Reproducibility challenges, solutions, and the psychology of analyzing data

Martin Schweinsberg (ESMT Berlin)

National Academy of Sciences Kavly Symposium on the Frontiers of Science (8-12 April 2022 in Irvine, California)

We all analyze data to do science.

POLITICAL COMMUNICATION
2020, VOL. 37, NO. 4, 447–467
https://doi.org/10.1080/10584609.2020.1718257

Routledge
Taylor & Francis Group

Troll Factories: Manufacturing Specialized Disinformation on Twitter

Darren L. Linvill[a] and Patrick L. Warren[b]

[a]Department of Communication, Clemson University, Clemson, South Carolina, USA; [b]John E. Walker Department of Economics, Clemson University, Clemson, South Carolina, USA

ABSTRACT
We document methods employed by Russia's Internet Research Agency to influence the political agenda of the United States from September 9, 2009 to June 21, 2018. We qualitatively and quantitatively analyze Twitter accounts with known IRA affiliation to better understand the form and function of Russian efforts. We identified five handle categories: *Right Troll, Left Troll, News Feed, Hashtag Gamer*, and *Fearmonger*. Within each type, accounts were used consistently, but the behavior across types was different, both in terms of "normal" daily behavior and in how they responded to external events. In this sense, the Internet Research Agency's agenda-building effort was "industrial" – mass produced from a system of interchangeable parts, where each class of part fulfilled a specialized function.

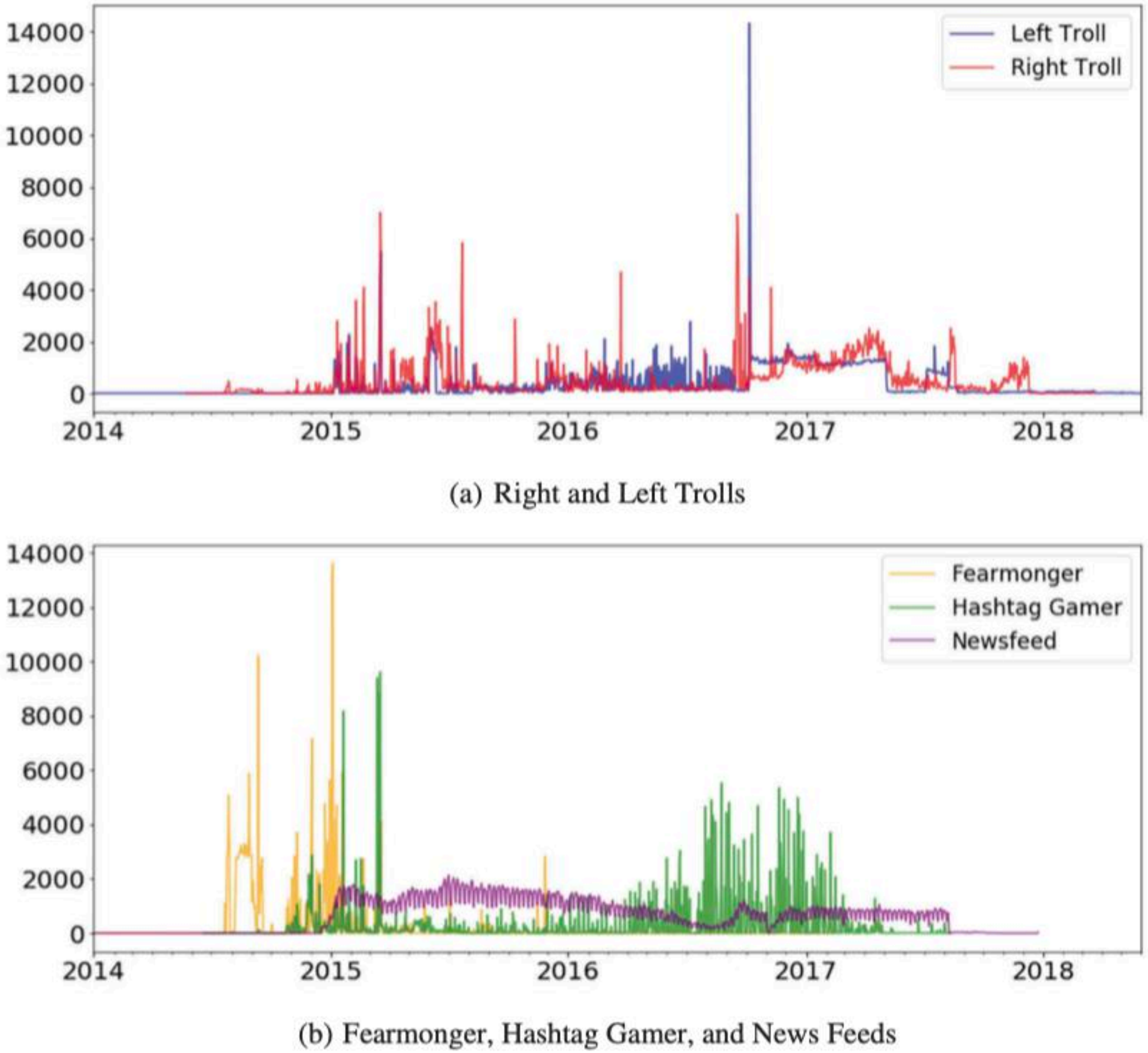
KEYWORDS
Twitter; social media; disinformation; Internet Research Agency

In February 2018, the U.S. Justice Department indicted 13 Russian nationals for interference with the 2016 U.S. Presidential election (The United States of America v. Internet Research Agency LLC, 2018). The indictment named the Internet Research Agency (IRA), based in St. Petersburg, as central to a Russian effort to sow discord in the U.S. political system, largely through social media. The IRA intervened in the 2016 election (Mueller, 2019), with some even suggesting they may have tipped the balance of the election in favor of candidate Donald Trump (Jamieson, 2018). Regardless of the effect on the election, however, it is undeniable that many real people were taken in by IRA messaging, including mainstream journalists (Lukito & Wells, 2018) and political activists (Birnbaum, 2019).

Researchers have moved to try to understand the strategy and impact of what is, perhaps, the most important foreign influence operation of the social media age. Concentrating on the discussions on Twitter around the Black Lives Matter movement, Arif, Stewart, and Starbird (2018) showed that the IRA fostered antagonism and undermined trust in authorities. Looking at accounts discussing vaccines, Broniatowski et al. (2018) showed that IRA trolls amplified both sides of the contentious debate. In the context of the Twitter discussion of the Malaysian Airlines flight (MH17) downed in eastern Ukraine, Golovchnko, Hartmann, and Adler-Nissen (2018) showed that the IRA appeared in the conversation but had no substantial effects on its progress. Finally, Badawy, Lerman, and Ferrara (2018) investigated what sort of accounts shared the content the IRA produced, in the context of the 2016 U.S. presidential election, finding that more conservative and more "bot like" accounts, with fewer followers but more status updates, were more likely to share IRA content.

CONTACT Darren L. Linvill darrenl@clemson.edu Department of Communication, Clemson University, Clemson, SC, USA.

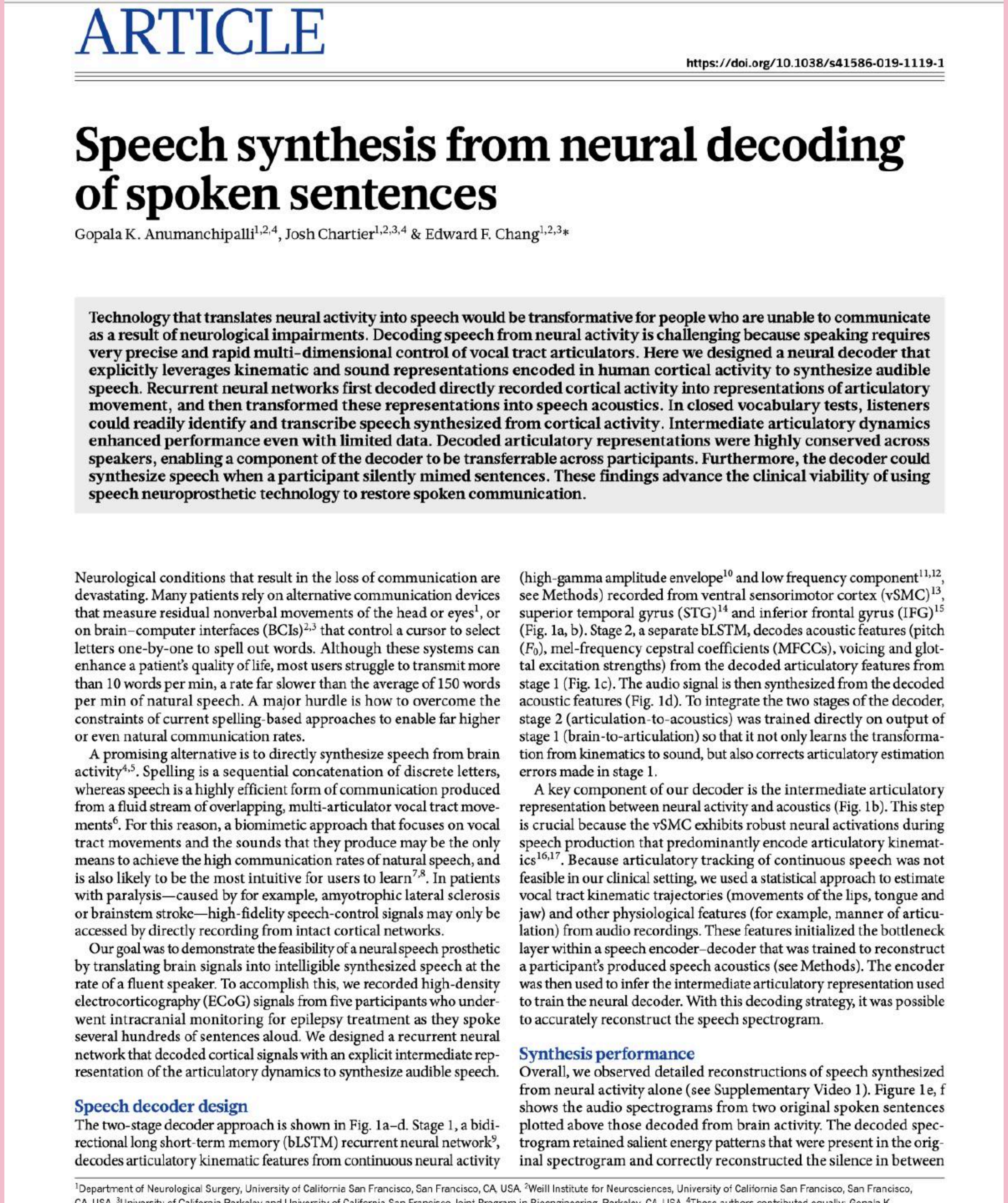
Linvill, D. L., & Warren, P. L. (2020). Troll factories: Manufacturing specialized disinformation on Twitter. Political Communication, 37(4), 447-467.



igure 2. Daily tweets by English-language accounts, by account type, January 1, 2014–June 21, 201

Linvill, D. L., & Warren, P. L. (2020). Troll factories: Manufacturing specialized disinformation on Twitter. Political Communication, 37(4), 447-467.

>9 million tweets


ARTICLE

https://doi.org/10.1038/s41586-019-1119-1

Speech synthesis from neural decoding of spoken sentences

Gopala K. Anumanchipalli[1,2,4], Josh Chartier[1,2,3,4] & Edward F. Chang[1,2,3]*

Technology that translates neural activity into speech would be transformative for people who are unable to communicate as a result of neurological impairments. Decoding speech from neural activity is challenging because speaking requires very precise and rapid multi-dimensional control of vocal tract articulators. Here we designed a neural decoder that explicitly leverages kinematic and sound representations encoded in human cortical activity to synthesize audible speech. Recurrent neural networks first decoded directly recorded cortical activity into representations of articulatory movement, and then transformed these representations into speech acoustics. In closed vocabulary tests, listeners could readily identify and transcribe speech synthesized from cortical activity. Intermediate articulatory dynamics enhanced performance even with limited data. Decoded articulatory representations were highly conserved across speakers, enabling a component of the decoder to be transferrable across participants. Furthermore, the decoder could synthesize speech when a participant silently mimed sentences. These findings advance the clinical viability of using speech neuroprosthetic technology to restore spoken communication.

