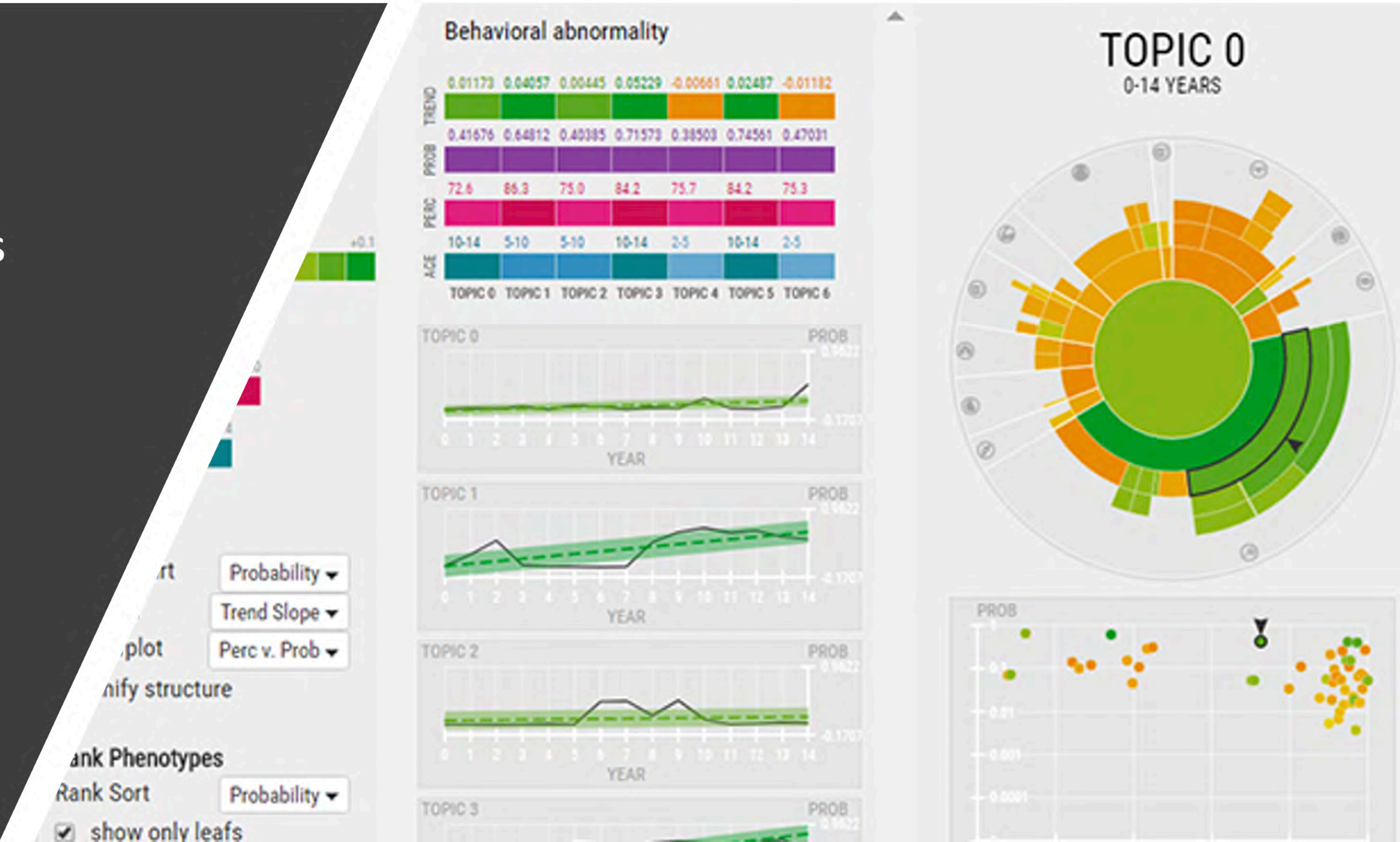
Fanny CHEVALIER

Departments of Computer Science & Statistics
University of Toronto

Research interests:

- Human-computer interaction
- Data visualization

http://fannychevalier.net



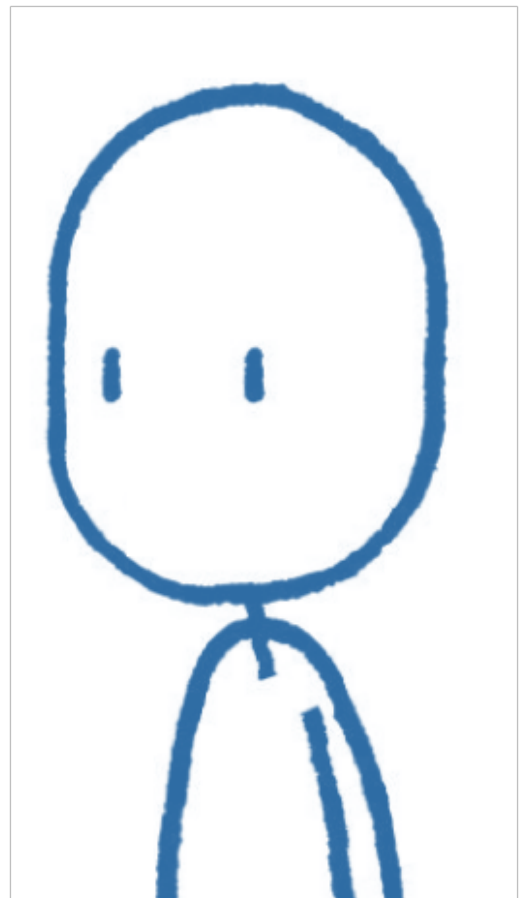
Data Insights: Bridging the Gap Between Numbers and Knowledge

