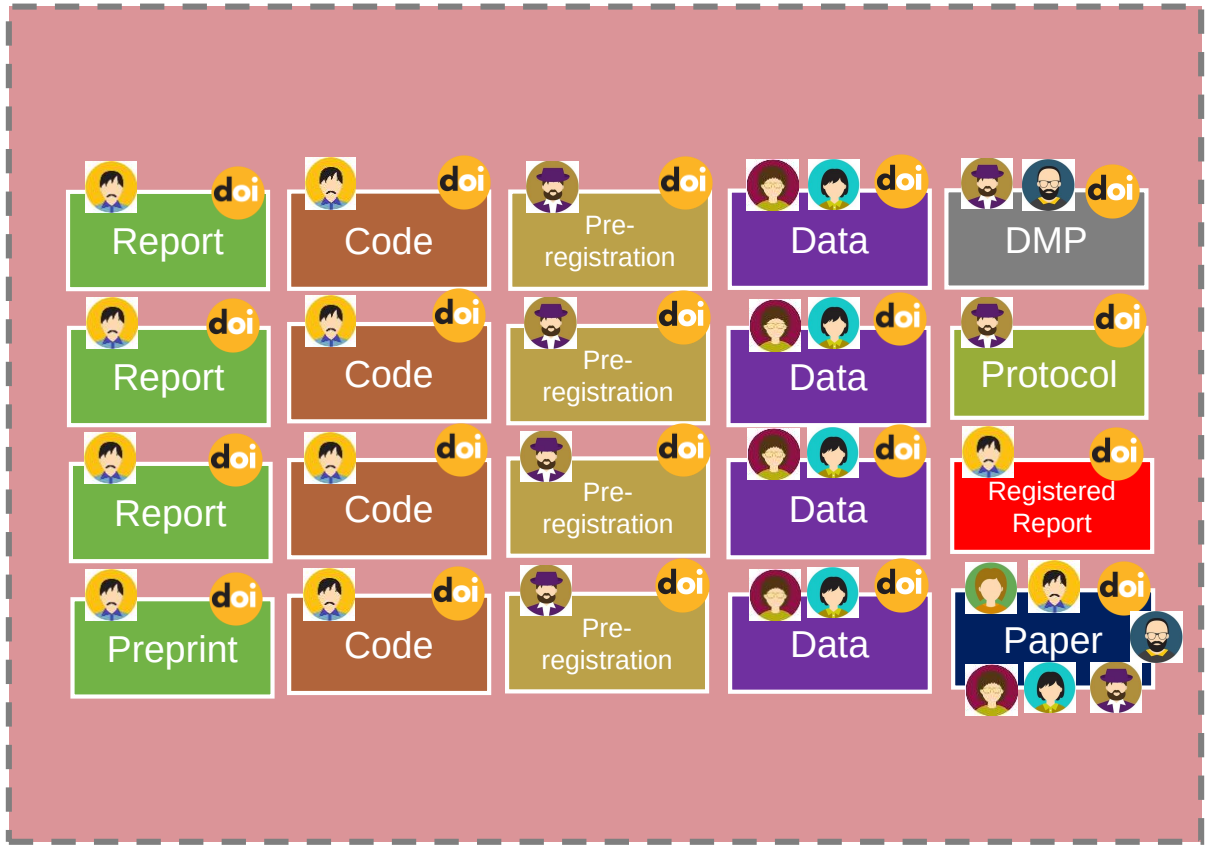
Number of Subjects,
Attributes Measured (Demographic),
Attributes Measured (Medical),
Attributes Measuerd (Occupational),
Attributes Measuerd (Psychological)