Neurological conditions that result in the loss of communication are devastating. Many patients rely on alternative communication devices that measure residual nonverbal movements of the head or eyes[1], or on brain–computer interfaces (BCIs)[2,3] that control a cursor to select letters one-by-one to spell out words. Although these systems can enhance a patient's quality of life, most users struggle to transmit more than 10 words per min, a rate far slower than the average of 150 words per min of natural speech. A major hurdle is how to overcome the constraints of current spelling-based approaches to enable far higher or even natural communication rates.

A promising alternative is to directly synthesize speech from brain activity[4,5]. Spelling is a sequential concatenation of discrete letters, whereas speech is a highly efficient form of communication produced from a fluid stream of overlapping, multi-articulator vocal tract movements[6]. For this reason, a biomimetic approach that focuses on vocal tract movements and the sounds that they produce may be the only means to achieve the high communication rates of natural speech, and is also likely to be the most intuitive for users to learn[7,8]. In patients with paralysis—caused by for example, amyotrophic lateral sclerosis or brainstem stroke—high-fidelity speech-control signals may only be accessed by directly recording from intact cortical networks.

Our goal was to demonstrate the feasibility of a neural speech prosthetic by translating brain signals into intelligible synthesized speech at the rate of a fluent speaker. To accomplish this, we recorded high-density electrocorticography (ECoG) signals from five participants who underwent intracranial monitoring for epilepsy treatment as they spoke several hundreds of sentences aloud. We designed a recurrent neural network that decoded cortical signals with an explicit intermediate representation of the articulatory dynamics to synthesize audible speech.

Speech decoder design

The two-stage decoder approach is shown in Fig. 1a–d. Stage 1, a bidirectional long short-term memory (bLSTM) recurrent neural network[9], decodes articulatory kinematic features from continuous neural activity (high-gamma amplitude envelope[10] and low frequency component[11,12], see Methods) recorded from ventral sensorimotor cortex (vSMC)[13], superior temporal gyrus (STG)[14] and inferior frontal gyrus (IFG)[15] (Fig. 1a, b). Stage 2, a separate bLSTM, decodes acoustic features (pitch (F_0), mel-frequency cepstral coefficients (MFCCs), voicing and glottal excitation strengths) from the decoded articulatory features from stage 1 (Fig. 1c). The audio signal is then synthesized from the decoded acoustic features (Fig. 1d). To integrate the two stages of the decoder, stage 2 (articulation-to-acoustics) was trained directly on output of stage 1 (brain-to-articulation) so that it not only learns the transformation from kinematics to sound, but also corrects articulatory estimation errors made in stage 1.

A key component of our decoder is the intermediate articulatory representation between neural activity and acoustics (Fig. 1b). This step is crucial because the vSMC exhibits robust neural activations during speech production that predominantly encode articulatory kinematics[16,17]. Because articulatory tracking of continuous speech was not feasible in our clinical setting, we used a statistical approach to estimate vocal tract kinematic trajectories (movements of the lips, tongue and jaw) and other physiological features (for example, manner of articulation) from audio recordings. These features initialized the bottleneck layer within a speech encoder–decoder that was trained to reconstruct a participant's produced speech acoustics (see Methods). The encoder was then used to infer the intermediate articulatory representation used to train the neural decoder. With this decoding strategy, it was possible to accurately reconstruct the speech spectrogram.

Synthesis performance

Overall, we observed detailed reconstructions of speech synthesized from neural activity alone (see Supplementary Video 1). Figure 1e, f shows the audio spectrograms from two original spoken sentences plotted above those decoded from brain activity. The decoded spectrogram retained salient energy patterns that were present in the original spectrogram and correctly reconstructed the silence in between

[1]Department of Neurological Surgery, University of California San Francisco, San Francisco, CA, USA. [2]Weill Institute for Neurosciences, University of California San Francisco, San Francisco,

Anumanchipalli, G. K., Chartier, J., & Chang, E. F. (2019). Speech synthesis from neural decoding of spoken sentences. Nature, 568(7753), 493-498. https://doi.org/10.1038/s41586-019-1119-1

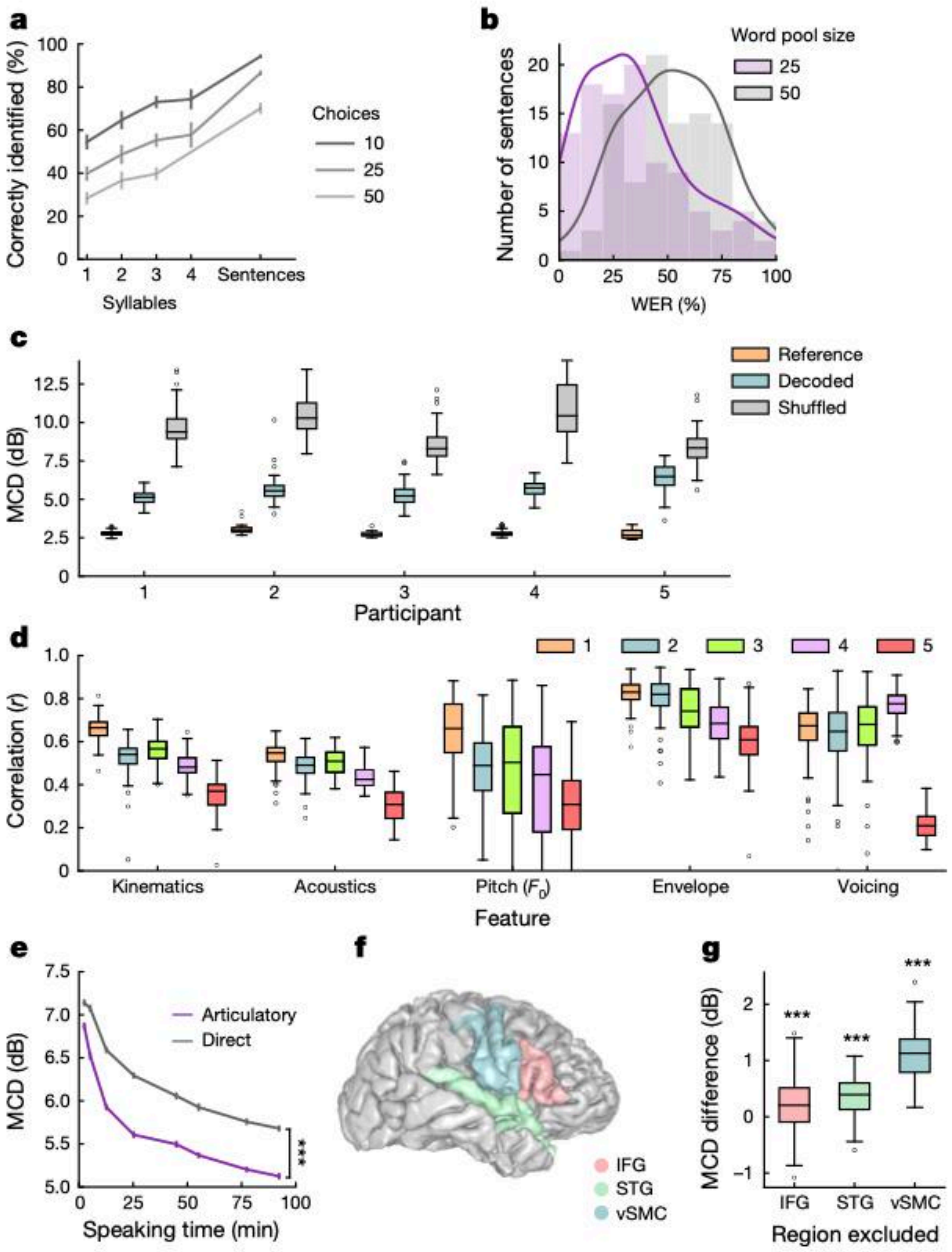


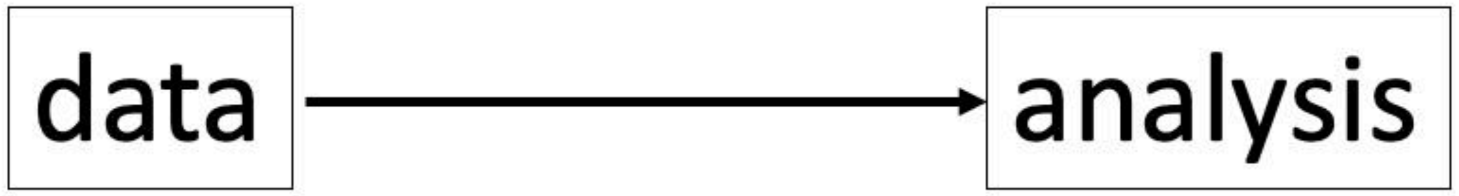
Fig. 2 | Synthesized speech intelligibility and feature-specific

Anumanchipalli, G. K., Chartier, J., & Chang, E. F. (2019). Speech synthesis from neural decoding of spoken sentences. Nature, 568(7753), 493-498. https://doi.org/10.1038/s41586-019-1119-1

"(...)high-density electrocorticography (ECoG) signals from five participants who underwent intracranial monitoring for epilepsy treatment as they spoke several hundreds of sentences aloud(...).

Data analysis is the common language we share to advance our respective fields.

data


data
analysis


data
analysis
results


data
analysis
results

Journal of Experimental Social Psychology 48 (2012) 226–231

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

ELSEVIER

Reports

Starting high and ending with nothing: The role of anchors and power in negotiations

Martin Schweinsberg [a,*], Gillian Ku [a], Cynthia S. Wang [b,c], Madan M. Pillutla [a]

[a] London Business School, UK
[b] University of Michigan, National Center for Institutional Diversity and Ross School of Business, USA
[c] National University of Singapore, NUS Business School, Singapore

ARTICLE INFO

Article history:
Received 13 September 2010
Revised 4 July 2011
Available online 23 July 2011

Keywords:
Anchor
First offer
Impasse
Negotiation

ABSTRACT

Most research suggests that negotiators gain value by making first offers in negotiations. The current research examines the proposition that extreme first offers offend their recipients and cause them to walk away, resulting in an impasse. Results across two experiments support this proposition. As a result, extreme offers can be risky: even though they can anchor counteroffers and final outcomes, bringing benefit to the offerer, they only do so when impasses are avoided. In addition, we find support for the proposition that power moderates the relationship between extreme offers and impasses: although low- and high-power negotiators are equally offended by extreme offers, it is the low-power negotiators who walk away from the negotiation.

Introduction

Negotiations research has repeatedly shown that final prices are positively correlated with first offers — the more a seller asks for, the higher the final price; the less a buyer offers, the lower the final price (Chertkoff & Conley, 1967; Galinsky & Mussweiler, 2001). As a result, negotiators are often advised to make aggressive first offers i.e., high for sellers and low for buyers (Thompson, 2008).

However, practitioners often refrain from making extreme offers (i.e., unreasonably high for sellers and unreasonably low for buyers) because they are concerned that such offers will offend counterparts and cause them to walk away. These intuitions are consistent with findings showing that impasses are not uncommon, especially if one party is offended by the other (c.f., Pillutla & Murnighan, 1996; Van Kleef, 2010). In the current research, we empirically examine the proposition that extreme first offers will result in more impasses. Additionally, we explore how negotiators' relative power may affect their reactions to first offers by examining whether low- or high-power negotiators are more likely to walk away from extreme offers and why.

First offers in negotiations

A well-established explanation for the positive correlation between first offers and final negotiation outcomes is the anchoring and insufficient adjustment heuristic (Tversky & Kahneman, 1974), which suggests that individuals make estimations by starting with an initial value (the anchor) and adjusting (insufficiently) away from it until a final estimate is reached. Thus, when sellers make aggressive first offers, buyers anchor on the high offers and adjust their value estimations insufficiently (Mussweiler & Strack, 1999), leading to higher counteroffers and final outcomes than if the sellers had made lower first offers (Galinsky & Mussweiler, 2001).