Fanny Chevalier — University of Toronto

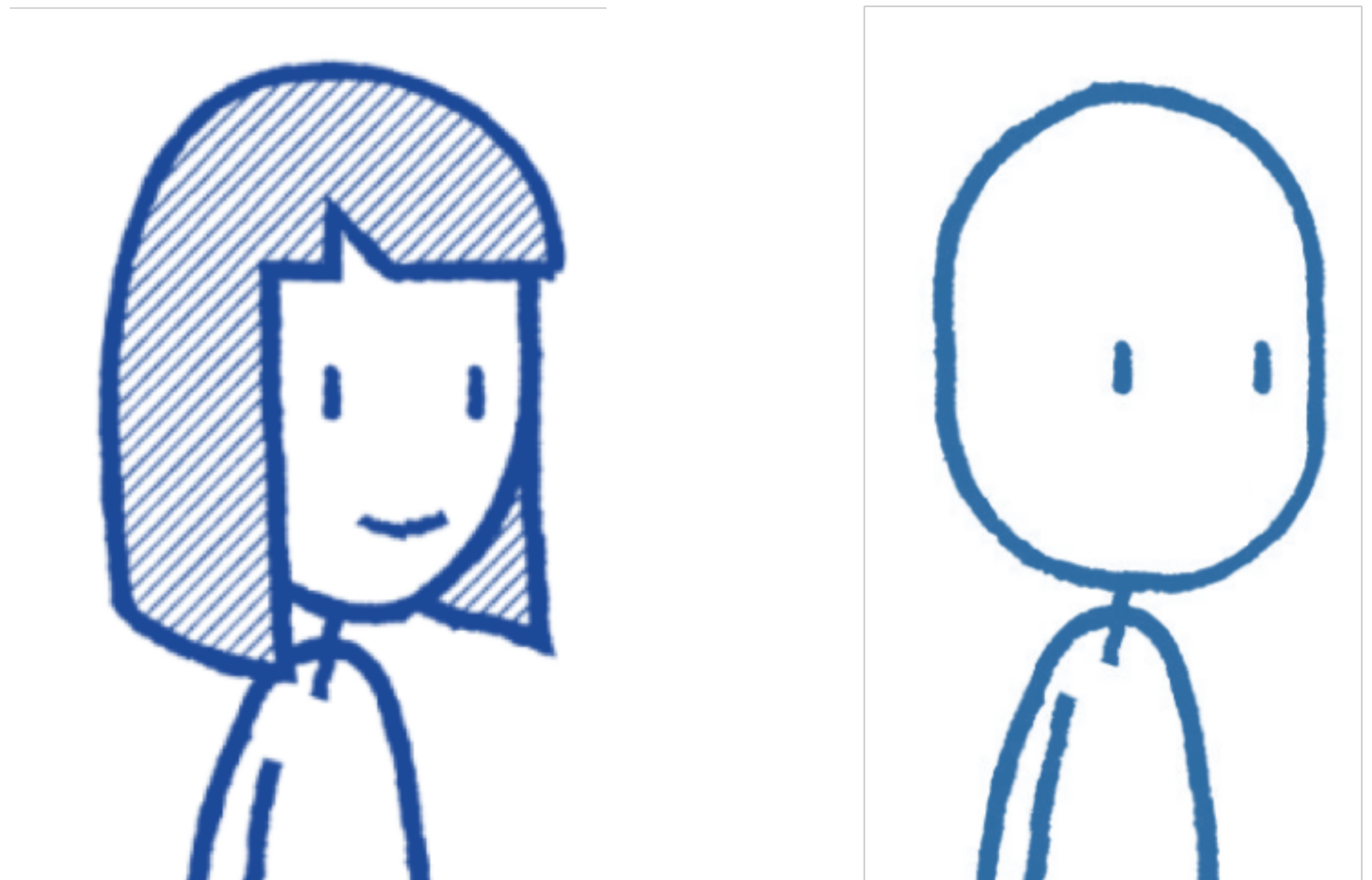
Scientists make sense of data

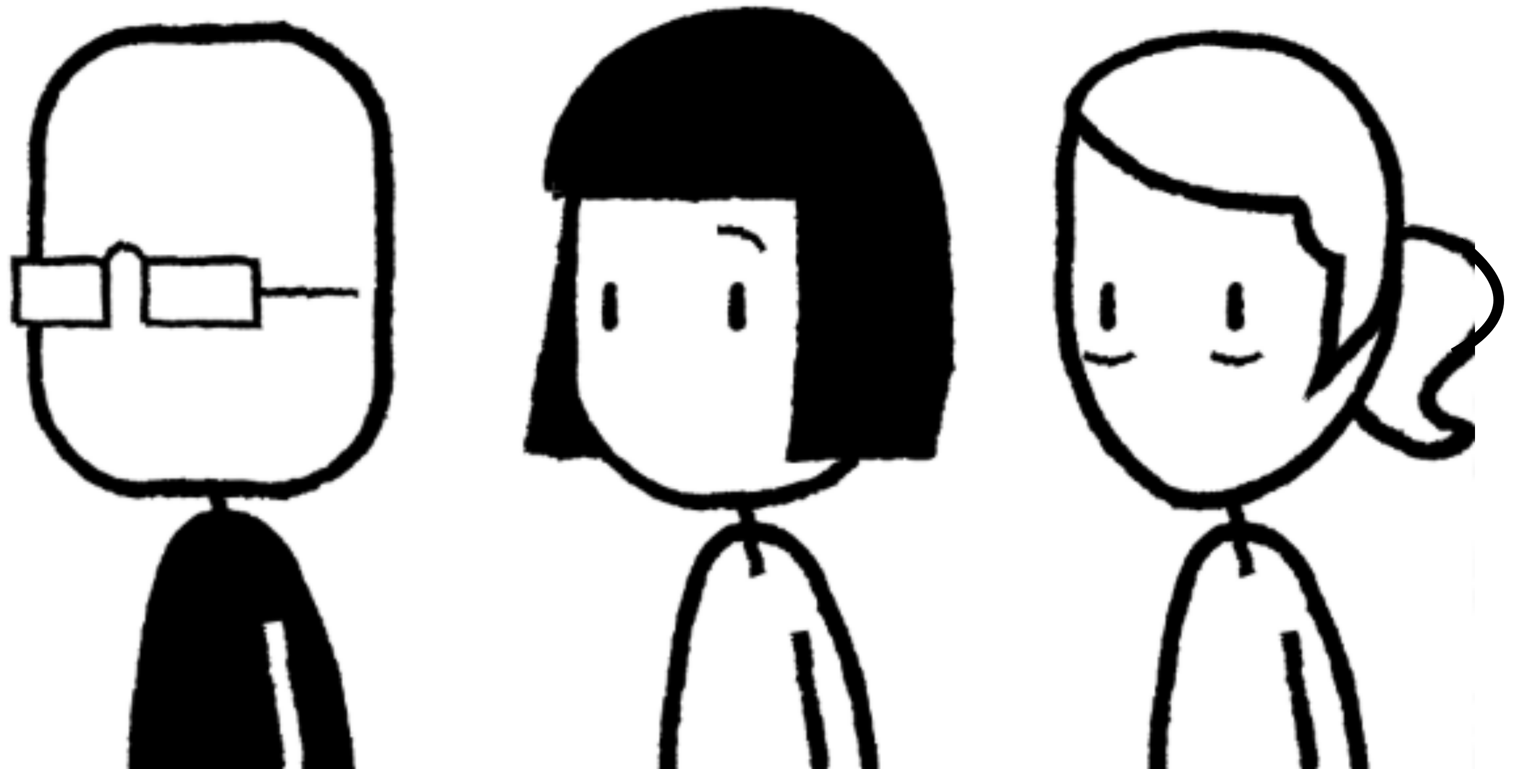
Scientists make sense of data

… communicate to other scientists



Scientists make sense of data
... communicate to **other scientists**
