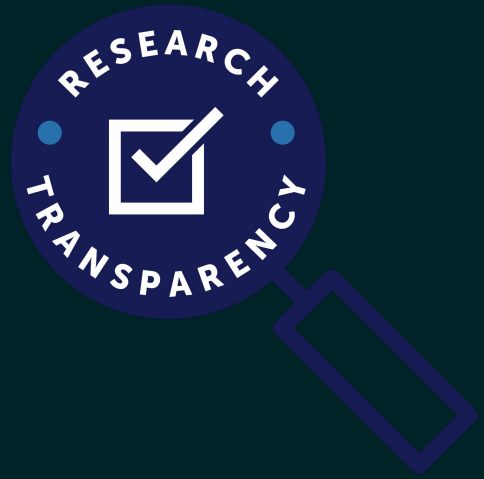
Myth-busting concerns on preregistration



Gabriele Paolacci (RSM)
Julia Rose (ESE)
Antonia Krefeld-Schwalb (RSM)

Erasmus University Rotterdam

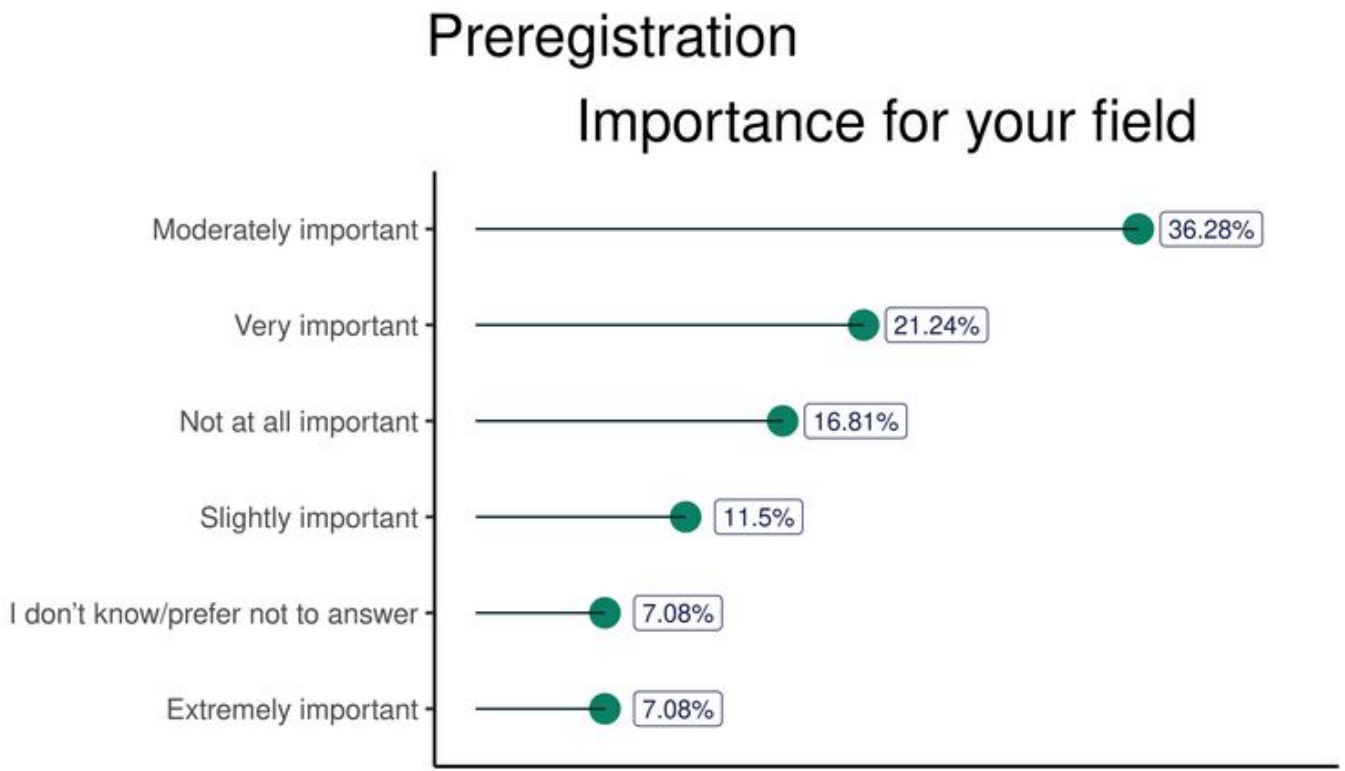
Top 3

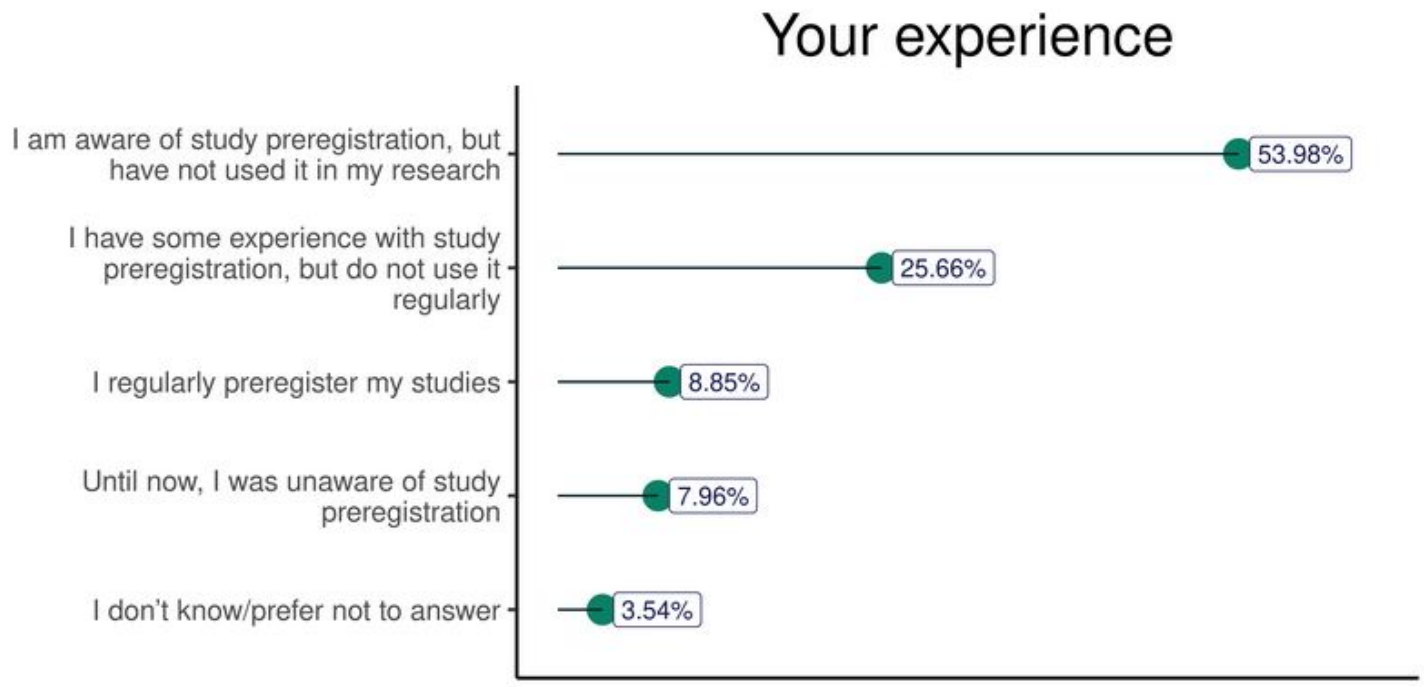
Agenda

Liz Guzman-Ramirez (ESE & RSM)	Introduction to the session
Antonia Krefeld-Schwalb (RSM)	Concern 1 I need to look at my data before I can decide how to best analyse it
Julia Rose (ESE)	Concern 2 Preregistration prevents exploratory research
Gabriele Paolacci (RSM)	Concern 3 It might delay data collection