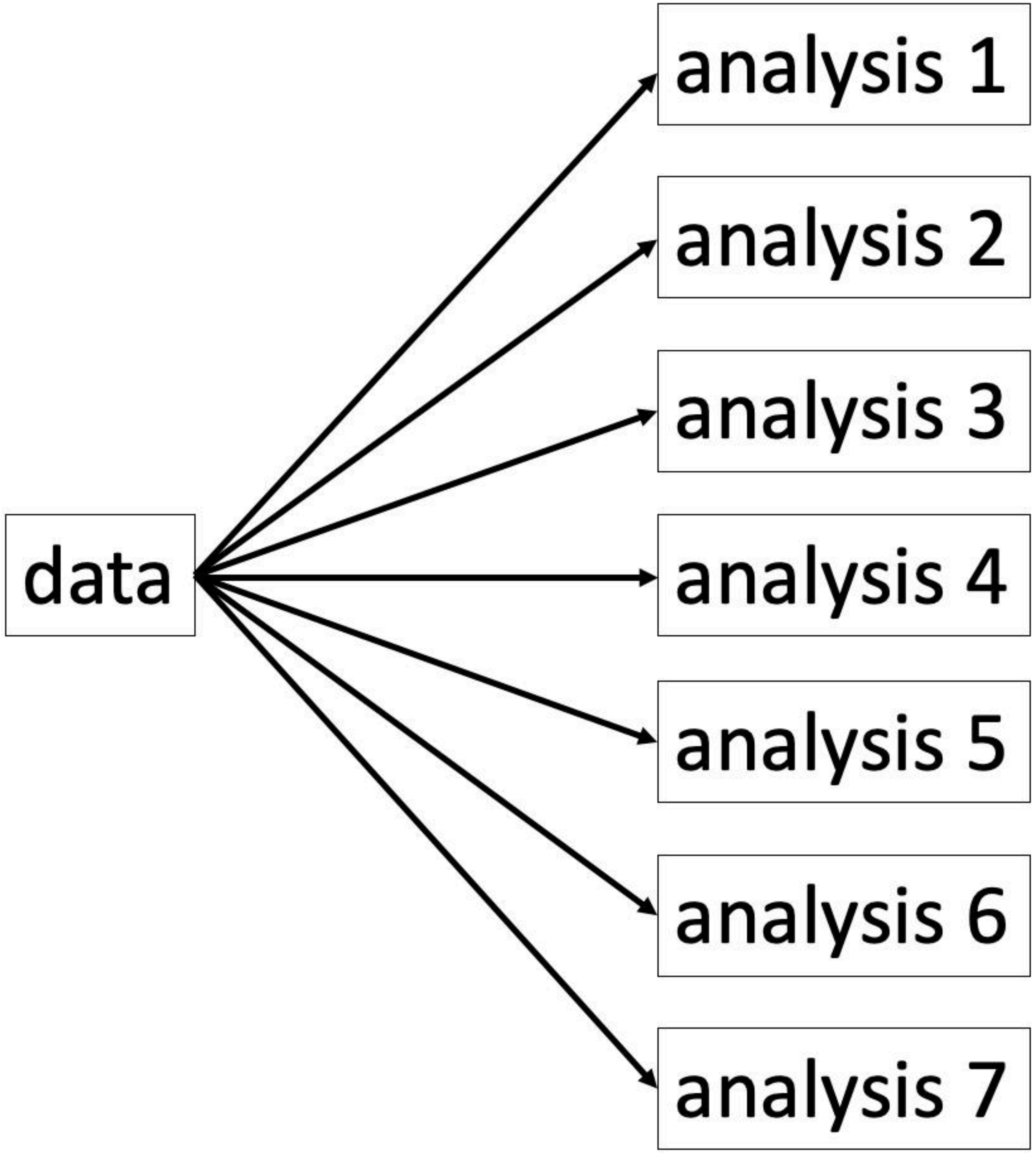
The effect of anchors on judgments and behavior is extensive and robust, extending from numerical estimates (Northcraft & Neale, 1987; Tversky & Kahneman, 1974) to judicial verdicts (Englich & Mussweiler, 2001) and self-perceptions (Gilovich, Medvec, & Savitsky, 2000). Anchors affect judgments even when individuals know that the anchor is randomly-generated (e.g., Tversky & Kahneman, 1974) and implausibly low or high. For instance, when asked to estimate Gandhi's age at his death, participants were biased by the anchors of 9 and 140 years, even though these fantastic anchors were clearly implausible (Strack & Mussweiler, 1997).

Extrapolating these findings to negotiations is problematic because judgments and behaviors made in interpersonal contexts such as negotiations may differ from the impersonal contexts examined in earlier studies. In negotiations, extreme offers may offend recipients because they violate norms of appropriate behavior (c.f., Pillutla & Murnighan, 1996) and may provoke recipients to walk away from the negotiation (c.f., Brooks & Schweitzer, 2011). Extreme anchors may therefore decrease value to negotiators if their counterparts walk away, leaving the negotiator without a deal.

Previous research (e.g., Galinsky & Mussweiler, 2001) has failed to examine impasses caused by first offers, probably because the typical negotiation experiment conceals the risk of impasses and artificially inflates agreement rates (Galinsky, Ku, & Mussweiler, 2009).

* Corresponding author at: Organisational Behaviour Department, London Business School, Regent's Park, London NW1 4SA, UK.
E-mail address: martins.phd2008@london.edu (M. Schweinsberg).


data
analysis
results



What would be the result(s)???

When many analysts analyze the same data set and test the same hypotheses, they report different results, at times even in the opposite direction.

Silberzahn et al. (2018):

- 29 teams of analysts

- all analysts analyzed same dataset on soccer

- to answer same research question: *are soccer referees more likely to give red cards to dark-skin-toned players than to light-skin-toned players?*

Empirical Article

ASSOCIATION FOR PSYCHOLOGICAL SCIENCE
Advances in Methods and Practices in Psychological Science
2018, Vol. 1(3) 337–356

DOI: 10.1177/2515245917747646
www.psychologicalscience.org/AMPPS
SAGE

Many Analysts, One Data Set: Making Transparent How Variations in Analytic Choices Affect Results

R. Silberzahn[1], E. L. Uhlmann[2], D. P. Martin[3], P. Anselmi[4], F. Aust[5], E. Awtrey[6], Š. Bahník[7], F. Bai[8], C. Bannard[9], E. Bonnier[10], R. Carlsson[11], F. Cheung[12], G. Christensen[13], R. Clay[14], M. A. Craig[15], A. Dalla Rosa[4], L. Dam[16], M. H. Evans[17], I. Flores Cervantes[18], N. Fong[19], M. Gamez-Djokic[20], A. Glenz[21], S. Gordon-McKeon[22], T. J. Heaton[23], K. Hederos[24], M. Heene[25], A. J. Hofelich Mohr[26], F. Högden[5], K. Hui[27], M. Johannesson[10], J. Kalodimos[28], E. Kaszubowski[29], D. M. Kennedy[30], R. Lei[15], T. A. Lindsay[26], S. Liverani[31], C. R. Madan[32], D. Molden[33], E. Molleman[16], R. D. Morey[34], L. B. Mulder[16], B. R. Nijstad[16], N. G. Pope[35], B. Pope[36], J. M. Prenoveau[37], F. Rink[16], E. Robusto[4], H. Roderique[38], A. Sandberg[24], E. Schlüter[39], F. D. Schönbrodt[25], M. F. Sherman[37], S. A. Sommer[40], K. Sotak[41], S. Spain[42], C. Spörlein[43], T. Stafford[44], L. Stefanutti[4], S. Tauber[16], J. Ullrich[21], M. Vianello[4], E.-J. Wagenmakers[45], M. Witkowiak[46], S. Yoon[19], and B. A. Nosek[3,47]

[1]Organisational Behaviour, University of Sussex Business School; [2]Organisational Behaviour Area, INSEAD Asia Campus; [3]Department of Psychology, University of Virginia; [4]Department of Philosophy, Sociology, Education and Applied Psychology, University of Padua; [5]Department of Psychology, University of Cologne; [6]Department of Management, University of Cincinnati; [7]Department of Management, Faculty of Business Administration, University of Economics, Prague; [8]Department of Management and Marketing, Hong Kong Polytechnic University; [9]Department of Psychology, University of Liverpool; [10]Department of Economics, Stockholm School of Economics; [11]Department of Psychology, Linnaeus University; [12]School of Public Health, University of Hong Kong; [13]Berkeley Institute for Data Science, University of California, Berkeley; [14]Department of Psychology, College of Staten Island, City University of New York; [15]Department of Psychology, New York University; [16]Faculty of Economics and Business, University of Groningen; [17]Division of Neuroscience and Experimental Psychology, University of Manchester; [18]Westat, Rockville, Maryland; [19]Department of Marketing and Supply Chain Management, Temple University; [20]Department of Management and Organizations, Kellogg School of Management, Northwestern University; [21]Department of Psychology, University of Zurich; [22]Washington, D.C.; [23]School of Mathematics and Statistics, University of Sheffield; [24]Swedish Institute for Social Research (SOFI), Stockholm University; [25]Department of Psychology, Ludwig-Maximilians-Universität München; [26]College of Liberal Arts, University of Minnesota; [27]School of Management, Xiamen University; [28]College of Business, Oregon State University; [29]Department of Psychology, Federal University of Santa Catarina; [30]School of Business, University of Washington Bothell; [31]School of Mathematical Sciences, Queen Mary University of London; [32]School of Psychology, University of Nottingham; [33]Department of Psychology, Northwestern University; [34]School of Psychology, Cardiff University; [35]Department of Economics, University of Maryland; [36]Department of Economics, Brigham Young University; [37]Department of Psychology, Loyola University Maryland; [38]Rotman School of Management, University of Toronto; [39]Department of Social Sciences and Cultural Studies, Institute of Sociology, Justus Liebig University, Giessen; [40]United States Military Academy at West Point; [41]Department of Marketing and Management, SUNY Oswego; [42]John Molson School of Business, Concordia University; [43]Lehrstuhl für Soziologie, insb. Sozialstrukturanalyse, Otto-Friedrich-Universität Bamberg; [44]Department of Psychology, University of Sheffield; [45]Department of Psychological Methods, University of Amsterdam; [46]Poznań, Poland; and [47]Center for Open Science, Charlottesville, Virginia

Corresponding Authors:
R. Silberzahn, University of Sussex Business School, Jubilee Building, Brighton BN1 9SL, United Kingdom
E-mail: r.silberzahn@gmail.com

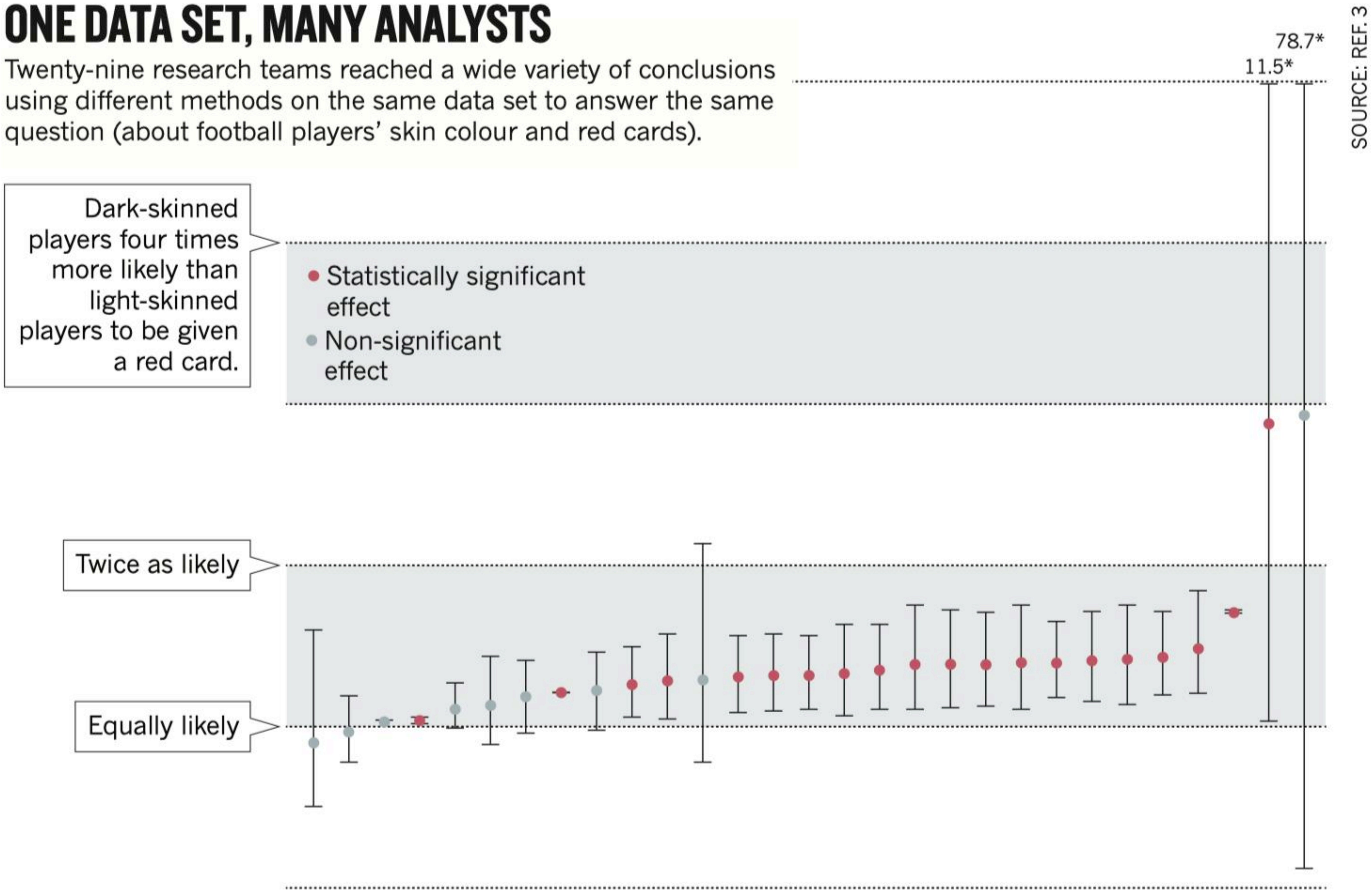
E. L. Uhlmann, INSEAD, Organisational Behaviour Area, 1 Ayer Rajah Ave., 138676 Singapore
E-mail: eric.luis.uhlmann@gmail.com