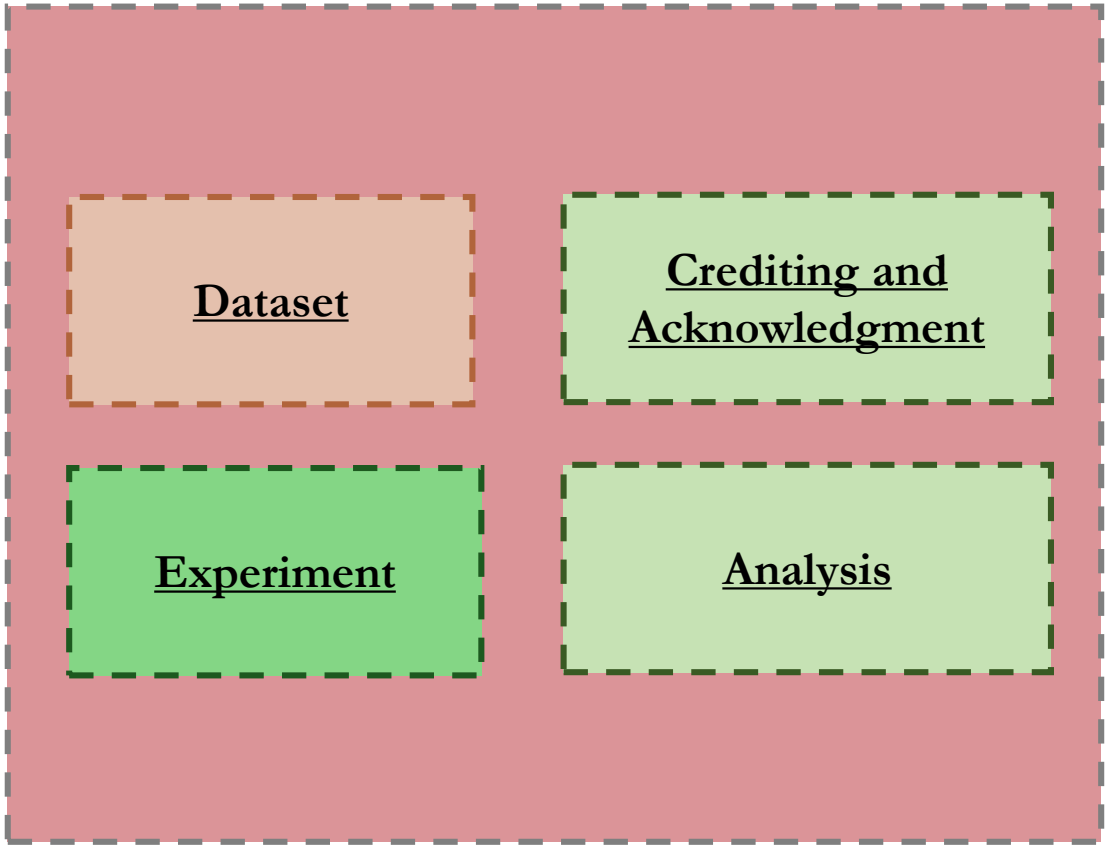
Analysis

Analysis Pipeline,
Statistical Model,
Programming Language

FAIR Data

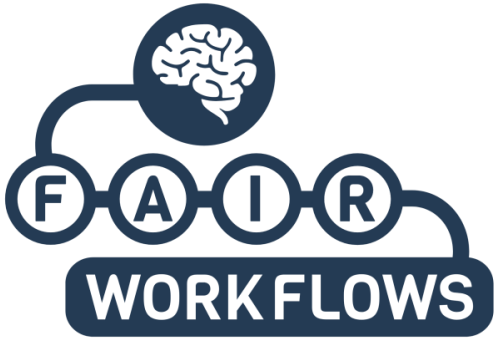
Rich Data + *Rich Metadata*





FAIR for science

FAIR for open scientists



FAIR for data

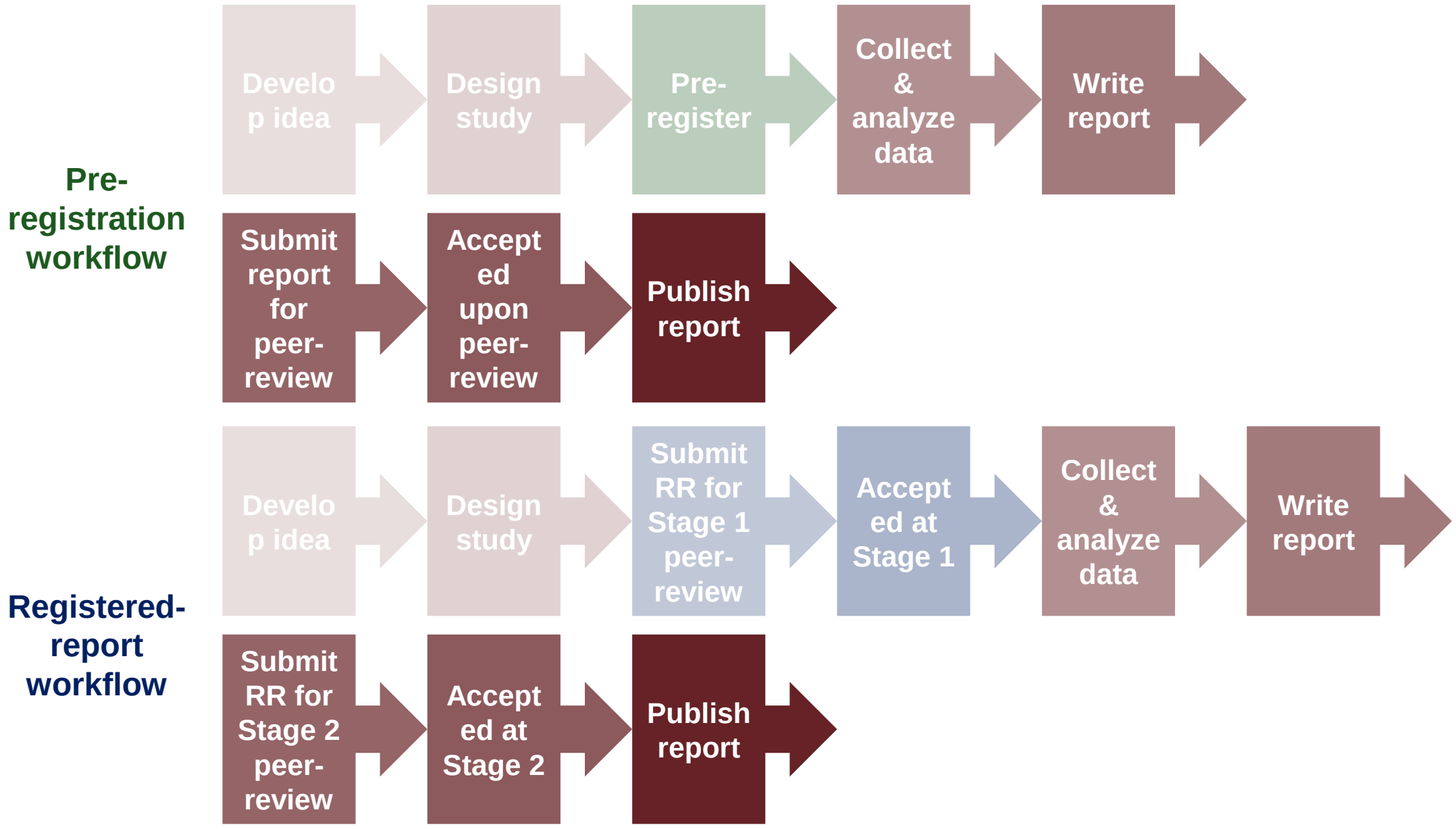
SECTION 2:
A WORKSHOP ON
PRE-REGISTRATION