... and to **lay audiences**.





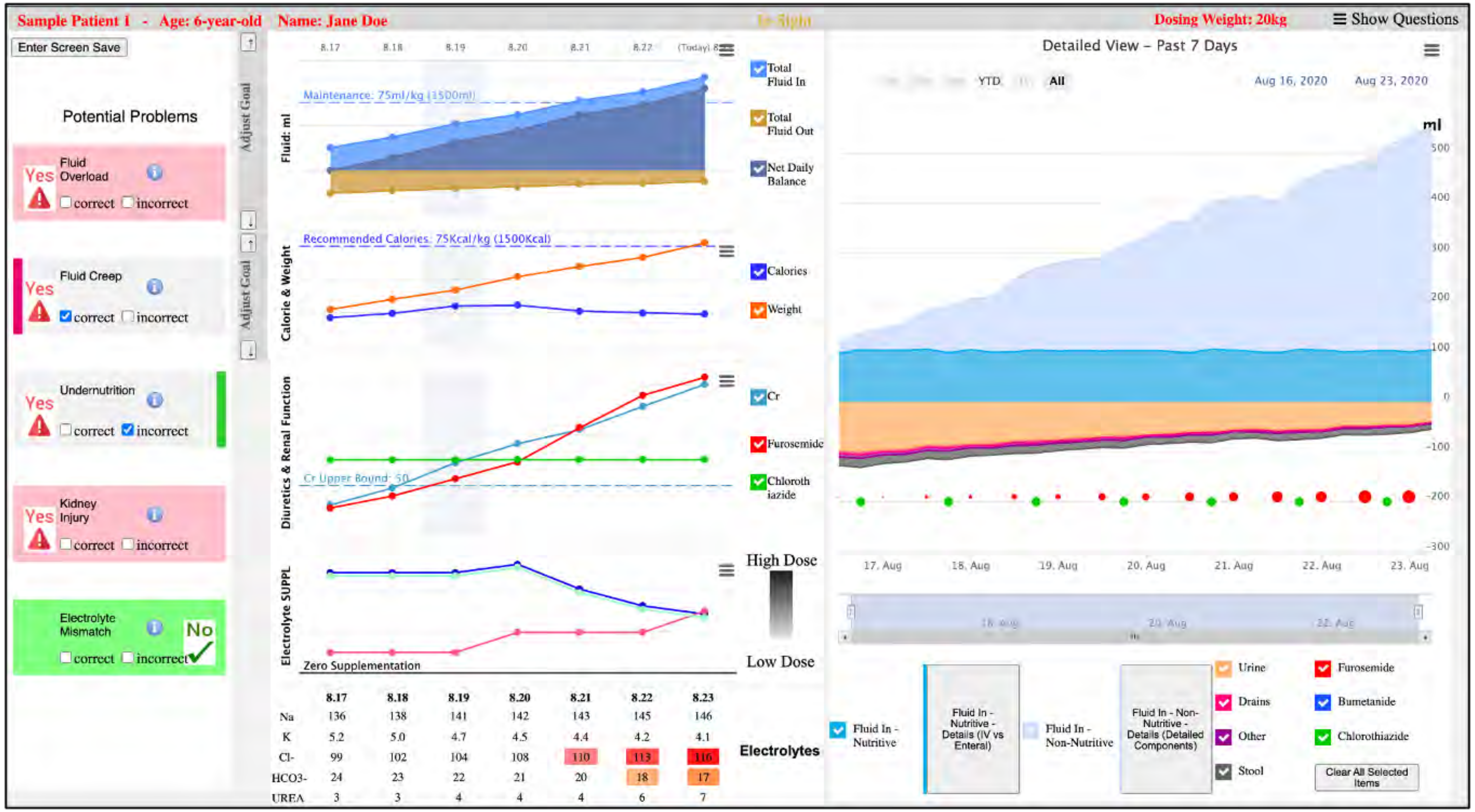
Scientists | … to other scientists | … to lay audiences

Making sense of data in critical environments


Dräger
87
93
36,4


Many displays
Constant alarms
Complex interfaces

Get to the Point!