On April 2021, an Open Science Survey was conducted to RSM & ESE researchers, topics where:

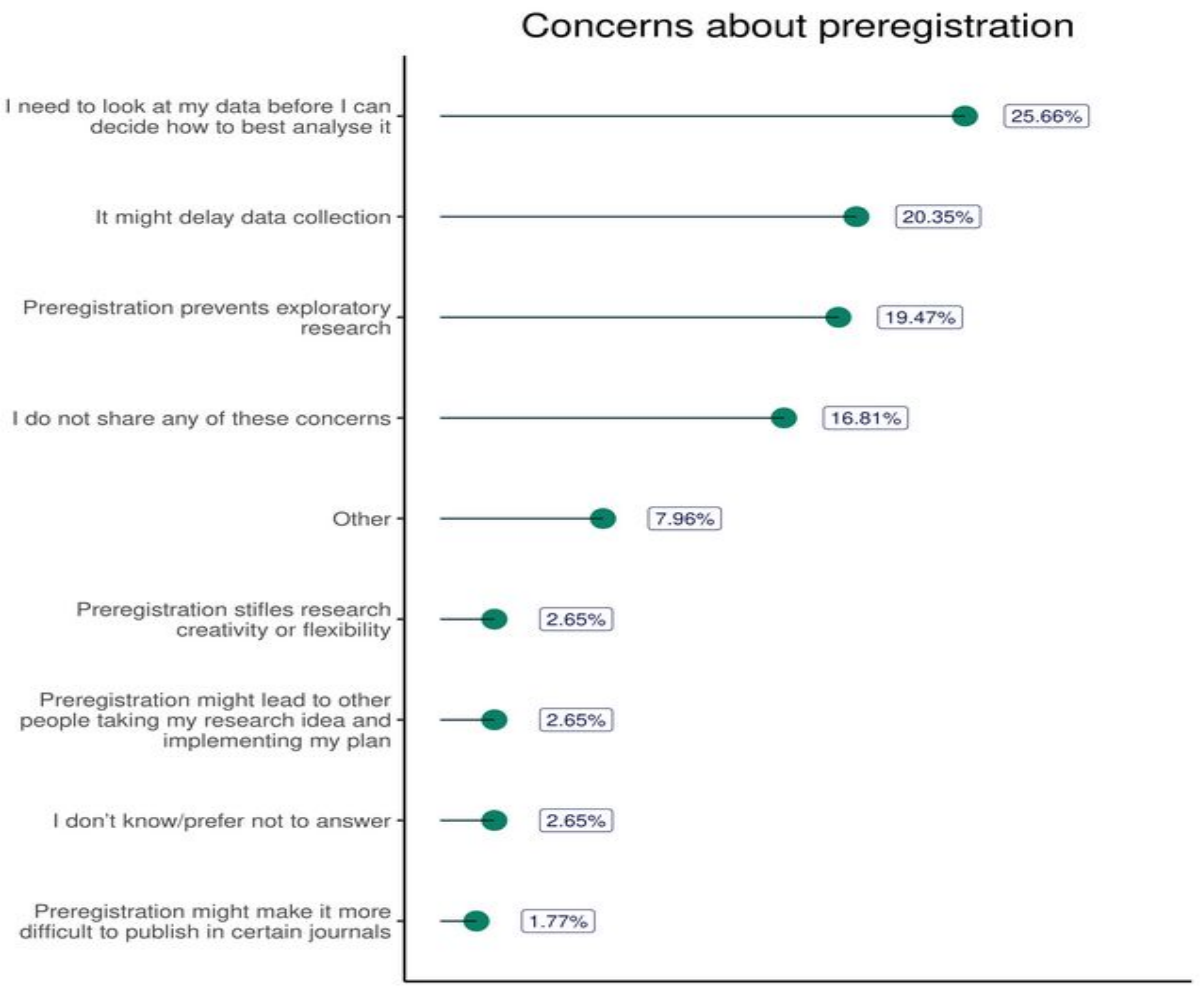
Preregistration
Open materials/code
Open data
Pre-publication archiving
Open access
General experience with open science practices
Awareness of available tools
Role of ERIM