Pre-registration

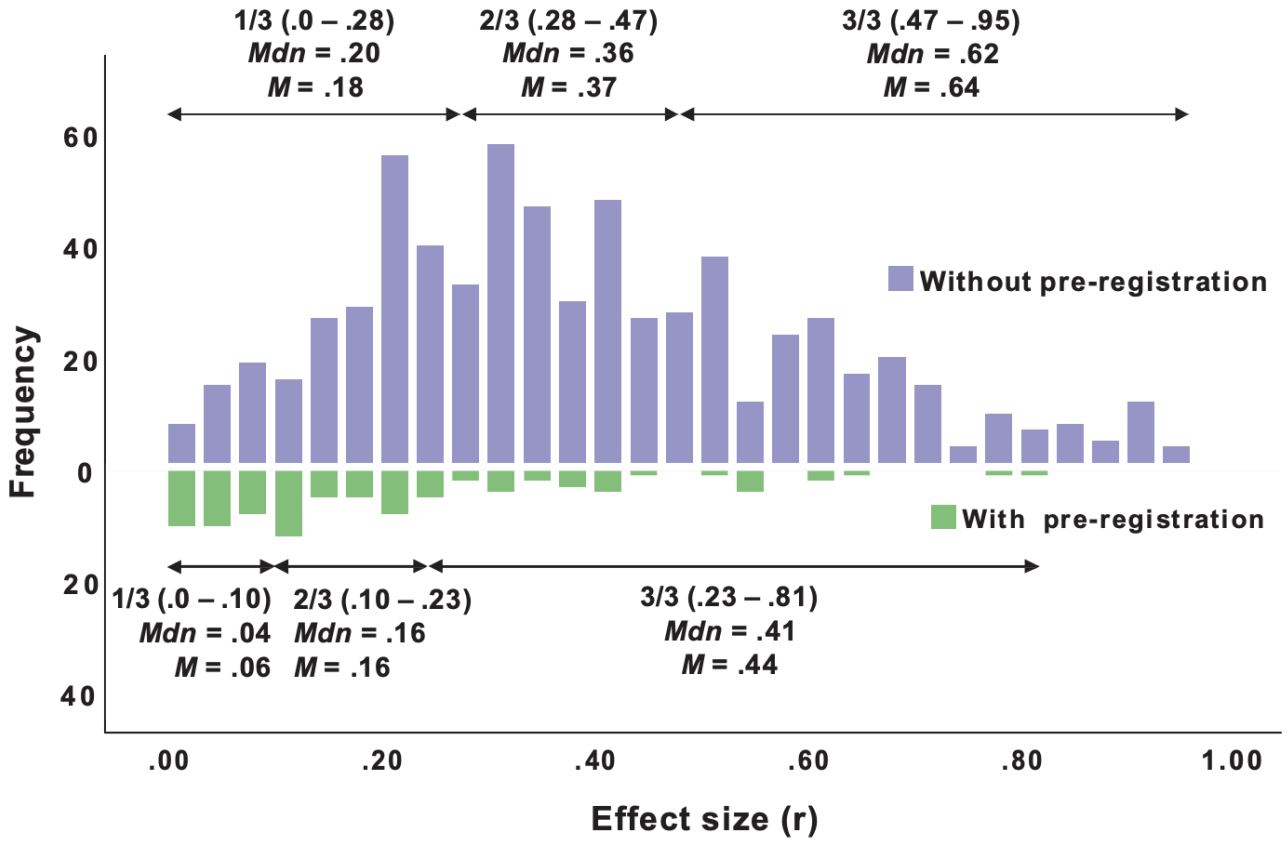
The practice of registering the **hypotheses**, **methods**, and/or **analyses** of a scientific study

before it is conducted.

Pre-registration is different from registered report.