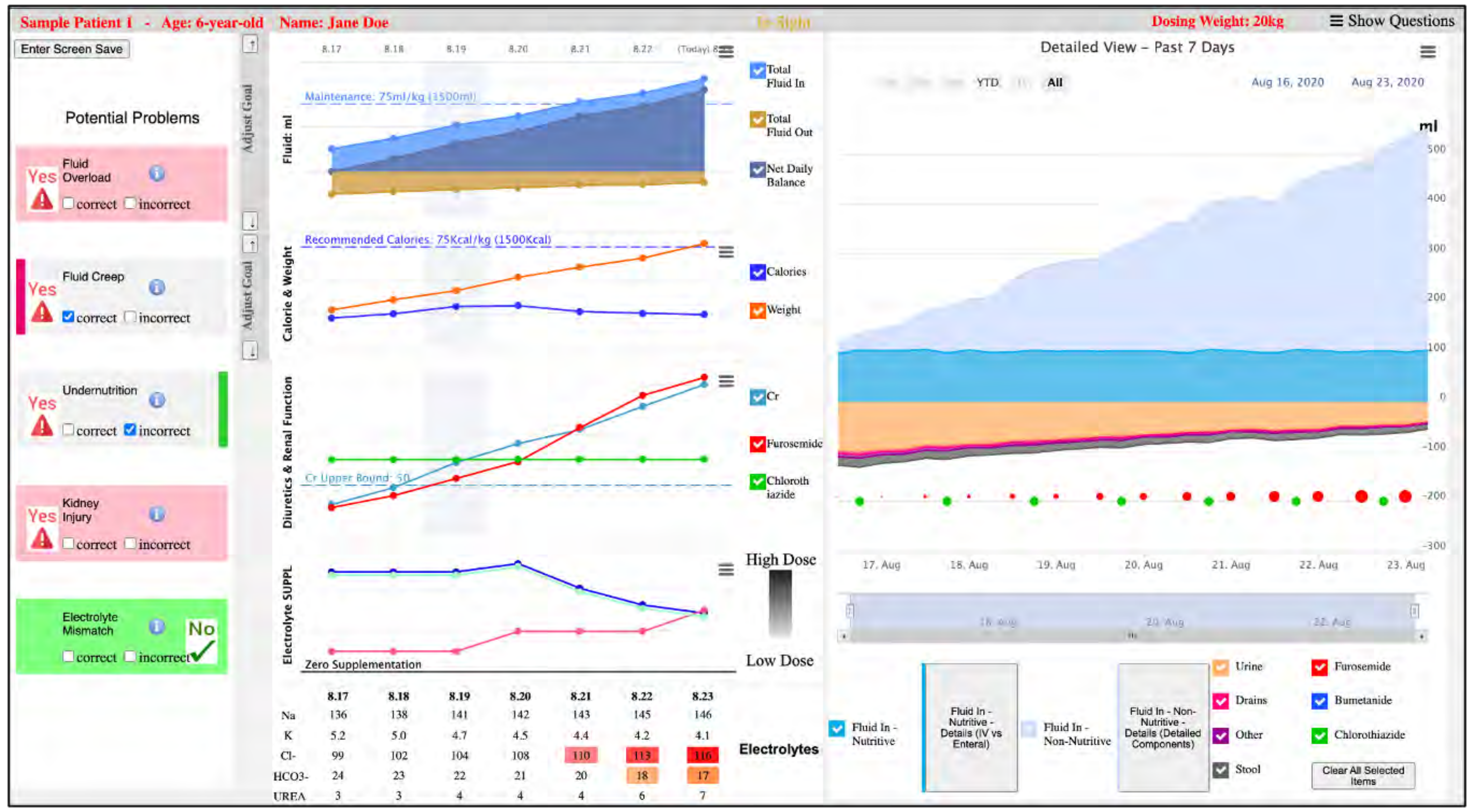


Zhang, Ehrmann, Mazwi, Eytan, Ghassemi, Chevalier (2022). **Get to the Point! Problem-Based Curated Data Views to Augment Care for Critically Ill Patients**. SIGCHI Conference on Human Factors in Computing Systems (CHI '22). Article No. 278.

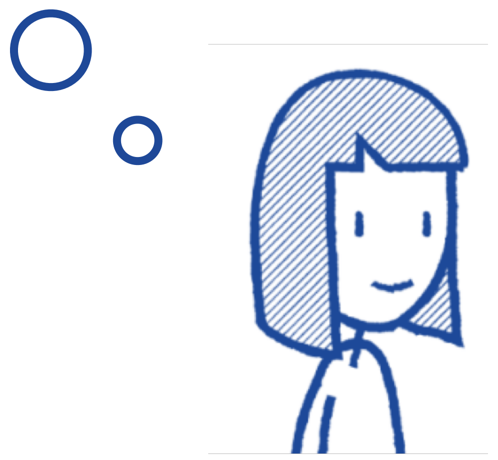
Problem-Based Curated Data Views

- [ ] Fluid Overload
- [ ] Fluid Creep
- [ ] Undernutrition
- [ ] Kidney Injury
- [ ] Electrolyte Mismatch

Checklist



Visualization



Zhang, Ehrmann, Mazwi, Eytan, Ghassemi, Chevalier (2022). **Get to the Point! Problem-Based Curated Data Views to Augment Care for Critically Ill Patients.** SIGCHI Conference on Human Factors in Computing Systems (CHI '22). Article No. 278.


Dräger
otential Problems
Fluid Overload
Fluid Creep
Undernutrition
Kidney Injury
Electrolyte Mismatch
87
36.4

Leave Screen Save

Potential Problems

Yes	Fluid Overload	☐ correct ☐ incorrect	
Yes	Fluid Creep	☐ correct ☐ incorrect	
Yes	Undernutrition	☐ correct ☐ incorrect	
Yes	Kidney Injury	☐ correct ☐ incorrect	
	Electrolyte Mismatch	☐ correct ☐ incorrect	No