Preregistration
Importance for your field
Moderately important
Very important
Not at all important
Slightly important
I don't know/prefer not to answer
Extremely important
36.28%
21.24%
16.81%
11.5%
7.08%
7.08%


Your experience
I am aware of study preregistration, but have not used it in my research
53.98%
I have some experience with study preregistration, but do not use it regularly
25.66%
I regularly preregister my studies
8.85%
Until now, I was unaware of study preregistration
7.96%
I don't know/prefer not to answer
3.54%

Which of the following open science practices would you like ERIM to provide information or support for?

response	percentage
Preregistration	50%
Open Materials and/or Code	18.42%
Open access publishing	14.47%
I don't know/prefer not to answer	6.58%
Pre-publication archiving	5.26%
Other	2.63%
Open Data	2.63%


Concerns about preregistration
I need to look at my data before I can decide how to best analyse it
25.66%
It might delay data collection
20.35%
Preregistration prevents exploratory research
19.47%
I do not share any of these concerns
16.81%
Other
7.96%
Preregistration stifles research creativity or flexibility
2.65%
Preregistration might lead to other people taking my research idea and implementing my plan
2.65%
I don't know/prefer not to answer
2.65%
Preregistration might make it more difficult to publish in certain journals
1.77%

Agenda

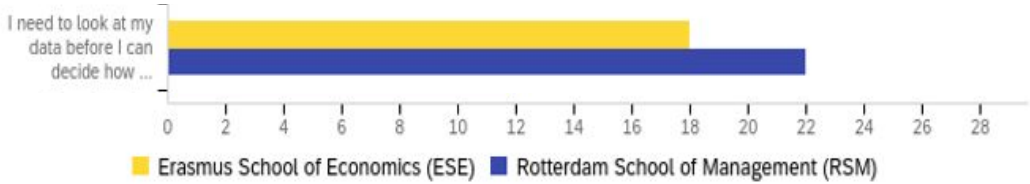
Liz Guzman-Ramirez (ESE & RSM)	Introduction to the session
Antonia Krefeld-Schwalb (RSM)	Concern 1 I need to look at my data before I can decide how to best analyse it
Julia Rose (ESE)	Concern 2 Preregistration prevents exploratory research
Gabriele Paolacci (RSM)	Concern 3 It might delay data collection