D. P. Martin, University of Virginia, Department of Psychology, 918 Emmet St. N, Charlottesville, VA 22903
E-mail: dpmartin42@gmail.com

B. A. Nosek, Center for Open Science, 210 Ridge McIntire Rd., Suite 500, Charlottesville, VA 22903-5083
E-mail: nosek@virginia.edu

ONE DATA SET, MANY ANALYSTS

Twenty-nine research teams reached a wide variety of conclusions using different methods on the same data set to answer the same question (about football players' skin colour and red cards).



Point estimates and 95% confidence intervals. *Truncated upper bounds.

Article

Variability in the analysis of a single neuroimaging dataset by many teams

https://doi.org/10.1038/s41586-020-2314-9

Received: 14 November 2019

Accepted: 7 April 2020

Published online: 20 May 2020

Check for updates

A list of authors and affiliations appears in the online version of the paper.

Data analysis workflows in many scientific domains have become increasingly complex and flexible. Here we assess the effect of this flexibility on the results of functional magnetic resonance imaging by asking 70 independent teams to analyse the same dataset, testing the same 9 ex-ante hypotheses[1]. The flexibility of analytical approaches is exemplified by the fact that no two teams chose identical workflows to analyse the data. This flexibility resulted in sizeable variation in the results of hypothesis tests, even for teams whose statistical maps were highly correlated at intermediate stages of the analysis pipeline. Variation in reported results was related to several aspects of analysis methodology. Notably, a meta-analytical approach that aggregated information across teams yielded a significant consensus in activated regions. Furthermore, prediction markets of researchers in the field revealed an overestimation of the likelihood of significant findings, even by researchers with direct knowledge of the dataset[2–5]. Our findings show that analytical flexibility can have substantial effects on scientific conclusions, and identify factors that may be related to variability in the analysis of functional magnetic resonance imaging. The results emphasize the importance of validating and sharing complex analysis workflows, and demonstrate the need for performing and reporting multiple analyses of the same data. Potential approaches that could be used to mitigate issues related to analytical variability are discussed.

Data analysis workflows in many areas of science have a large number of analysis steps that involve many possible choices (that is, "researcher degrees of freedom"[6,7]). Simulation studies show that variability in analytical choices can have substantial effects on results[8], but its degree and effect in practice is unclear. Recent work in psychology addressed this through a "many analysts" approach[9], in which the same dataset was analysed by a large number of groups, uncovering substantial variability in behavioural results across analysis teams. In the Neuroimaging Analysis Replication and Prediction Study (NARPS), we applied a similar approach to the domain of functional magnetic resonance imaging (fMRI), the analysis workflows of which are complex and highly variable. Our goal was to assess—with the highest possible ecological validity—the degree and effect of analytical flexibility on fMRI results in practice. In addition, we estimated the beliefs of researchers in the field regarding the degree of variability in analysis outcomes using prediction markets to test whether peers in the field could predict the results[2–5].

Variability of results across teams

The first aim of NARPS was to assess the real-world variability of results across independent teams analysing the same dataset. The dataset included fMRI data from 108 individuals, each performing one of two versions of a task that was previously used to study decision-making under risk[10]. The two versions were designed to address a debate on the effect of gain and loss distributions on neural activity in this task[10–12]. A full description of the dataset is available in a Data Descriptor[1]; the dataset is openly available at https://doi.org/10.18112/openneuro.ds001734.v1.0.4.

Seventy teams (69 of whom had previous fMRI publications) were provided with the raw data, and an optional preprocessed version of the dataset (with fMRIPrep[13]). They were asked to analyse the data to test nine ex-ante hypotheses (Extended Data Table 1), each consisting of a description of activity in a specific brain region in relation to a particular feature of the task. They were given up to 100 days to report whether each hypothesis was supported on the basis of a whole-brain-corrected analysis (yes or no). In addition, each team submitted a detailed report of the methods of analysis that they had used, together with unthresholded and thresholded statistical maps supporting each hypothesis test (Extended Data Tables 2, 3a). To perform an ecologically valid study testing the sources of variability that contribute to published literature 'in the wild', the instructions to the teams were as minimal as possible. The only instructions were to perform the analysis as they usually would in their own research laboratory and report the binary decision on the basis of their own criteria for a whole-brain-corrected result for the specific region described in the hypothesis. The dataset, reports and collections were kept private until after the prediction markets were closed.

Overall, the rates of reported significant findings varied across hypotheses (Fig. 1, Extended Data Table 1). Only one hypothesis (hypothesis 5) showed a high rate of significant findings (84.3%), whereas three other hypotheses showed consistent non-significant findings across teams (5.7% significant findings). For the remaining five hypotheses, the results were variable, with 21.4% to 37.1% of teams reporting a significant result. The extent of the variation in results across teams was quantified by the fraction of teams that reported a result different from the majority of teams (that is, the absolute distance from consensus). On average across the 9 hypotheses, 20% of teams

Bovotnik-Rezer et al. (2020):

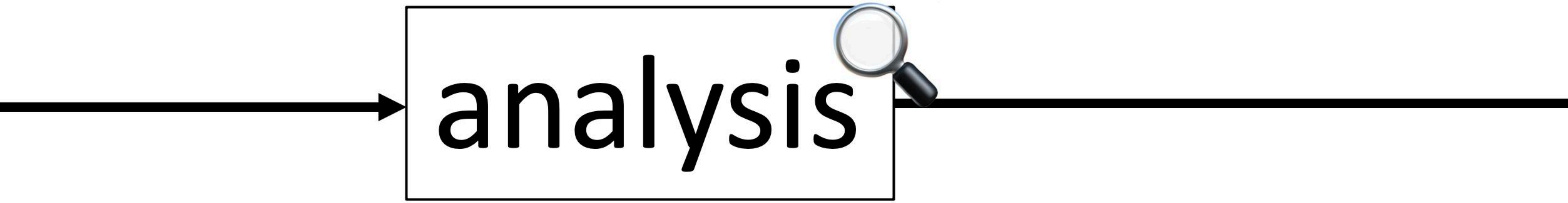
- 70 teams tested 9 hypotheses on functional magnetic resonance images

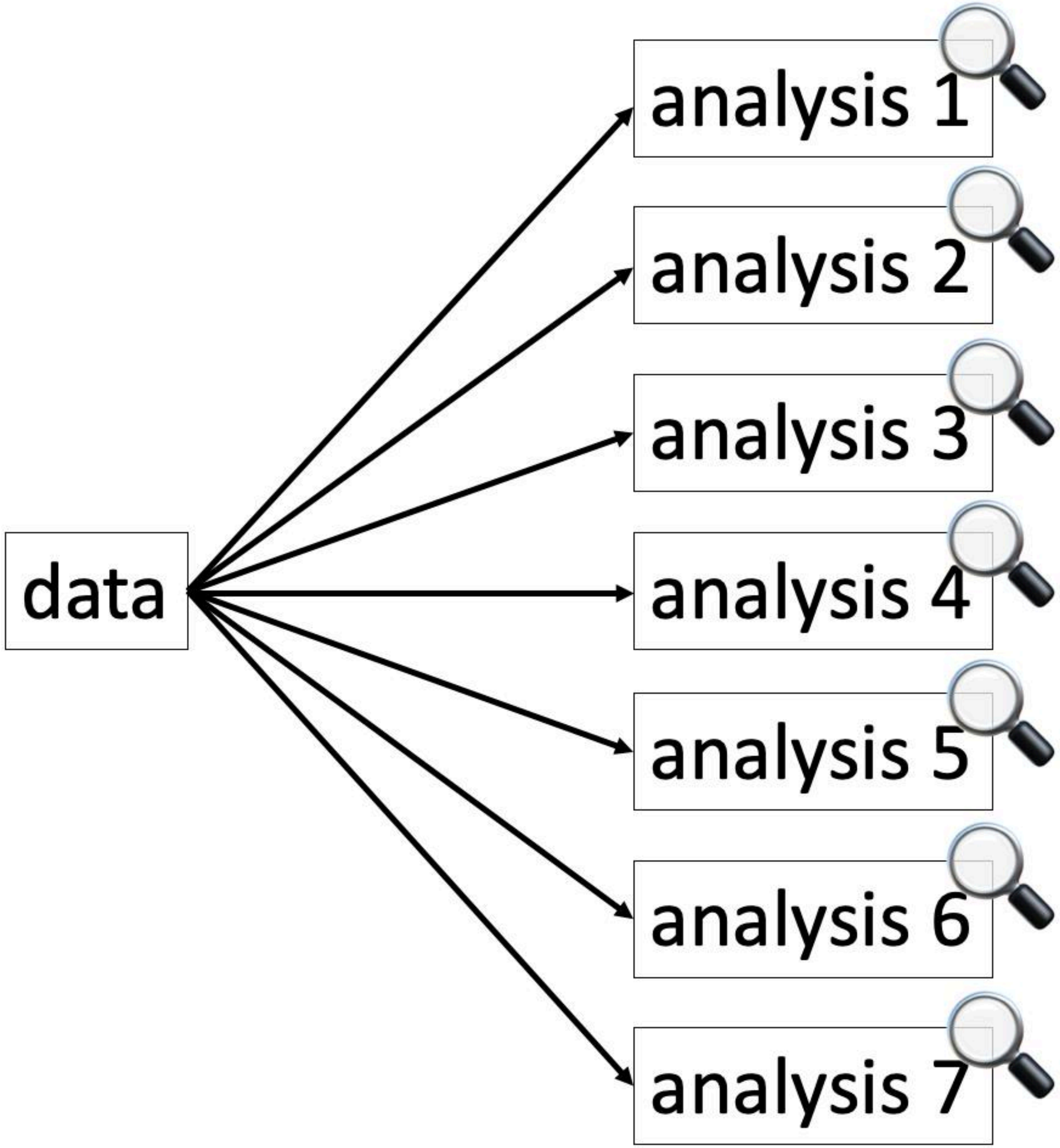
- no 2 teams chose same analytic approach