Sample Patient 1 - Age: 6-year-old
Name: Jane Doe
In-Sight
Dosing Weight: 20kg
Show Questions
Enter Screen Save
Potential Problems
Fluid Overload
Yes
correct
incorrect
Fluid Creep
Yes
correct
incorrect
Undernutrition
Yes
correct
incorrect
Kidney Injury
Yes
correct
incorrect
Electrolyte Mismatch
No
correct
incorrect
Adjust Goal
Adjust Goal
8.17 8.18 8.19 8.20 8.21 8.22 (Today) 8.23
Fluid: ml
Maintenance: 75ml/kg (1500ml)
Total Fluid In
Total Fluid Out
Net Daily Balance
Calorie & Weight
Recommended Calories: 75Kcal/kg (1500Kcal)
Calories
Weight
Diuretics & Renal Function
Cr Upper Bound: 50
Cr
Furosemide
Chlorothiazide
Electrolyte SUPPL
Zero Supplementation
High Dose
Low Dose
Electrolytes

	8.17	8.18	8.19	8.20	8.21	8.22	8.23
Na	136	138	141	142	143	145	146
K	5.2	5.0	4.7	4.5	4.4	4.2	4.1
Cl-	99	102	104	108	110	113	116
HCO3-	24	23	22	21	20	18	17
UREA	3	3	4	4	4	6	7

Detailed View – Past 7 Days
1m 3m 6m YTD 1y All
Aug 16, 2020
Aug 23, 2020
ml
500
400
300
200
100
0
-100
-200
-300
17. Aug 18. Aug 19. Aug 20. Aug 21. Aug 22. Aug 23. Aug
18. Aug 20. Aug 22. Aug
Fluid In - Nutritive
Fluid In - Nutritive - Details (IV vs Enteral)
Fluid In - Non-Nutritive
Fluid In - Non-Nutritive - Details (Detailed Components)
Urine
Drains
Other
Stool
Furosemide
Bumetanide
Chlorothiazide
Clear All Selected Items


Sample Patient 1 - Age: 6-year-old Name: Jane Doe
In-Sight
Dosing Weight: 20kg
Show Questions
Enter Screen Save
otential Problems
Yes
Fluid Overload
correct incorrect
Yes
Fluid Creep
correct incorrect
Yes
Undernutrition
correct incorrect
Yes
Kidney Injury
correct incorrect
Electrolyte Mismatch
No
correct incorrect
Adjust Goal
8.17
8.18
8.19
8.20
8.21
8.22
(Today) 8.23
Maintenance: 75ml/kg (1500ml)
Fluid: ml
Total Fluid In
Total Fluid Out
Net Daily Balance

Problem-Based Curated Data Views

- Beyond yet-another dashboard: **design for user relevance**
- Balance **guidance** via checklist vs **discovery** via freeform exploration