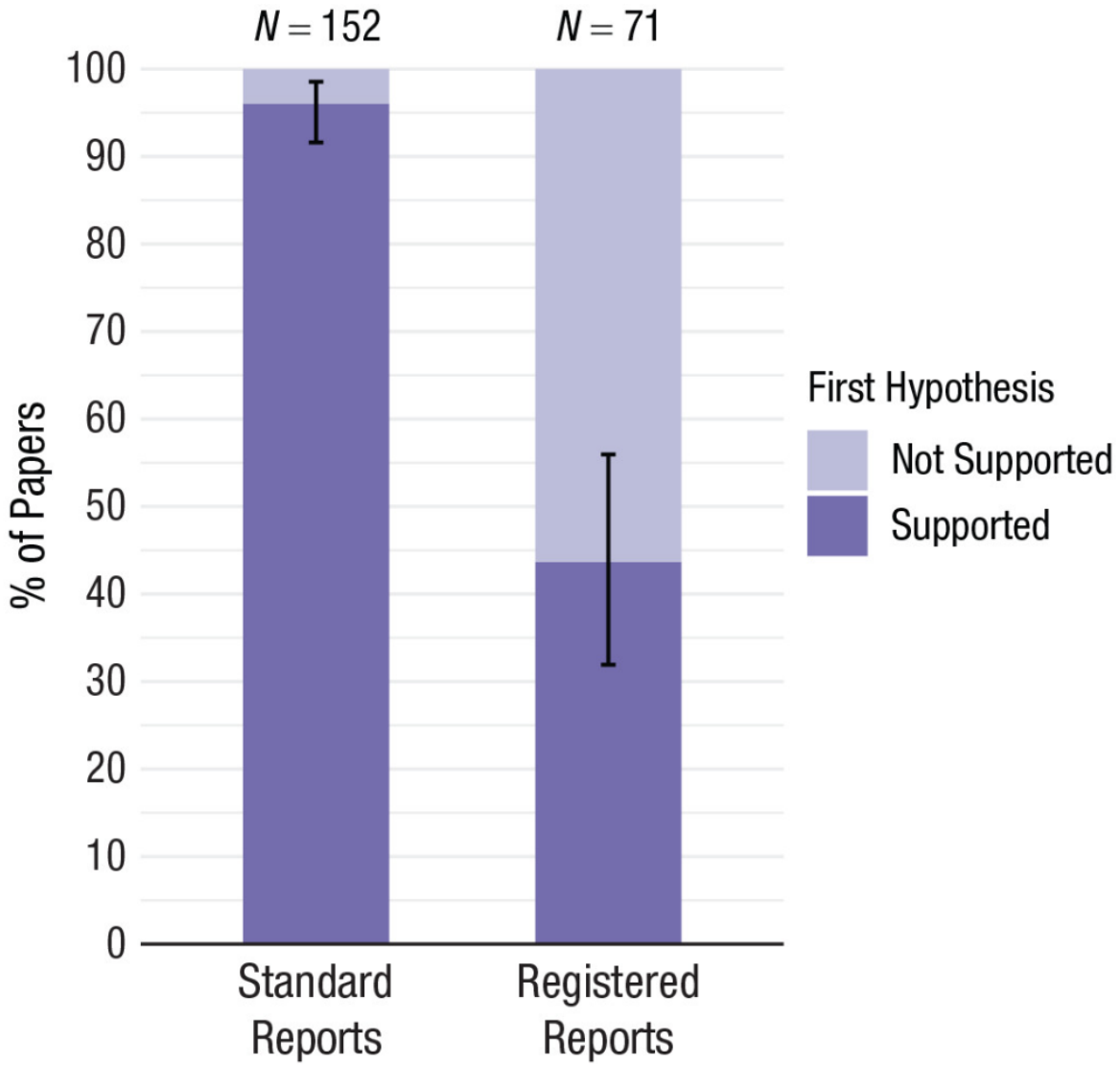


Pre-registration reduces the likelihood of positive results.



Schäfer, T., & Schwarz, M. A. (2019). The Meaningfulness of Effect Sizes in Psychological Research: Differences Between Sub-Disciplines and the Impact of Potential Biases. *Frontiers in Psychology*, *10.*

Registered report reduces the likelihood of positive results.



Scheel, A. M., Schijen, M. R. M. J., & Lakens, D. (2021). An Excess of Positive Results: Comparing the Standard Psychology Literature With Registered Reports. *Advances in Methods and Practices in Psychological Science*, *4*(2), 251524592110074

	Pre-registration	Registered-reports
Advantages	1. Less time consuming 2. Easy to register	1. Guarantees publication regardless of results 2. Peer-review prior to experimentation improves experiment quality
Disadvantages	1. Does not guarantee publication 2. Lack of quality control (no peer-review)	1. More time-consuming 2. Two times of peer-view could take a long long time

The most popular platform for drafting and publishing pre-registrations

Pre-registration template selection

Normal recommended choice

- (●) OSF Preregistration

Template-free registration

- ( ) Open-Ended Registration

Qualitative research

- ( ) Qualitative Preregistration

Secondary analysis on existing data

- ( ) Secondary Data Preregistration

Meta-analysis

- ( ) Generalized Systematic Review Registration

Upload Stage 1 accepted RR manuscript

- ( ) Registered Report Protocol Preregistration

Pre-Data Collection

- ( ) OSF-Standard Pre-Data Collection Registration

Alternative template

- ( ) Preregistration Template from AsPredicted.org

Replication experiments

- ( ) Replication Recipe (Brandt et al., 2013): Post-Completion
- ( ) Replication Recipe (Brandt et al., 2014): Pre-Registration

Social psychology experiment

- ( ) Pre-Registration in Social Psychology (van 't Veer & Giner-Sorolla, 2016): Pre-Registration

Template overview

Metadata
Title
Description
Contributor
Category
Affliated institutions
License
Subjects

Study Information
Hypotheses

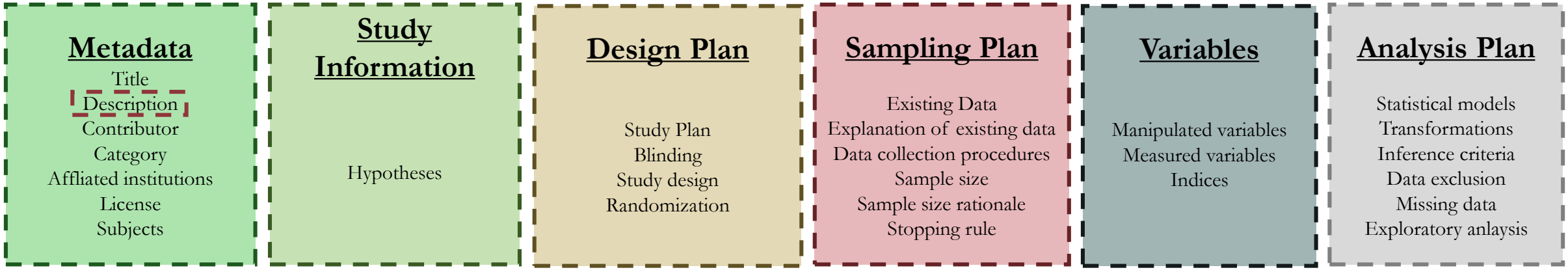
Design Plan
Study Plan
Blinding
Study design
Randomization

Sampling Plan
Existing Data
Explanation of existing data
Data collection procedures
Sample size
Sample size rationale
Stopping rule

Variables
Manipulated variables
Measured variables
Indices

Analysis Plan
Statistical models
Transformations
Inference criteria
Data exclusion
Missing data
Exploratory anlaysis

Template overview



Description *

This field can't be blank.