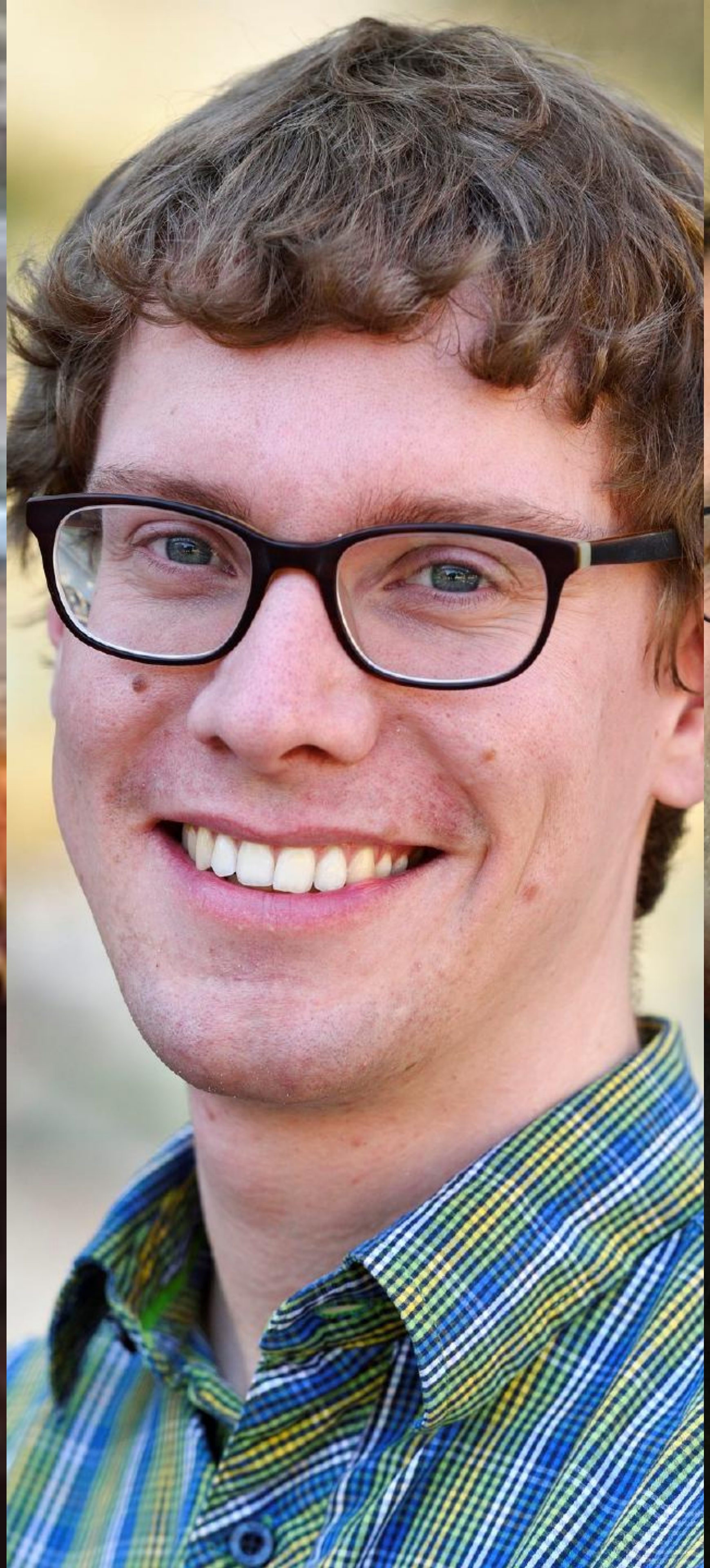
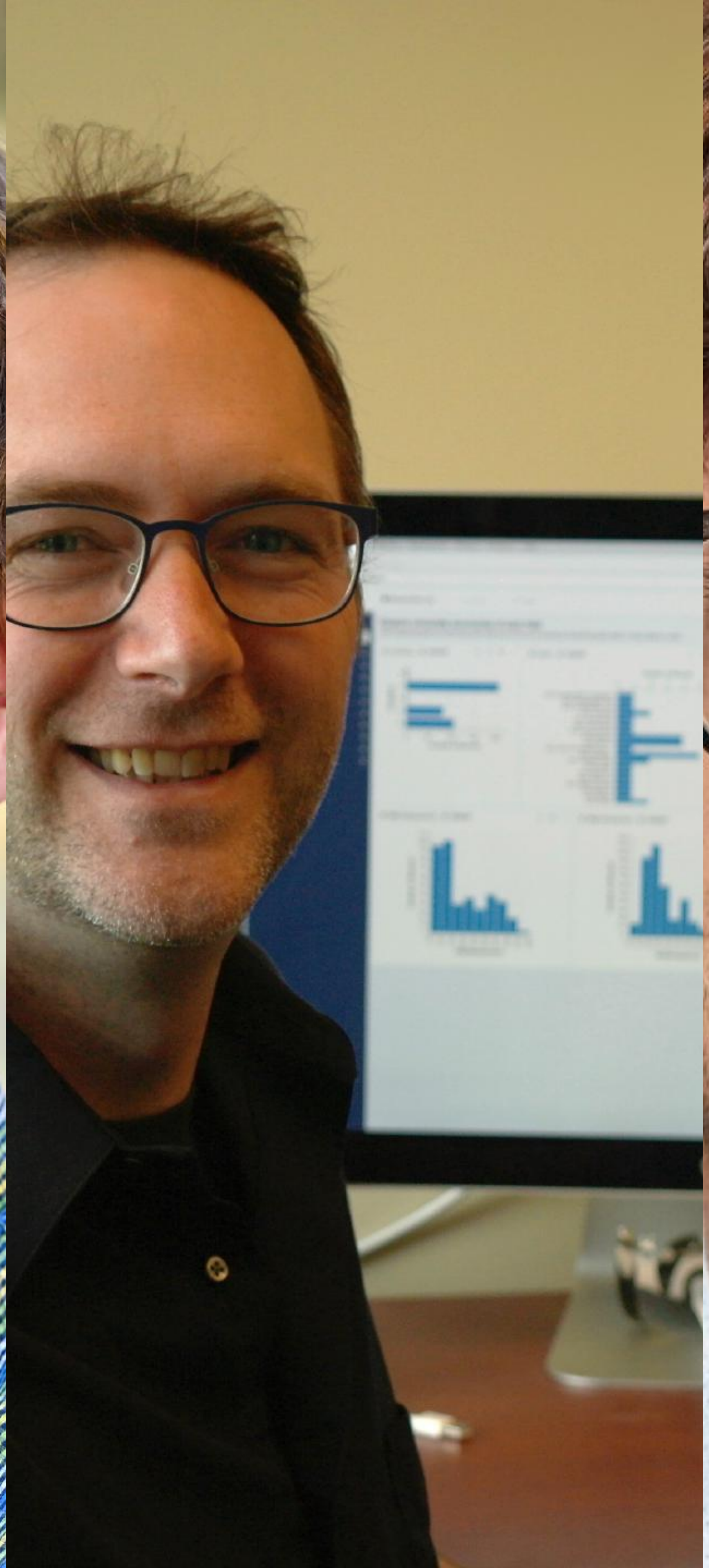
- different results despite substantial overlap in analytical approach

Why?


data
analysis
results


analysis


data
analysis 1
analysis 2
analysis 3
analysis 4
analysis 5
analysis 6
analysis 7


Schweinsberg et al. (2021):

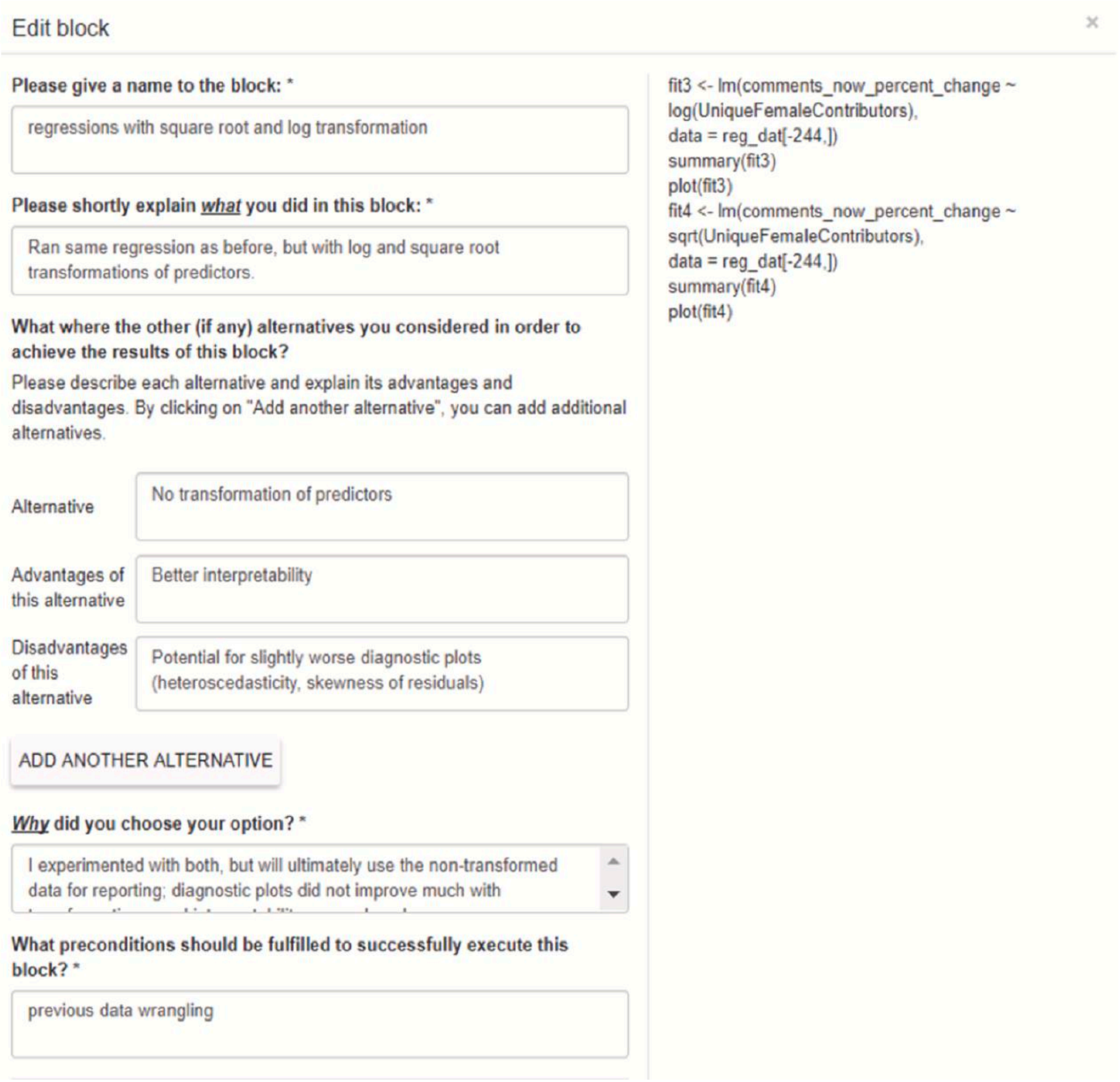
- 29 analyses tested two hypotheses
- analyzed same dataset: >3 million words in >7000 comments from the online academic forum Edge.org
  - 150 additional variables (contributor job title, gender, year of Ph.D., institution granting Ph.D., etc).

“Higher status participants are more verbose than lower status participants” (H2)

```
fit3 <- lm(comments_now_percent_change ~
log(UniqueFemaleContributors),
data = reg_dat[-244,])
summary(fit3)
plot(fit3)
fit4 <- lm(comments_now_percent_change ~
sqrt(UniqueFemaleContributors),
data = reg_dat[-244,])
summary(fit4)
plot(fit4)
```

Analyses conducted on DataExplained online platform:

— all analytical code saved


Edit block

Please give a name to the block: *

regressions with square root and log transformation

Please shortly explain *what* you did in this block: *

Ran same regression as before, but with log and square root transformations of predictors.

What where the other (if any) alternatives you considered in order to achieve the results of this block?

Please describe each alternative and explain its advantages and disadvantages. By clicking on "Add another alternative", you can add additional alternatives.