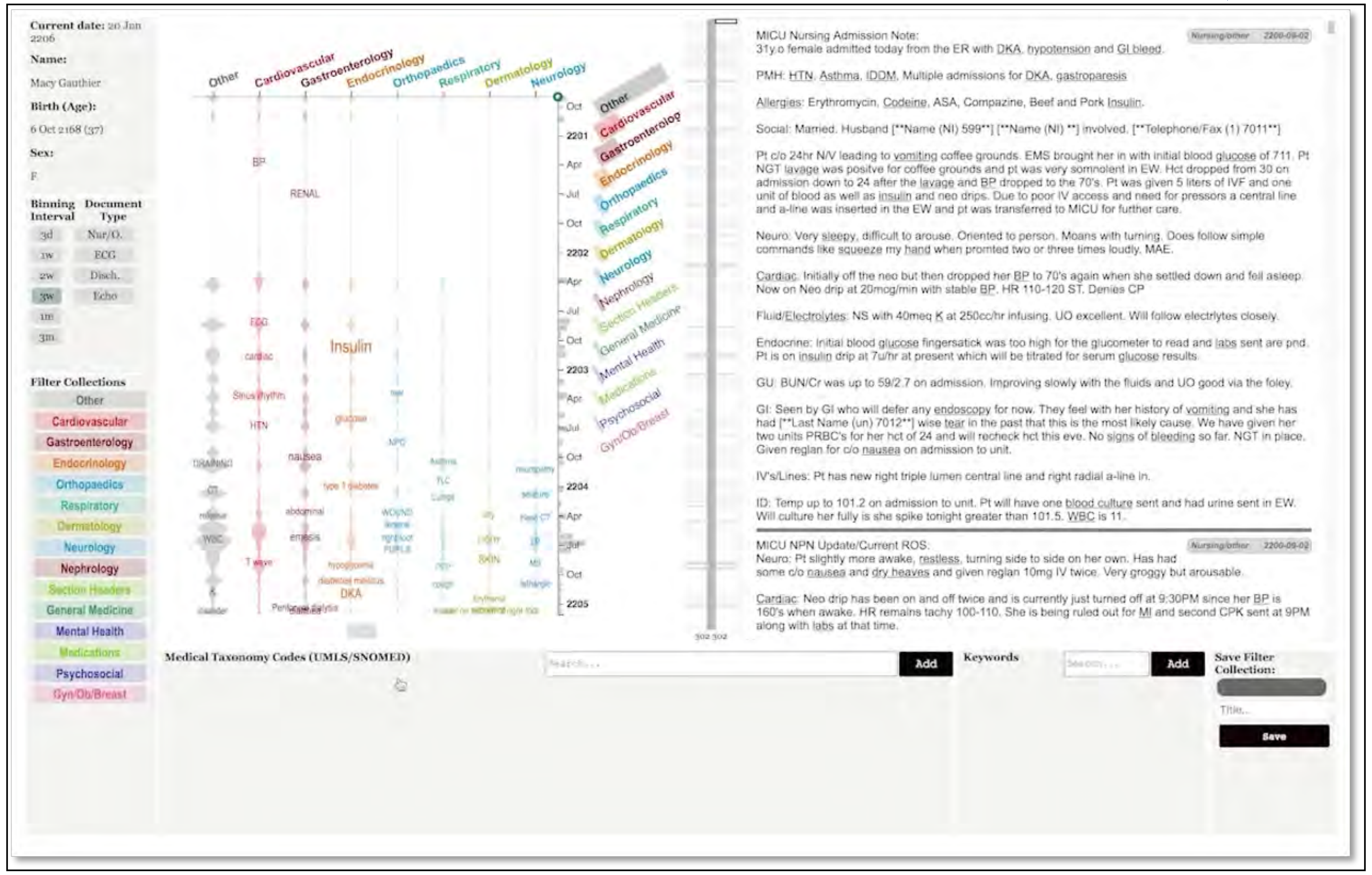
Scientists

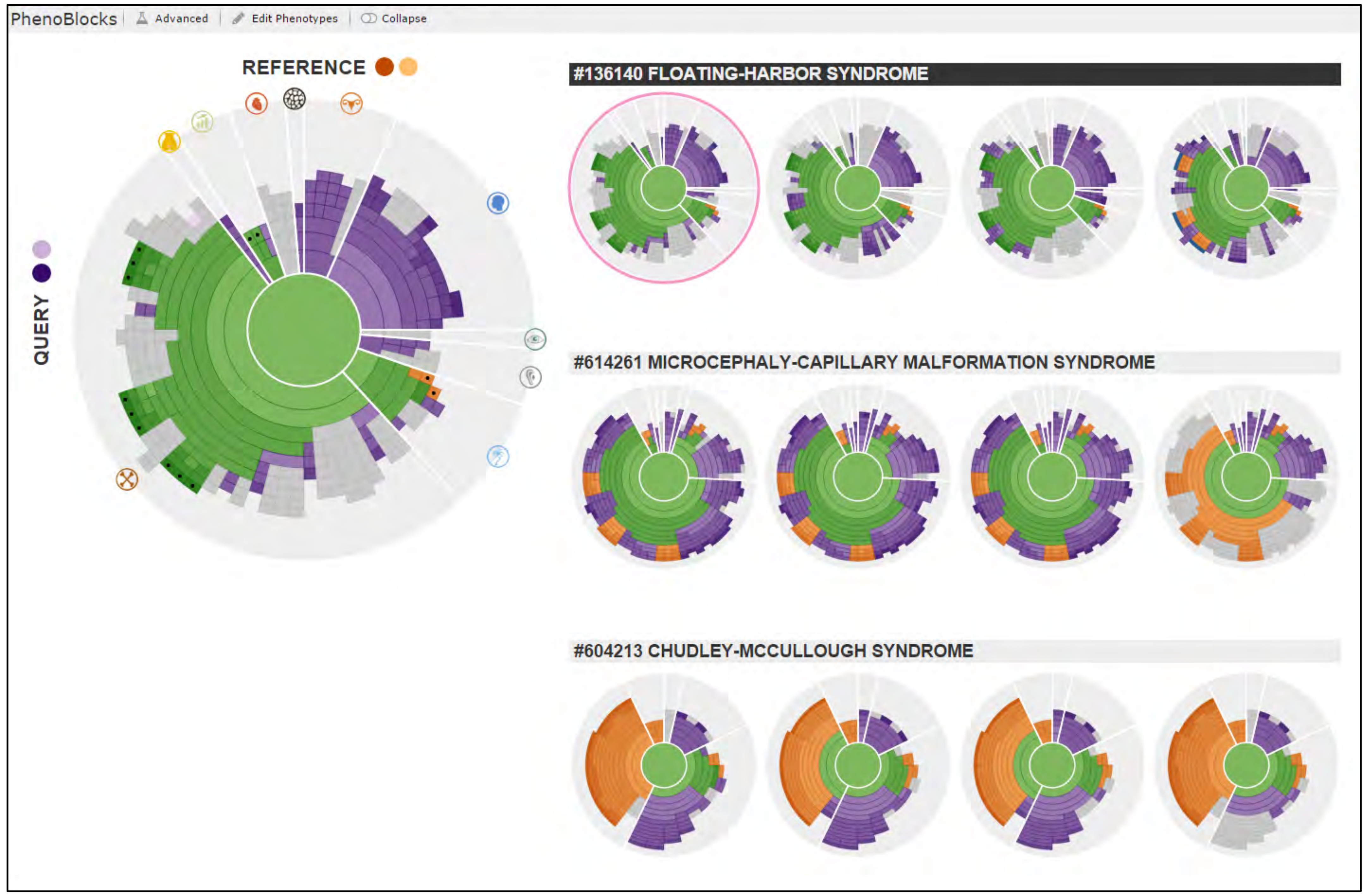
… to other scientists

… to lay audiences

Goals

Design to support experts' goals

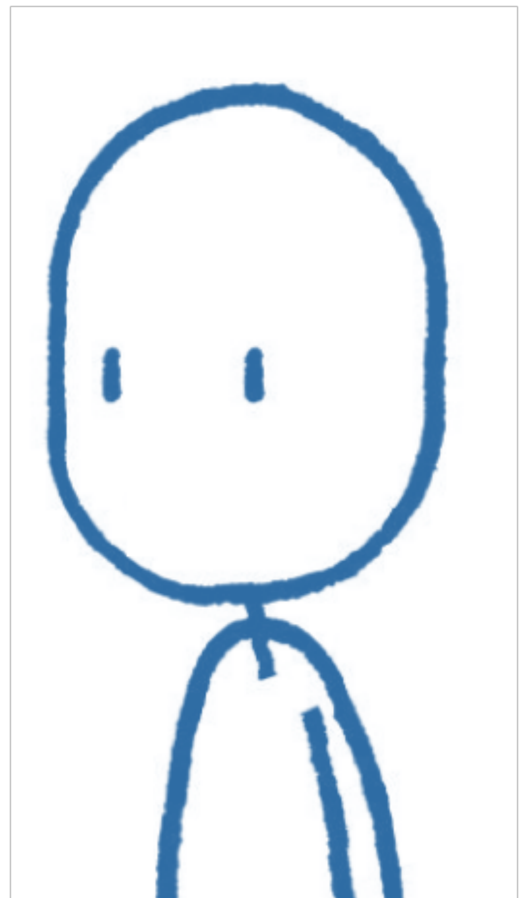




Nicole Sultanum, Devin Singh, Michael Brudno, Fanny Chevalier. **Doccurate: A Curation-Based Approach for Clinical Text Visualization** IEEE TVCG 2019.