Basically the abstract:

1. Research background
2. Research question
3. Design
4. Hypothesis

Template overview

Metadata
Title
Description
Contributor
Category
Affliated institutions
License
Subjects

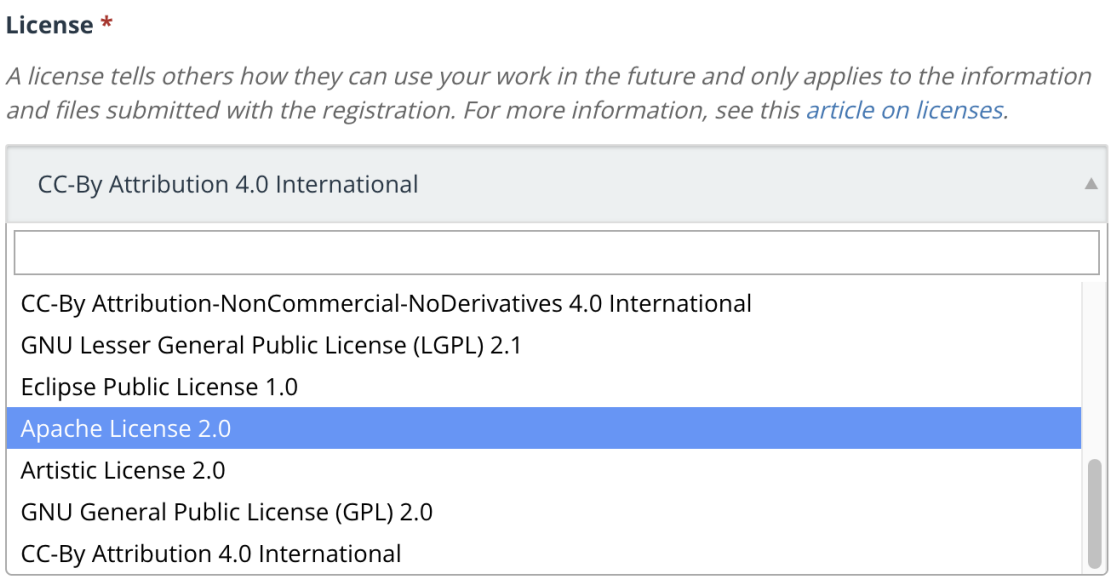
Study Information
Hypotheses

Design Plan
Study Plan
Blinding
Study design
Randomization

Sampling Plan
Existing Data
Explanation of existing data
Data collection procedures
Sample size
Sample size rationale
Stopping rule

Variables
Manipulated variables
Measured variables
Indices

Analysis Plan
Statistical models
Transformations
Inference criteria
Data exclusion
Missing data
Exploratory anlaysis



Clarity how can others use your pre-registration
without your consent:

1. Whether copyright holder has control over the content versioning;
2. Whether one can copy, distribute and modify the content;
3. Whether one should give attribution to the creator when re-using it;

Template overview

Metadata
Title
Description
Contributor
Category
Affliated institutions
License
Subjects

Study Information
Hypotheses

Design Plan
Study Plan
Blinding
Study design
Randomization

Sampling Plan
Existing Data
Explanation of existing data
Data collection procedures
Sample size
Sample size rationale
Stopping rule

Variables
Manipulated variables
Measured variables
Indices

Analysis Plan
Statistical models
Transformations
Inference criteria
Data exclusion
Missing data
Exploratory anlaysis

Hypotheses *

List specific, concise, and testable hypotheses. Please state if the hypotheses are directional or non-directional. If directional, state the direction. A predicted effect is also appropriate here. If a specific interaction or moderation is important to your research, you can list that as a separate hypothesis.

Hide example

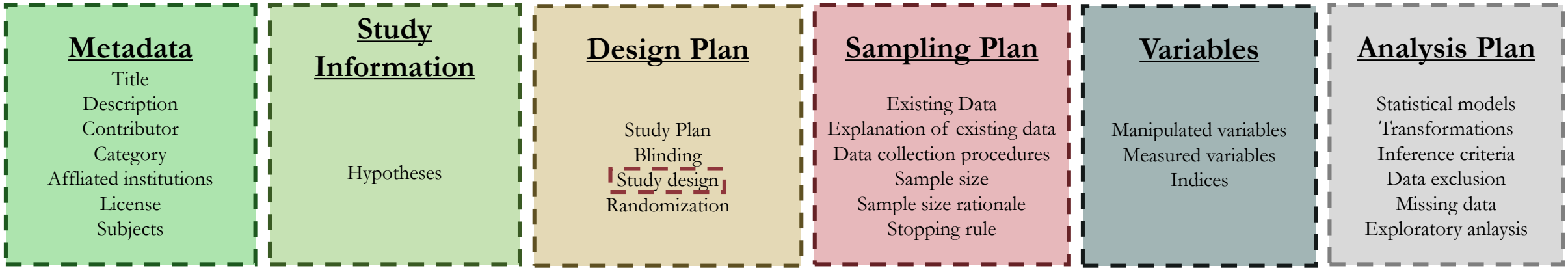
If taste affects preference, then mean preference indices will be higher with higher concentrations of sugar.

This field can't be blank.

A very specific statement of your hypotheses:

1. One point for each hypothesis;
2. Specific about the conditions of each hypothesis;
3. Specific about the empirical outcome measures;
4. Specific about the directions of the results;
5. You MUST report the results of this hypothesis testing in the manuscript as “pre-registered hypothesis”;
6. Any changes in real practice deviating from these statements must be

Template overview



Study design *

Describe your study design. The key is to be as detailed as is necessary given the specific parameters of the design. There may be some overlap between this question and the following questions. That is OK, as long as sufficient detail is given in one of the areas to provide all of the requested information. Examples include two-group, factorial, randomized block, and repeated measures. Is it a between (unpaired), within-subject (paired), or mixed design? Describe any counterbalancing required.