Alternative	No transformation of predictors
Advantages of this alternative	Better interpretability
Disadvantages of this alternative	Potential for slightly worse diagnostic plots (heteroscedasticity, skewness of residuals)

ADD ANOTHER ALTERNATIVE

***Why* did you choose your option?** *

I experimented with both, but will ultimately use the non-transformed data for reporting; diagnostic plots did not improve much with

What preconditions should be fulfilled to successfully execute this block? *

previous data wrangling

```
fit3 <- lm(comments_now_percent_change ~
log(UniqueFemaleContributors),
data = reg_dat[-244,])
summary(fit3)
plot(fit3)
fit4 <- lm(comments_now_percent_change ~
sqrt(UniqueFemaleContributors),
data = reg_dat[-244,])
summary(fit4)
plot(fit4)
```

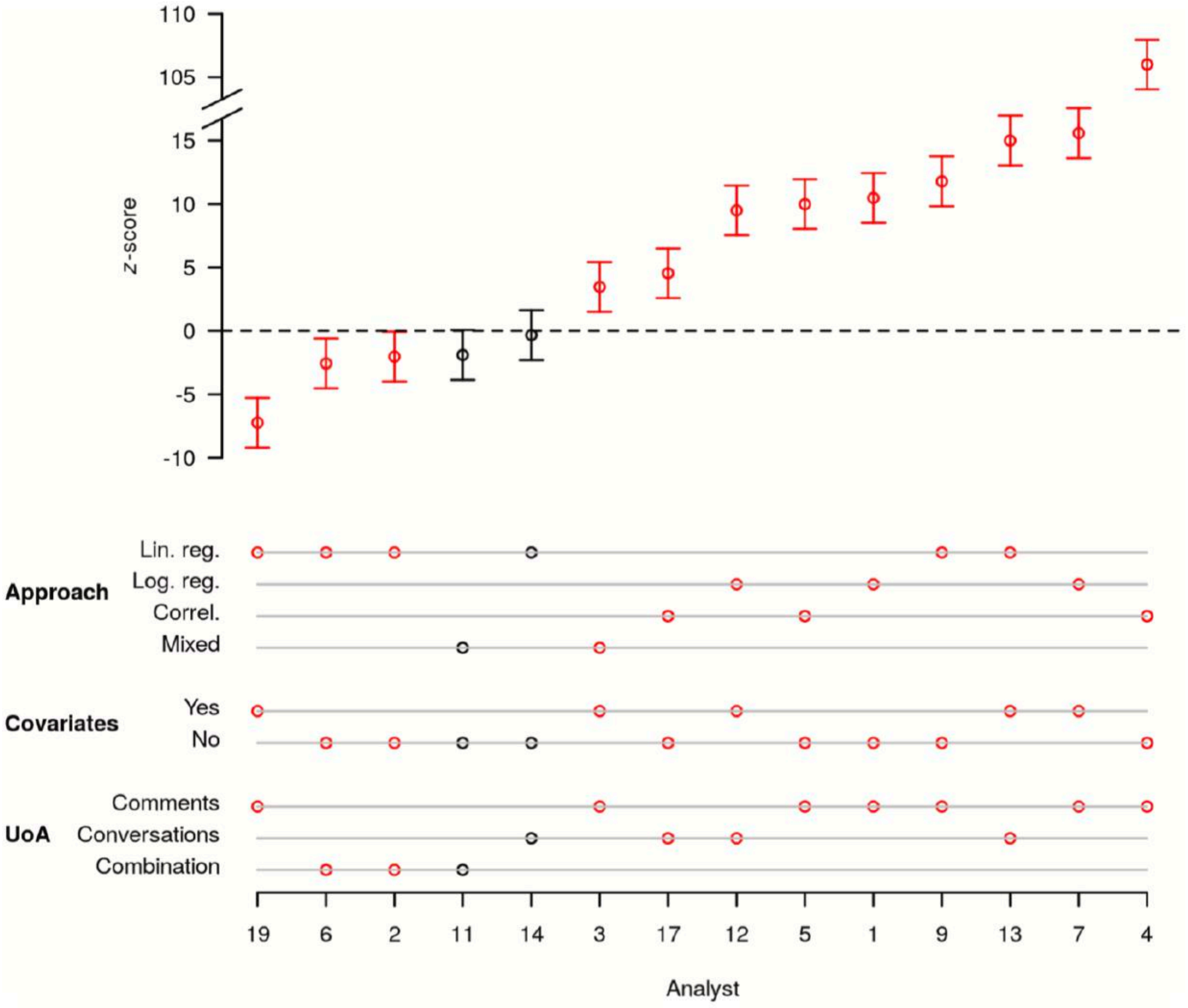
Analyses conducted on DataExplained online platform:

- all analytical code saved
- analysts asked about reasons for each analytical choice by platform:
  - *What did you do in this step?*
  - *Why did you choose this step?*
  - *What alternative(s) did you consider for this step?*

Analysts also laid out analytical path:

— split code into meaningful steps
— named each step
— put steps in order to reflect analytical workflow

Result 1: dispersion of results

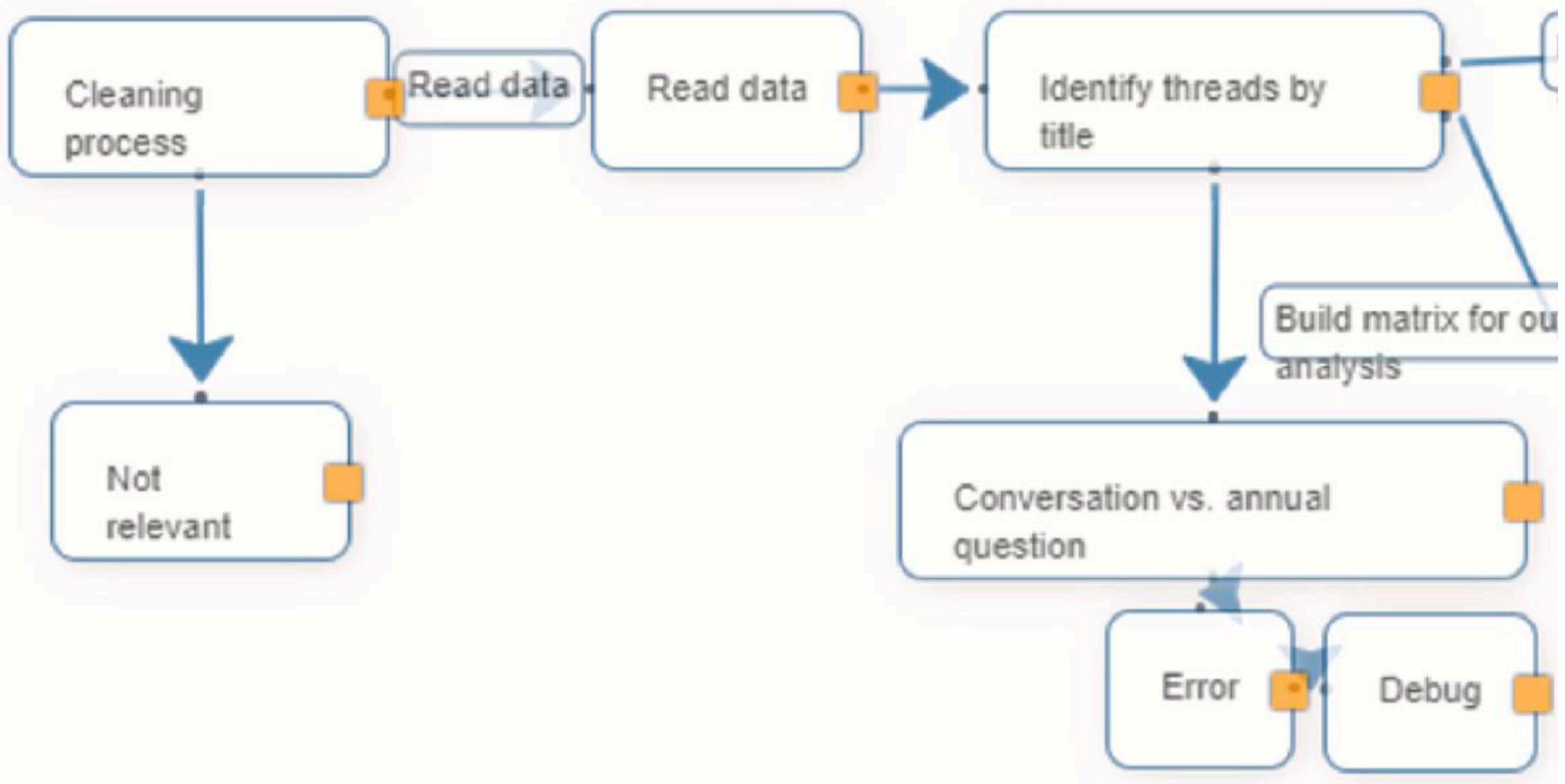


— some of which in opposite direction:

— despite no publication incentives (i.e. no incentive to *get* a signifiant result)

— and analysts acting in good faith

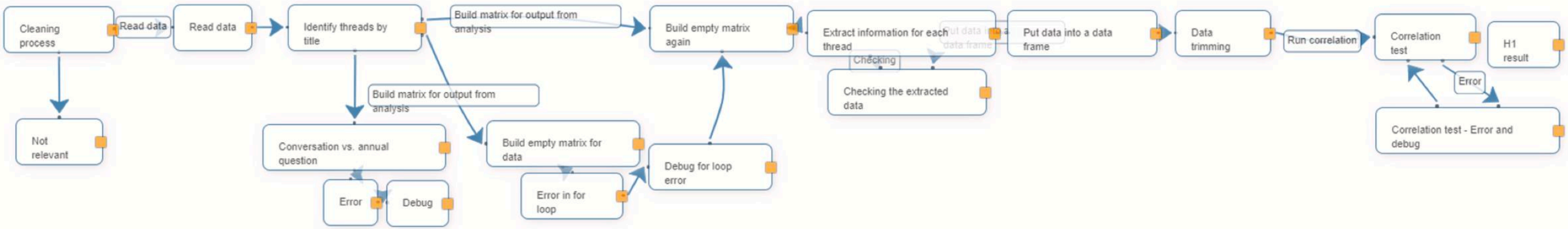
Result 2: Analyses are iterative with explorative loops

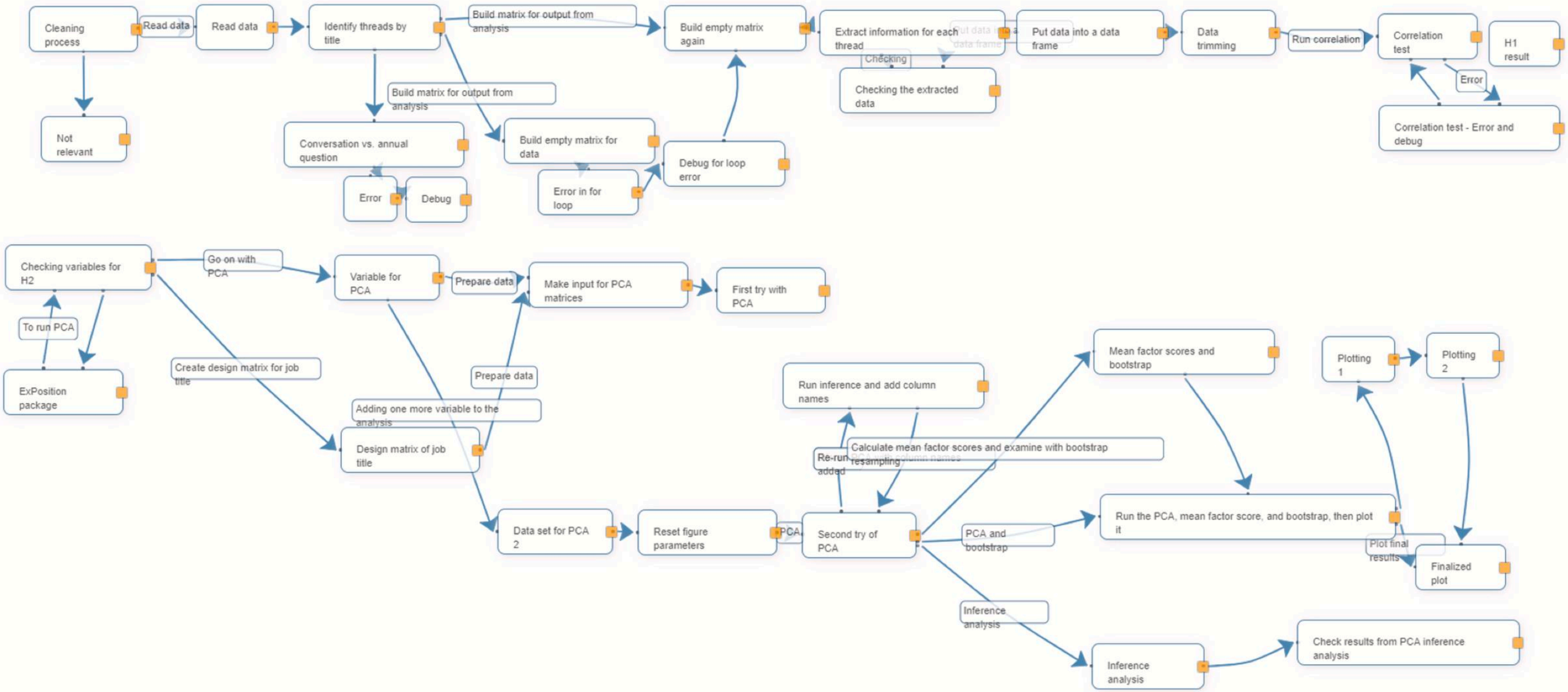


Analysts:

1. explore the data
2. make sense of what they found
3. then decide on next step

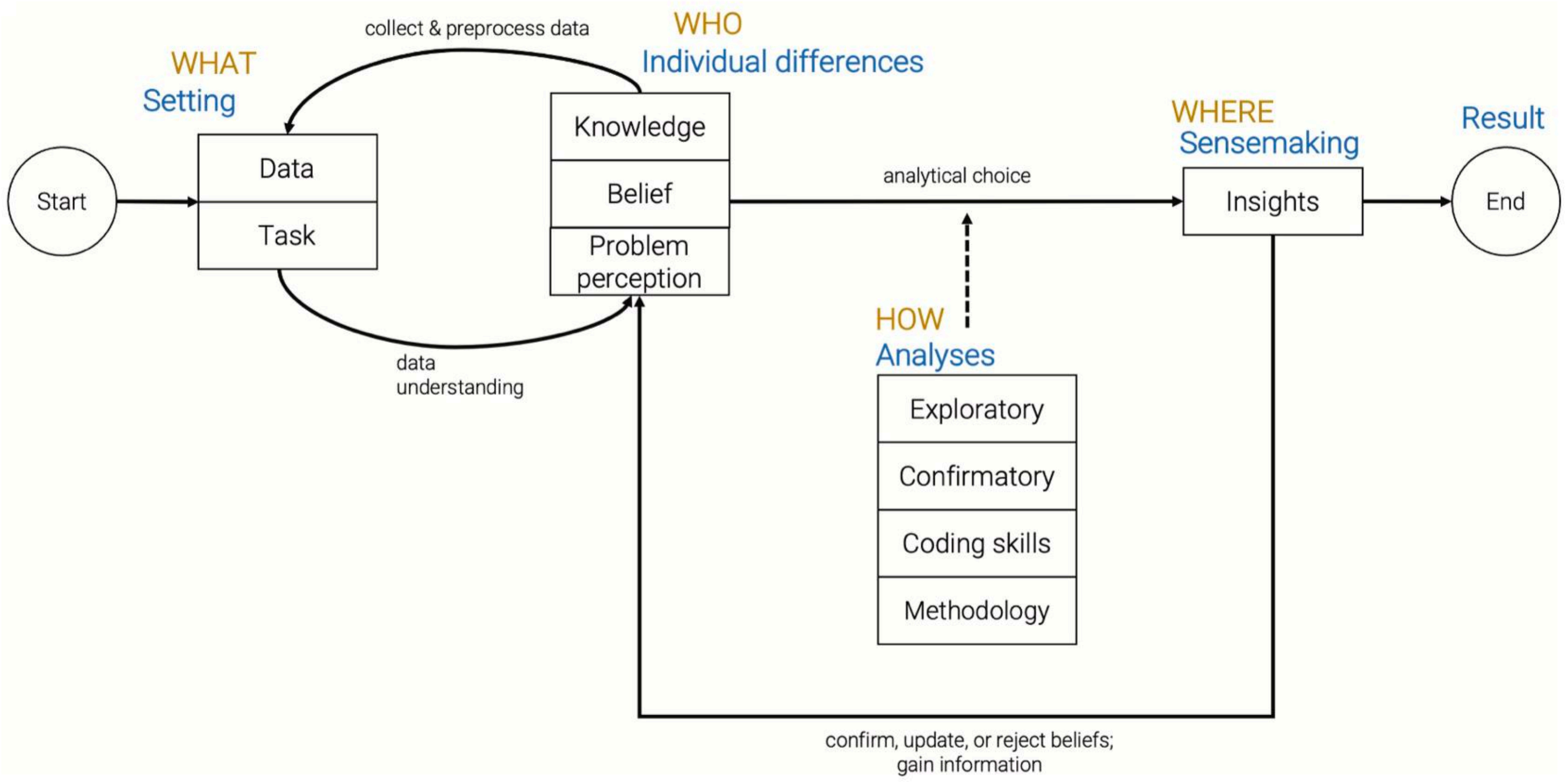
Explorative loops necessary so analysts make sense of data


Cleaning process
Read data
Read data
Identify threads by title
Build matrix for output from analysis
Build empty matrix again
Extract information for each thread
Put data into a data frame
Data trimming
Run correlation
Correlation test
H1 result
Not relevant
Build matrix for output from analysis
Conversation vs. annual question
Build empty matrix for data
Debug for loop error
Checking
Checking the extracted data
Error
Correlation test - Error and debug
Error
Debug
Error in for loop

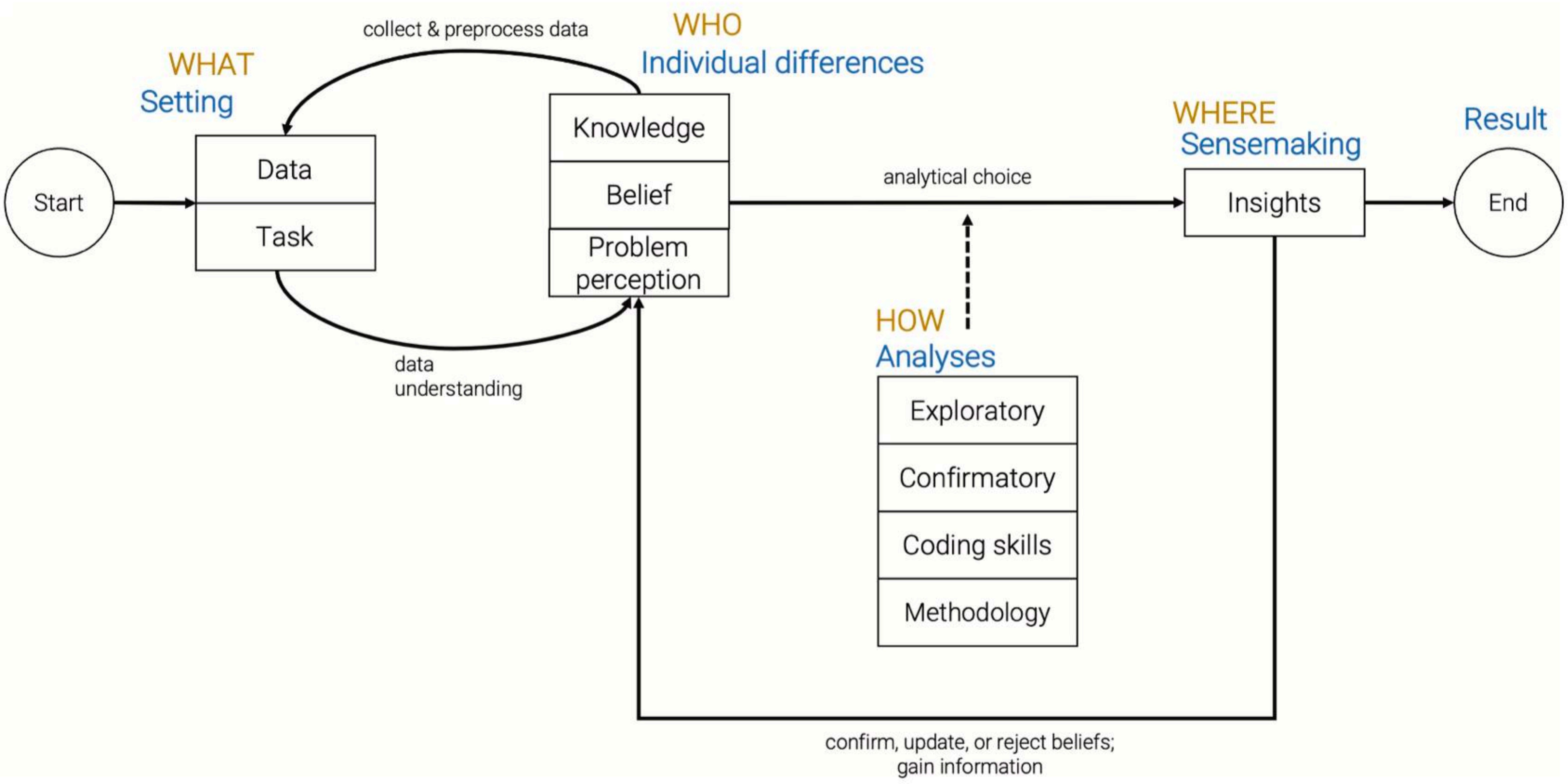


Cleaning process
Read data
Read data
Identify threads by title
Build matrix for output from analysis
Build empty matrix again
Extract information for each thread
Put data into a data frame
Data trimming
Run correlation
Correlation test
H1 result
Not relevant
Build matrix for output from analysis
Conversation vs. annual question
Build empty matrix for data
Debug for loop error
Checking
Checking the extracted data
Error
Correlation test - Error and debug
Error
Debug
Error in for loop
Checking variables for H2
Go on with PCA
Variable for PCA
Prepare data
Make input for PCA matrices
First try with PCA
To run PCA
ExPosition package
Create design matrix for job title
Prepare data
Adding one more variable to the analysis
Design matrix of job title
Mean factor scores and bootstrap
Plotting 1
Plotting 2
Run inference and add column names
Calculate mean factor scores and examine with bootstrap resampling
Re-run added
Data set for PCA 2
Reset figure parameters
PCA
Second try of PCA
PCA and bootstrap
Run the PCA, mean factor score, and bootstrap, then plot it
Plot final results
Finalized plot
Inference analysis
Inference analysis
Check results from PCA inference analysis

Result 3. Model of psychological process: Analytical choices are likely subjective:

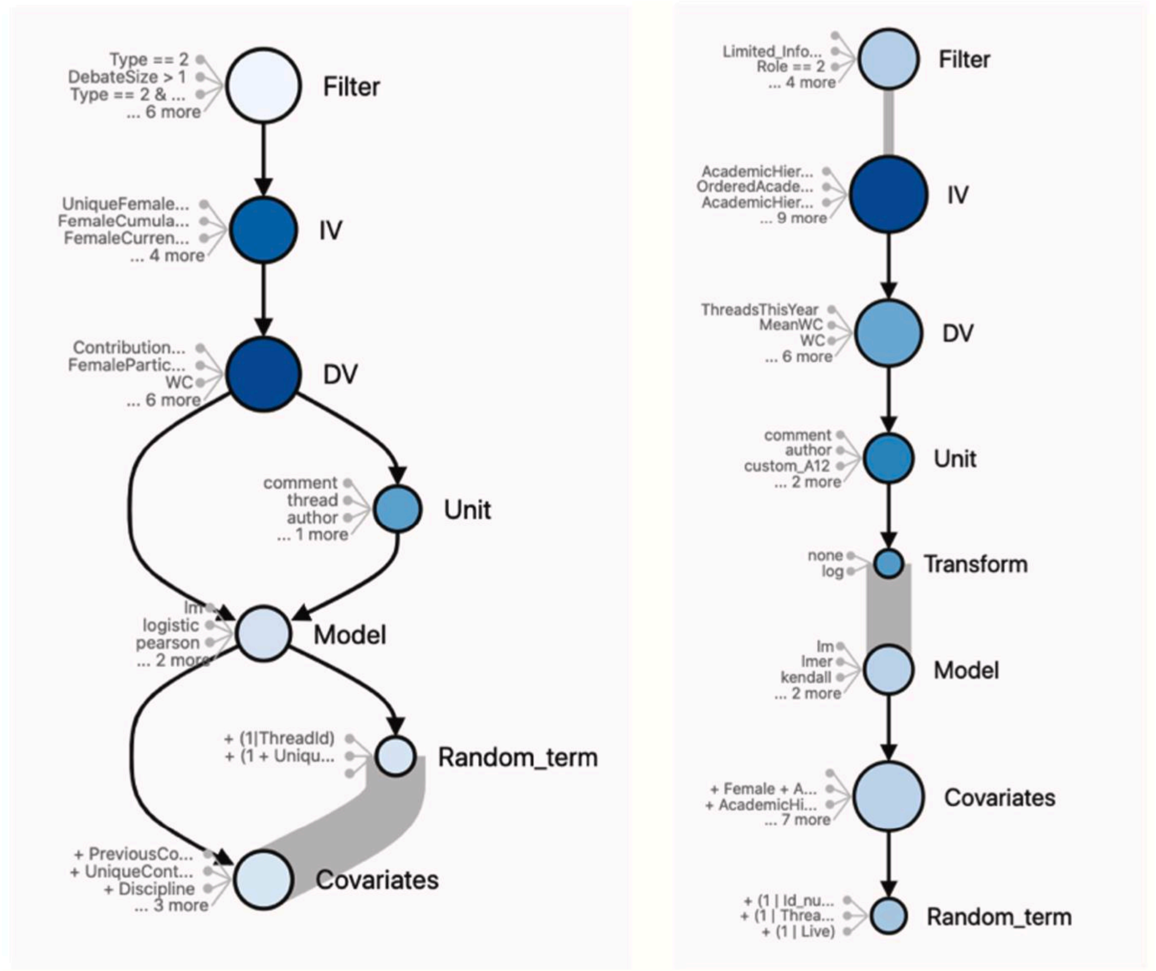


- qualitative analysis of analysts' code & thought process explanation
- subjective construal required throughout the analytical process