Michael Glueck, Peter Hamilton, Fanny Chevalier, Simon Breslav, Azam Khan, Daniel Wigdor, Michael Brudno. **PhenoBlocks: Phenotype Comparison Visualizations** IEEE TVCG 2015.

Scientists make sense of data
... communicate to other scientists

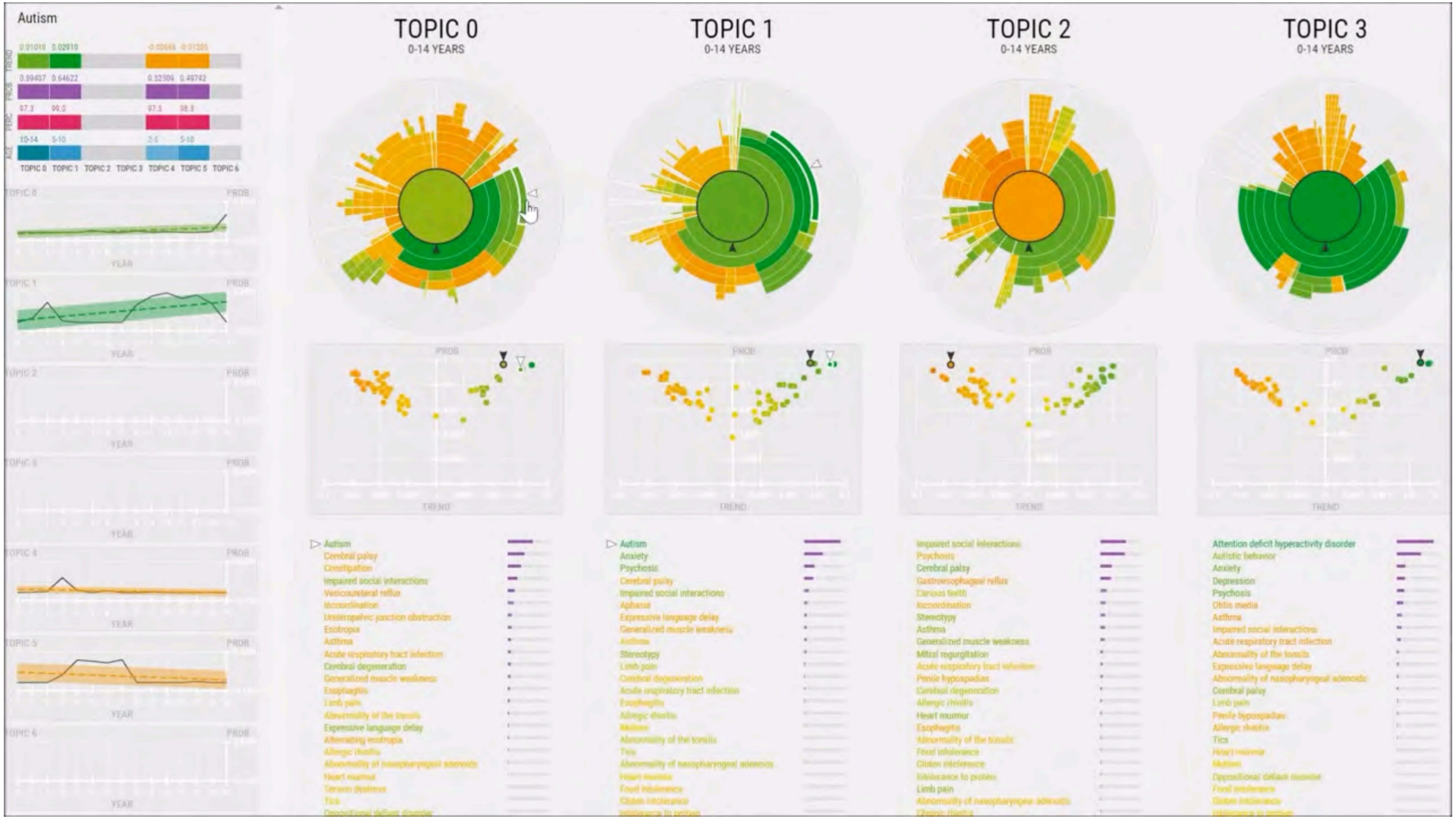


Scientists

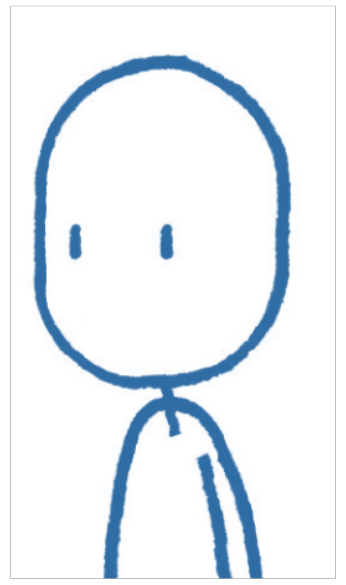
… to other scientists

… to lay audiences

Clinical Expert



ML Expert



Michael Glueck, Mahdi Pakdaman Naeini, Finale Doshi-Velez, Fanny Chevalier, Azam Khan, Daniel Wigdor **PhenoLines: Phenotype Comparison Visualizations for Disease Subtyping via Topic Models**. IEEE TVCG 2017.

Scientists

… to other scientists

… to lay audiences

AP Photo/Wilfredo Lee

Do hurricanes with more feminine names cause more deaths?