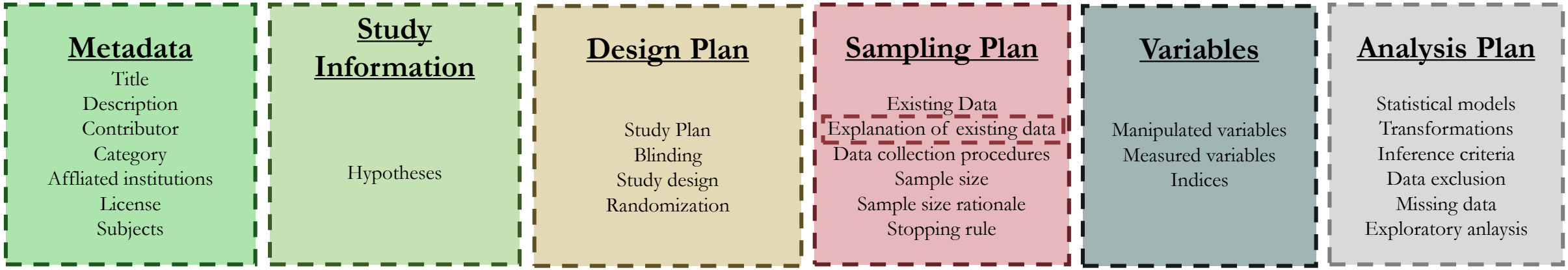
Show example

This field can't be blank.

Equivalent to experimental procedure and design in your manuscript:

1. Should be specific enough in experimental parameters so that others can replicate your experiment procedure;
2. You can also add figures as supplementary material;

Template overview



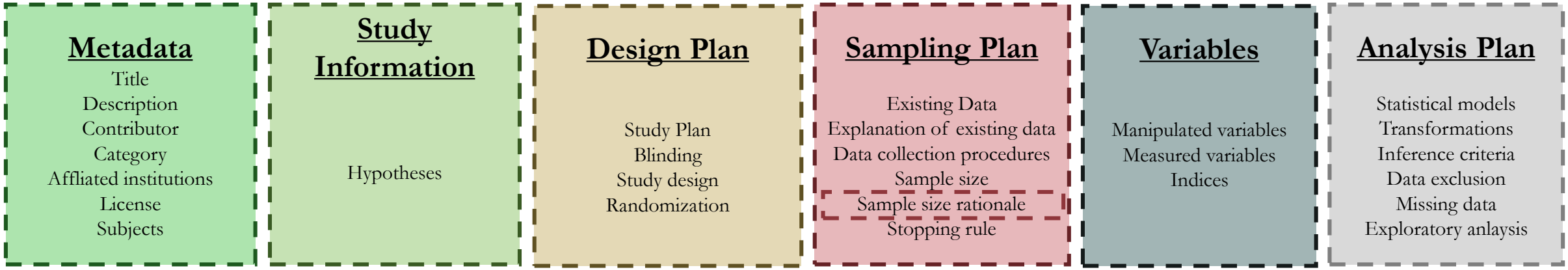
Explanation of existing data

If you indicate that you will be using some data that already exist in this study, please describe the steps you have taken to assure that you are unaware of any patterns or summary statistics in the data. This may include an explanation of how access to the data has been limited, who has observed the data, or how you have avoided observing any analysis of the specific data you will use in your study.

You can also pre-register analysis on an existing data:

1. Be honest whether you are already aware of the results of the pre-registered analysis;

Template overview



Sample size rationale

This could include a power analysis or an arbitrary constraint such as time, money, or personnel.

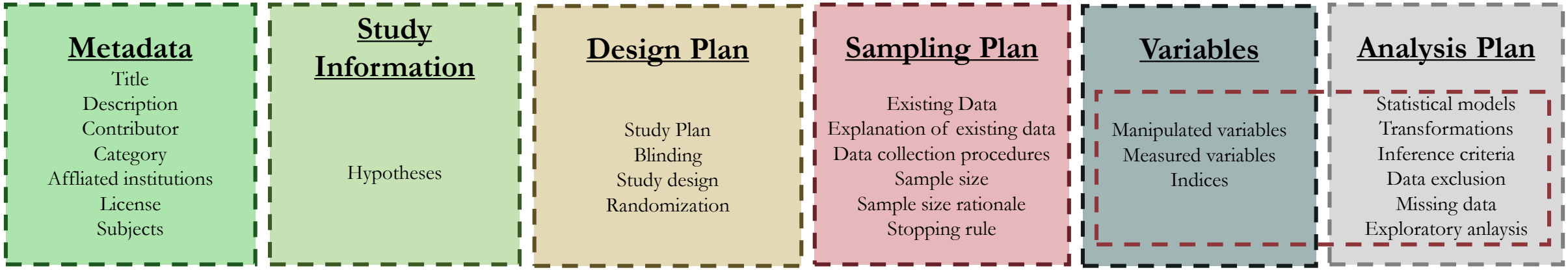
Hide example

We used the software program G*Power to conduct a power analysis. Our goal was to obtain .95 power to detect a medium effect size of .25 at the standard .05 alpha error probability.

Motivate the planned sample size:

1. Power analysis has to be done to justify the sample size;
2. One should only pre-register a study if the planned sample size has enough power to detect the hypothesized effect size.