collect & preprocess data
WHO
Individual differences
WHAT
Setting
Knowledge
Belief
Problem perception
Data
Task
Start
analytical choice
WHERE
Sensemaking
Insights
Result
End
HOW
Analyses
Exploratory
Confirmatory
Coding skills
Methodology
data understanding
confirm, update, or reject beliefs;
gain information

Result 4. BOBA multiverse analysis (Liu et al., 2020):



H1 H2

Fig. 8. Analytic decision graphs. Nodes represent analytic branches, and edges indicate order and dependency between branches. The size of a node encodes the number of alternative analytic approaches. Color maps to sensitivity, with darker color indicating a more sensitive branch. Here, sensitivity is computed using the k-samples Anderson-Darling test.

— Operationalization key cause of variance in *this* study

Now what?

Mindset:
Humble towards own findings
+
open to alternatives

Exciting times to do science!

Analyses increasingly open and transparent: sharing data, code, and processes

Solutions and potential next steps:

— transparently share data, code, and study materials (Nosek et al., 2021)

This can help you *save* pages!

Hypothesis 5. In value-driven conflicts, the perceived discrepancy in positions becomes greater over the course of the negotiation if information is exchanged than if it is not. In utility-driven conflicts, the perceived discrepancy in positions becomes smaller over the course of the negotiation if information is exchanged than if it is not.

Besides the preregistered hypotheses and preregistered manipulation checks, we exploratively examined further dependent variables (compromises, subjective evaluations of the process, and relationship with the counterpart).[3] The data (including all measures) and a detailed overview of the results (Supplement 1) are available as supplemental materials. The preregistration and the complete study materials can be accessed via the project folder on osf (Schuster, 2020a).

- authors share data, code, and study materials separately online
- helps save space (journal page limits) and is comprehensive

1.8MB | Public | 0 | ···

Value relevant conflicts_ Study 1

Contributors: Carolin Schuster, Johann M. Majer, Roman Trötschel

Date created: 2017-05-10 09:29 AM | Last Updated: 2020-05-08 07:04 AM

Category: Project

Description: *experimental scenario study on the effect of information exchange in value relevant conflicts*

Files

Filter

Name	Modified
Value relevant conflicts_ Study 1	
– OSF Storage (United States)	
Preregistration_pilot 1.pdf	2017-05-10 09:33 AM
Preregistration_Study 1_online negotiatio...	2017-07-26 07:31 AM
Study 1 - online questionnaire.pdf	2019-07-11 06:12 AM
Study 1 data.sav	2020-03-31 10:18 AM
Supplemental 1.docx	2020-03-31 10:18 AM

Citation

Recent Activity

Carolin Schuster changed permissions for Value relevant conflicts_ Study 1
2020-05-08 07:04 AM

Carolin Schuster added Johann M. Majer and Roman Trötschel as contributor(s) to Value relevant conflicts_ Study 1
2020-05-08 07:04 AM

Carolin Schuster made Value relevant conflicts_ Study 1 public
2020-04-29 04:35 AM

Carolin Schuster updated file Study 1 data.sav in OSF Storage in Value relevant conflicts_ Study 1
2020-03-31 10:18 AM

Carolin Schuster added file Supplemental 1.docx to OSF Storage in Value relevant conflicts_ Study 1
2020-03-31 10:18 AM

Carolin Schuster removed file Detailed Overview of Results_Study 1.xlsx from OSF Storage in Value relevant conflicts_ Study 1

Journal of Experimental Social Psychology 90 (2020) 103993

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

Whatever we negotiate is not what I like: How value-driven conflicts impact negotiation behaviors, outcomes, and subjective evaluations☆,☆☆

Carolin Schuster*, Johann M. Majer, Roman Trötschel

Leuphana University Lüneburg, Germany

ARTICLE INFO

Keywords:
Value conflict
Integrative negotiation
Subjective value
Information exchange

ABSTRACT

Value conflicts have been shown to impair negotiation behaviors and outcomes (Harinck & Ellemers, 2014). The present studies aim to replicate and extend this finding in a paradigm where the parties' values were different, but not opposed. We hypothesized that activating values, rather than utilities, as motives in a negotiation would not only impair negotiation behavior and outcomes, but also subjective evaluations of the negotiated agreements. We further predicted that information provided about the counterparts' priorities would be a less effective facilitator of integrative negotiating in value-driven than in utility-driven conflicts. Two preregistered experiments (N = 176/310) confirm that a value motive leads to an increased aversion to trade-offs and to more compromise offers (Studies 1 and 2), and to lower individual and joint outcomes (Study 2). The results also show that the activation of value rather than utility motives in the parties trigger subjective perceptions of clashing values rather than conflicting interests, even though the values were not opposed. By triggering these perceptions, the value motives indirectly lead to worse subjective evaluations of the outcome, the process, the self, and (in Study 2) the relationship, even when controlling for objective outcomes. Providing information about the counterpart's underlying motives did not produce conclusive differential effects on negotiation behaviors and outcomes, possibly because the shared information did not increase the perceived discrepancy between counterparts. Theoretical and practical implications of the results for value-driven conflicts are discussed.

1. Introduction

Some of the most pressing societal conflicts involve core values. For example, in 2019 climate activists across the globe have demonstrated for a sustainable future (Taylor, Watts, & Bartlett, 2019). Among other things, they demand a greater use of renewable energy such as wind energy. However, opponents of wind energy in Germany have also brought up value-based concerns about its impact on health, the beauty of the landscape, and even the protection of wildlife (Hessler, 2019), values that in theory seem similar to those of climate activists. Unfortunately, negotiations about issues that are based on values seem to be particularly hard to solve and often end with frustration on both sides. For example, the conflict about wind energy has not been solved while electricity prices are rising (Sönnichsen, 2020) and fossil fuels, which are unpopular on both sides, still account for almost 50% of the energy mix (Hessler, 2019).

Negotiating integrative (or win-win) agreements that maximize the extent to which both parties' interests are satisfied typically requires both parties to make systematic concessions. Instead of making moderate concessions on all issues (compromising), better joint outcomes are reached by making strong concessions on some issues that are more important for the other party than oneself and trading these off against strong reciprocal concessions of the other party on issues that are more important for oneself than the other (making integrative trade-offs). This requires both parties to understand the other's underlying interests and priorities and make concessions accordingly (Pruitt & Carnevale, 1993; Thompson, 1991; Trötschel, Hüffmeier, Loschelder, Schwartz, & Gollwitzer, 2011).

However, personal values are closely tied to one's identity. When the values of parties involved in a conflict clash, the parties may feel that their identities are threatened (Harinck & Ellemers, 2014). Core values can even function as moral imperatives such that it is a taboo to even think about conceding on issues related to them (Tetlock, 2003; Tetlock, Kristel, Elson, Green, & Lerner, 2000). Therefore, parties in

☆ This paper has been recommended for acceptance by Rachel Barkan.

☆☆ Authors' note: We would like to thank Pauline Tenbrüggen, Helen Sauter, and Jonas Brammer for their help with preparing the studies and collecting data. In addition, we thank Marcel Hartmann and Wiebke Heyse for their support in preparing supplemental materials.

* Corresponding author at: Leuphana University Lüneburg, Universitätsallee 1, 21335 Lüneburg, Germany
E-mail address: carolin.schuster@leuphana.de (C. Schuster).

https://doi.org/10.1016/j.jesp.2020.103993
Received 11 July 2019; Received in revised form 31 March 2020; Accepted 9 April 2020
0022-1031/

Journals increasingly reward transparency with badges

- Authors *want* to have these badges
- Signal transparency and rigour

Solutions and potential next steps:

- transparently share data, code, and study materials (Nosek et al., 2021)
- multiverse analyses (systematically perform *all possible analyses*) to identify how fragile / robust findings are (Liu et al., 2020; Steegen et al., 2016)

Solutions and potential next steps:

- transparently share data, code, and study materials (Nosek et al., 2021)
- multiverse analyses (systematically perform *all possible analyses*) to identify how fragile / robust findings are (Liu et al., 2020; Steegen et al., 2016)
- replicate findings before, rather than after publication (Schweinsberg et al., 2016) / adversarial collaboration in ML: "*Can you make this go away*"?

Thank you!
Questions for the group?

References

Botvinik-Nezer, R., Holzmeister, F., Camerer, C. F., Dreber, A., Huber, J., Johannesson, M., Kirchler, M., Iwanir, R., Mumford, J. A., Adcock, R. A., Avesani, P., Baczkowski, B. M., Bajracharya, A., Bakst, L., Ball, S., Barilari, M., Bault, N., Beaton, D., Beitner, J., . . . Schonberg, T. (2020). Variability in the analysis of a single neuroimaging dataset by many teams. Nature, 582(7810), 84-88. https://doi.org/10.1038/s41586-020-2314-9

Liu, Y., Kale, A., Althoff, T., & Heer, J. (2020). Boba: Authoring and visualizing multiverse analyses. IEEE Transactions on Visualization and Computer Graphics, 27(2), 1753-1763. https://doi.org/10.1109/TVCG.2020.3028985

Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., Fidler, F., Hilgard, J., Struhl, M. K., Nuijten, M. B., Rohrer, J. M., Romero, F., Scheel, A. M., Scherer, L. D., Schönbrodt, F. D., & Vazire, S. (2022). Replicability, robustness, and reproducibility in psychological science. Annual Review of Psychology, 73(1), 719-748. https://doi.org/10.1146/annurev-psych-020821-114157

Schweinsberg, M., Feldman, M., Staub, N., van den Akker, O. R., van Aert, R. C. M., van Assen, M. A. L. M., Liu, Y., Althoff, T., Heer, J., Kale, A., Mohamed, Z., Amireh, H., Venkatesh Prasad, V., Bernstein, A., Robinson, E., Snellman, K., Amy Sommer, S., Otner, S. M. G., Robinson, D., . . . Uhlmann, E. L. (2021). Same data, different conclusions: Radical dispersion in empirical results when independent analysts operationalize and test the same hypothesis. Organizational Behavior and Human Decision Processes, 165(July), 228-249. https://doi.org/https://

Silberzahn, R., Uhlmann, E. L., Martin, D. P., Anselmi, P., Aust, F., Awtrey, E., Bahnk, Bai, F., Bannard, C., Bonnier, E., Carlsson, R., Cheung, F., Christensen, G., Clay, R., Craig, M. A., Dalla Rosa, A., Dam, L., Evans, M. H., Flores Cervantes, I., . . . Nosek, B. A. (2018). Many analysts, one data set: Making transparent How variations in analytic choices affect results. Advances in Methods and Practices in Psychological Science, 1(3), 337-356. https://doi.org/10.1177/2515245917747646

Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. Perspectives on Psychological Science, 11(5), 702-712. https://doi.org/10.1177/1745691616658637