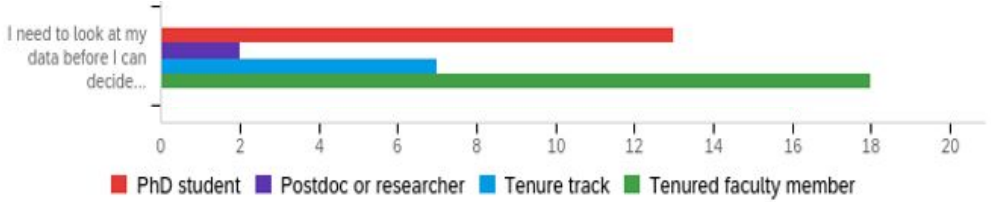
Concern 1

I need to look at my data before I can decide how to best analyse it

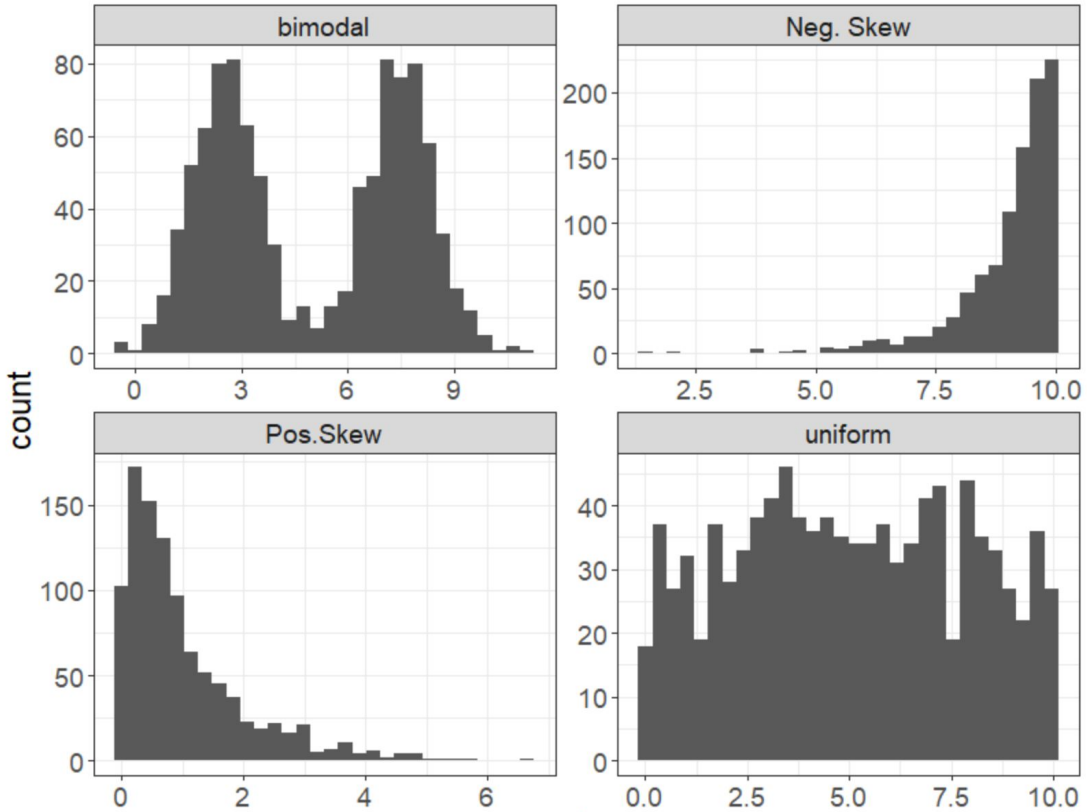
Antonia Krefeld-Schwalb (RSM)

26%


I need to look at my data before I can decide how ...
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28
Erasmus School of Economics (ESE)
Rotterdam School of Management (RSM)


I need to look at my data before I can decide...
0 2 4 6 8 10 12 14 16 18 20
PhD student
Postdoc or researcher
Tenure track
Tenured faculty member

Structuring the data to make it suitable for description and analysis can take a lot of time and effort



Some of the distributions in the data may be non-normal and require transformation for your analysis

How is it possible to know all that before data collection?

Three (not mutually exclusive) Solutions

Simulation

Simulate the expected data, e.g. with Simulation for Factorial Designs • faux)

Additional benefits:

- Statistical knowledge
- Coding skills
- Better informed analysis decisions

Simulate by Design
Simulate Correlated Variables
Simulate from Existing Data
Continuous Predictors
Plotting
Mixed Design Simulation
Randomised Reports
Distribution Conversions
Codebook Demo
Contrasts

Three (not mutually exclusive) Solutions

Simulation

Simulate the expected data, e.g. with Simulation for Factorial Designs • faux)

Additional benefits:

- Statistical knowledge
- Coding skills
- Better informed analysis decisions

Existing Datasets

Consider similar already existing datasets, e.g. your own data or data from relevant references

Additional benefits:

- Understand sources of noise
- Estimate amount of noise
- Identify the shape of the distributions
- Improves analysis decisions

Three (not mutually exclusive) Solutions

Simulation

Simulate the expected data, e.g. with Simulation for Factorial Designs • faux)

Additional benefits:

- Statistical knowledge
- Coding skills
- Better informed analysis decisions

Existing Datasets

Consider similar already existing datasets, e.g. your own data or data from relevant references

Additional benefits:

- Understand sour noise
- Estimate amount of noise
- Identify the shape of the distributions
- Improves analysis decisions

Pilot Study

Collect a very small sample before preregistration

Additional benefits:

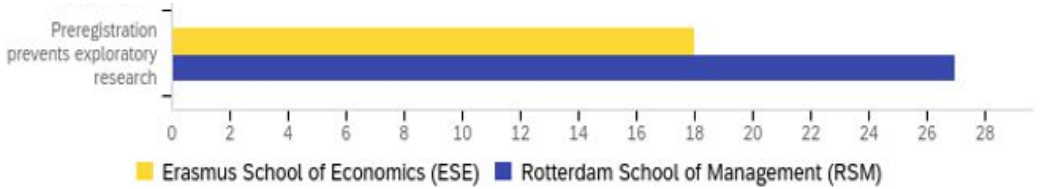
- Understand structure of the data
- Determine pre-processing steps
- Ensure correct data recording
- Write analysis skript to be included in the preregistration

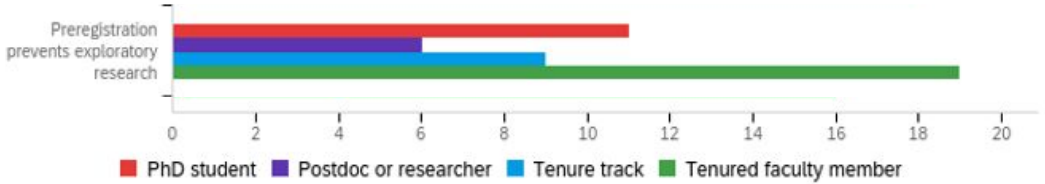
Concern 2

Preregistration prevents exploratory research

Julia Rose (ESE)

20%


Preregistration prevents exploratory research
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28
Erasmus School of Economics (ESE)
Rotterdam School of Management (RSM)


Preregistration prevents exploratory research
0 2 4 6 8 10 12 14 16 18 20
PhD student
Postdoc or researcher
Tenure track
Tenured faculty member

Concern 2: Preregistration prevents exploratory research

It will definitely prevent ***some kinds*** of exploratory research, yes!

Concern 2: Preregistration prevents exploratory research

It will definitely prevent ***some kinds*** of exploratory research, yes!

But let's look into some problematic cases:

CASE 1

Different methods and/or variables are used for different analyses, and the **most "significant" one** is then presented, and the decision on the methods finally included in the paper are ex-post rationalized ("p-hacking").

Concern 2: Preregistration prevents exploratory research

It will definitely prevent ***some kinds*** of exploratory research, yes!

But let's look into some problematic cases:

CASE 2

The main effect that was expected is not found in the overall sample, but it turns out that **some "hand-picked" sub-samples show significant results** (again, results are ex-post rationalized with fitting stories and the paper is sold around those results – "forking")

Concern 2: Preregistration prevents exploratory research

And you might now think:

> “She claims that all exploratory research is bad practise, those are extreme cases, and considered as ‘bad practise’ anyway!“

Concern 2: Preregistration prevents exploratory research

And you might now think:

> “She claims that all exploratory research is bad practise, those are extreme cases, and considered as ‘bad practise’ anyway!”

Yes, I fully agree - those were “extreme” cases.

Concern 2: Preregistration prevents exploratory research

And you might now think:

> "She claims that all exploratory research is bad practise, those are extreme cases, and considered as 'bad practise' anyway!"

Yes, I fully agree - those were "extreme" cases.

That is why there is still plenty of room for exploratory research.
As long as it is openly declared as "exploratory".

Concern 2: Preregistration prevents exploratory research

But then, how can I include exploratory research in my studies after a preregistration?

Concern 2: Preregistration prevents exploratory research

But then, how can I include exploratory research in my studies after a preregistration?

First a few points on "why to do" preregistration...