Metadata: Description

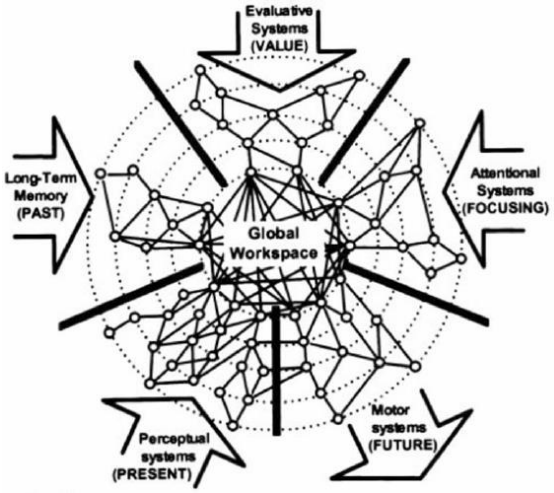


Basically the data analysis section:

1. In addition to the textual information, the best practice would be to write up the code and attach it to the analysis plan;
2. Given your data and these information, one should be able to completely replicate your analysis and results;
3. Any deviations in analysis in the final manuscript should be backed up by strong justifications;

Lessons from a large scale international open science project

Global Neuronal Workspace Theory



COGTATE

PREREGISTERED

OPEN DATA

OPEN MATERIALS

VS

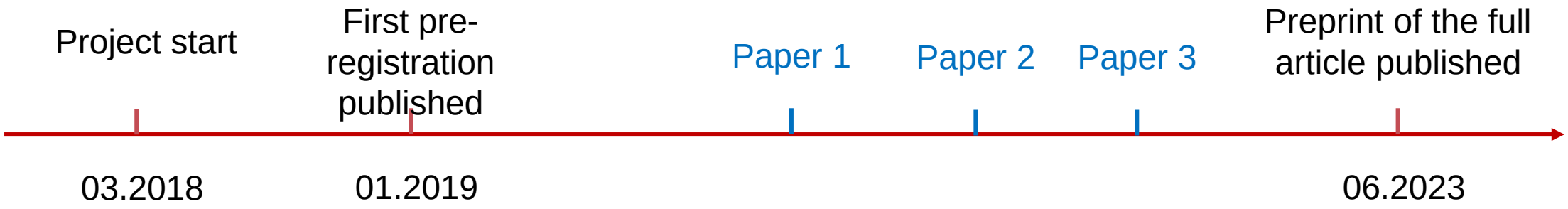
Φ

7 million euro funding

7 independent labs across the globe participated in data collection

Integrated Information Theory

Lessons from a large scale international open science project



Embargo your pre-registration

Separate the date of pre-registration and public access

Questions?

Quiz

1. Pre-registration only applies to quantitative research.

Quiz

2. I don’t need to pre-register the experiments, the sample size of which do not have enough power to detect the effect of interest.

Quiz

3. I do not need to clarify whether the data has been collected in the pre-registration as long as I haven’t conducted the pre-registered analysis.

Quiz

4. I can report the results of the analysis that I did not pre-register in my paper.

Quiz

5. I cannot amend my pre-registration once it is published.

Quiz

6. I have to report the results of every single analysis I pre-register.

Quiz

7. My PhD is supported by a three-year-grant and I have five research projects planned out. I want to do a registered-report for every project.

Quiz

8. All my pre-registered results are negative. No journals want to accept my paper. There is no way I have share my results to the community.

Quiz

9. I have got a five-year grant. I want to replicate a highly influential paper in my field. I think pre-registration is better for me compared to registered report.

Quiz

10. I can draft a license myself for my pre-registration.

Thanks to…

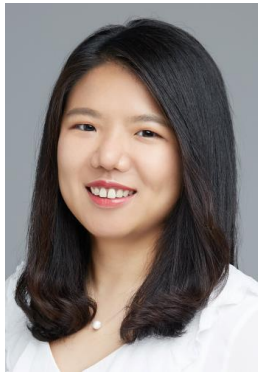
Scientific collaborators

Prof. Lucia Melloni

Dr. Darinka Truebutschek

Alex Lepauvre

Open Science collaborators



Xiaoli Chen

Dr. Helena Cousijn

Tanya Brown

External collaborators

DataCite
FIND, ACCESS, AND REUSE DATA

University of California
CDL
California Digital Library

ARDC
Australian Research Data Commons

ChronosHub

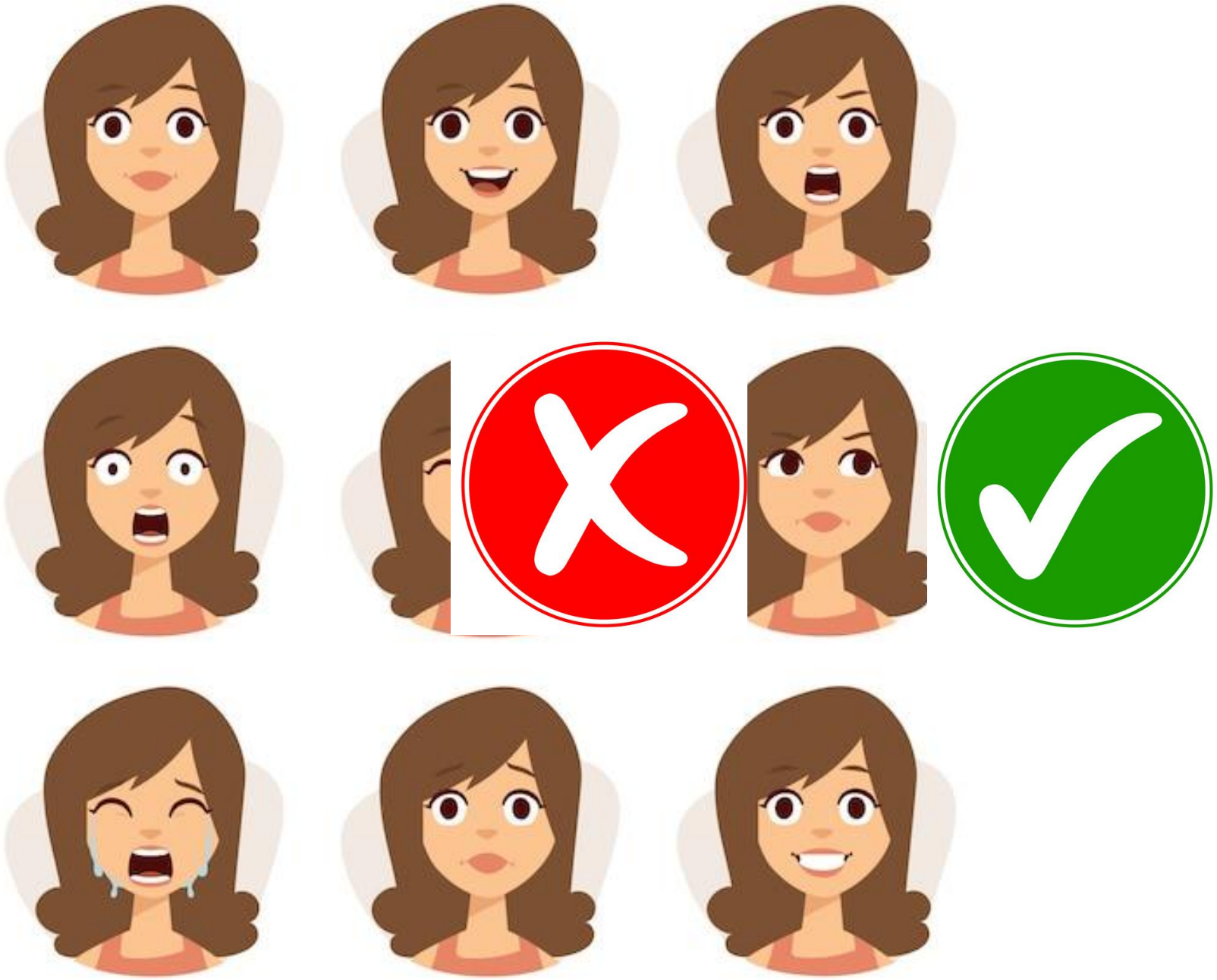
Stanford
University

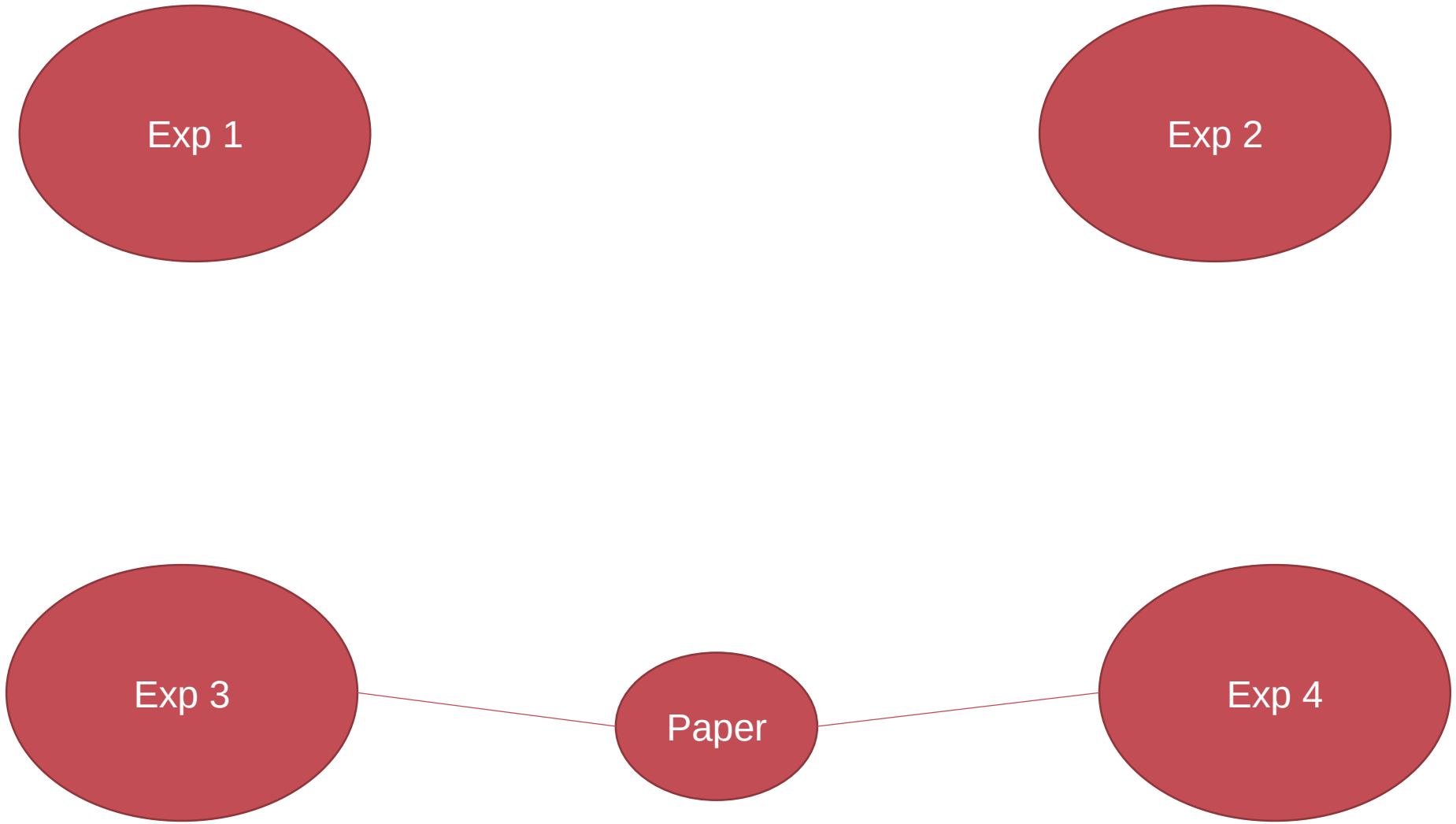
Max-Planck-Institut
für empirische Ästhetik

Crossref

DRYAD

COS
CENTER FOR
OPEN SCIENCE




Exp 1
Exp 2
Exp 3
Paper
Exp 4