- It gives a framework to guide the analyses
- All of the analyses that are included should also be included in the final paper
- This prevents, for example, p-hacking and forking
- ...and makes results more trustworthy ***to a certain extent***.

Concern 2: Preregistration prevents exploratory research

But then, how can I include exploratory research in my studies after a preregistration?

First a few points on "why to do" preregistration...

Okay, but **HOW to include exploratory analyses** then?

- ...and makes results more trustworthy *to a certain extent*.

Concern 2: Preregistration prevents exploratory research

How to include exploratory research in a preregistered project:

- In a separate section labelled “exploratory research”
- As an addition to the results obtained, and as an indication for directions for future research to validate those results further

Concern 2: Preregistration prevents exploratory research

My conclusion on exploratory research and preregistration:

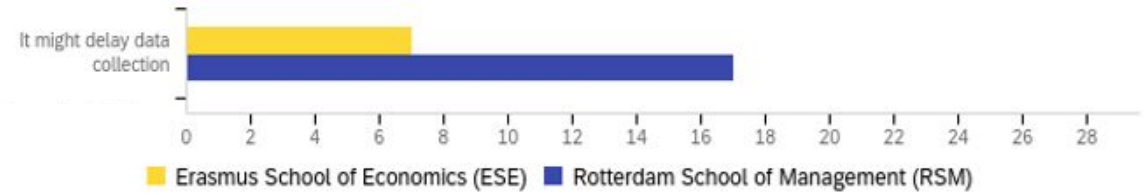
- It does not prevent exploratory research in general
- It adds to the overall transparency of research and the reliability of the results...
- ... as long as preregistration is viewed as one of many tools **and** not as a dogmatic "we can absolutely only include preregistered analyses in the final paper" type of "remedy"

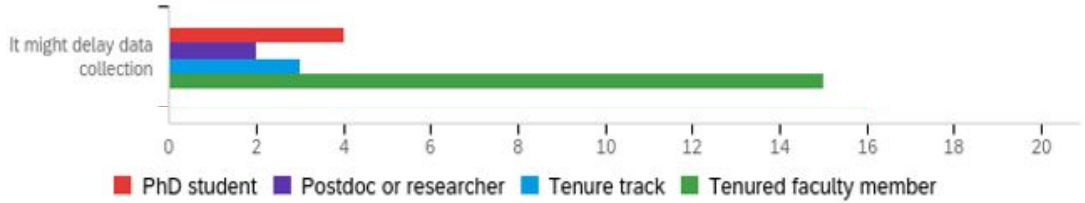
Concern 3

It might delay data collection

Gabriele Paolacci (RSM)

20%


It might delay data collection
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
Erasmus School of Economics (ESE)
Rotterdam School of Management (RSM)


It might delay data collection
0
2
4
6
8
10
12
14
16
18
20
PhD student
Postdoc or researcher
Tenure track
Tenured faculty member

Concern 3:
it might delay data collection

Concern 3:

it might delay data collection

it will delay data collection

Let’s put this in context...

it will delay data collection

Let's put this in context...

COSTS	BENEFITS
it will delay data collection	

Let's put this in context...

COSTS

it will delay data collection

BENEFITS

Let's put this in context...

COSTS

it will delay data collection

BENEFITS

Let's put this in context...

COSTS

it will delay data collection

BENEFITS

truly falsifiable predictions

Let's put this in context...

COSTS

it will delay data collection

BENEFITS

truly falsifiable predictions

more thoughtful research

Let's put this in context...

COSTS

it will delay data collection

BENEFITS

truly falsifiable predictions

more thoughtful research

higher chances of publication

Let's put this in context...

COSTS

it will delay data collection

BENEFITS

truly falsifiable predictions

more thoughtful research

higher chances of publication

reputational benefits

Let's put this in context...

COSTS

it will delay data collection

BENEFITS

truly falsifiable predictions

more thoughtful research

higher chances of publication

reputational benefits

more impact

Questions?