Female hurricanes are deadlier than male hurricanes

Kiju Jung[a,1], Sharon Shavitt[a,b,1], Madhu Viswanathan[a,c], and Joseph M. Hilbe[d]

[a]Department of Business Administration and [b]Department of Psychology, Institute of Communications Research, and Survey Research Laboratory, and [c]Women and Gender in Global Perspectives, University of Illinois at Urbana–Champaign, Champaign, IL 61820; and [d]Department of Statistics, T. Denny Sanford School of Social and Family Dynamics, Arizona State University, Tempe, AZ 85287-3701

Edited* by Susan T. Fiske, Princeton University, Princeton, NJ, and approved May 14, 2014 (received for review February 13, 2014)

Do people judge hurricane risks in the context of gender-based expectations? We use more than six decades of death rates from US hurricanes to show that feminine-named hurricanes cause significantly more deaths than do masculine-named hurricanes. Laboratory experiments indicate that this is because hurricane names lead to gender-based expectations about severity and this, in turn, guides respondents' preparedness to take protective action. This finding indicates an unfortunate and unintended consequence of the gendered naming of hurricanes, with important implications for policymakers, media practitioners, and the general public concerning hurricane communication and preparedness.

gender stereotypes | implicit bias | risk perception | natural hazard communication | bounded rationality

Estimates suggest that hurricanes kill more than 200 people in the United States annually, and severe hurricanes can cause fatalities in the thousands (1). As the global climate changes, the frequency and severity of such storms is expected to increase (2). However, motivating hurricane preparedness remains a major

violence and destruction (23, 24). We extend these findings to hypothesize that the anticipated severity of a hurricane with a masculine name (Victor) will be greater than that of a hurricane with a feminine name (Victoria). This expectation, in turn, will affect the protective actions that people take. As a result, a hurricane with a feminine vs. masculine name will lead to less protective action and more fatalities.

Archival Study

To test this hypothesis, we used archival data on actual fatalities caused by hurricanes in the United States (1950–2012). Ninety-four Atlantic hurricanes made landfall in the United States during this period (25). Nine independent coders who were blind to the hypothesis rated the masculinity vs. femininity of historical hurricane names on two items (1 = very masculine, 11 = very feminine, and 1 = very man-like, 11 = very woman-like), which were averaged to compute a masculinity-femininity index (MFI). A series of negative binomial regression analyses (26, 27) were performed to investigate effects of perceived masculinity-femininity of hurricane names (MFI), minimum pressure, normalized

PNAS

Female hurricanes are not deadlier than male hurricanes

Jung et al. (1) assert that hurricanes that made landfall in the United States killed more people when they had female names rather than male names. The article has stirred much controversy. Criticisms range from the inclusion of hurricanes from the era before they were given male names (2) ... lective interpretation and the ove... their results from the archival ... of their hypothesis (3), to the e... of their six behavioral experim... populations in at-risk situation...

The criticism of this letter ... one: the results of their archi... function of the selective incl... sors. Using the same data, m... variables, I show in Table 1 ... are not robust to the inclu... two-way interaction they o... analysis. Model 1 reprodu... main results. Models 2–4 s... that female- and male-n... were equally deadly cann... the interaction effect of ... metric pressure and its r... toll is included. A more ... letter should have stated ...

Models 2–4 show th... lower barometric pressu... tolls and that hurricanes ...

tolls had smaller death tolls when the hurricanes were strong (lower pressure), but higher death tolls when the hurricanes were weak (higher pressure). The latter result is driven by the pre-1978 sample (model 5). In the post-1978 sample, the interaction effect ...insignificant and the damage toll ...relationship

safety infrastruct... other characteri... after 1978. Eve... of safety infras... and an ove... vergence of... might expl... the death...

Table 1. Results from

Weather and Climate Extremes 12 (2016) 80–84

Contents lists available at ScienceDirect

ELSEVIER

Weather and Climate Extremes

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

Hurricane names: A bunch of hot air?

Gary Smith *

Department of Economics, Pomona College, United States

ARTICLE INFO

Article history:
Receiv...

...-named hurricanes are deadlier because people do not take them ser... ...is based on a questionable statistical analysis of a narrowly defined data ...not robust in that it is not confirmed by a straightforward analysis of ...ata.

...on (less than 39 mph), tropical storm (39–73 mph), hur... ...more than 73 mph), and major hurricane (more than ... Tropical storms and hurricanes are generally given ... Hurricane Sandy, but tropical depressions are not. ...al. (2014) examine a narrowly defined dataset: U.S. ...m Atlantic hurricanes that made landfall in the United ...n a strong, surprising conclusion is drawn from re... ...it can be instructive to see whether the conclusion is ...espect to the myriad decisions used to restrict the ...re are several issues: ...clude tropical storms? In 1994 Tropical Storm Al... ...lfall near Destin, Florida, with maximum sustained ...h and caused historic flooding in Alabama and ...lted in at least 30 deaths and caused $1 bill... ...4 dollars). Alberto was classif... ...l a hurricane, bec... ...mph th...

CrossMark

LETTER

LETTER

Are female hurricanes really deadlier than male hurricanes?

Jung et al. (1) claim to show that "feminine-named hurricanes cause significantly more deaths than do masculine-named hurricanes" (p. 1). This conclusion is mainly obtained by analyzing data on fatalities caused by hurricanes in the United States (1950–2012). By reanalyzing the same data, we show that the conclusion is based on biased presentation and invalid statistics.

The reasoning in ref. 1 is fundamentally based on the regression models reported in their table S2, in particular, model 4. However, due to the interaction terms combined with extreme values and weak significance, the analysis is based on a very fragile model; e.g., the model predicts almost 20,000 deaths for hurricane Sandy, which actually caused

Now, we explain our claim that the results are presented in a biased way. By holding the minimum pressure at its mean in prediction of counts of deaths, the authors only report the influence of MFI and normalized damage (figure 1 in ref. 1). This ignores the influence of the second interaction term MFI minimum pressure, which shows an opposite influence (see the estimated parameters on p. 5, first paragraph). By considering the counts of deaths under constant normalized damage, the results are contrary: male-named hurricanes with a low minimum pressure (strong hurricanes) are associated with more deaths than female ones (Fig. 1).

In the light of an alternating male-female

differences between male- or female-named hurricanes for deaths, minimum pressure, category, and damages.

To conclude, the analyses given in ref. 1 are examples of the fact that prediction models using interaction terms have to be handled and interpreted carefully; in particular, using insignificant variables is not expedient and may lead to statistical artifacts.

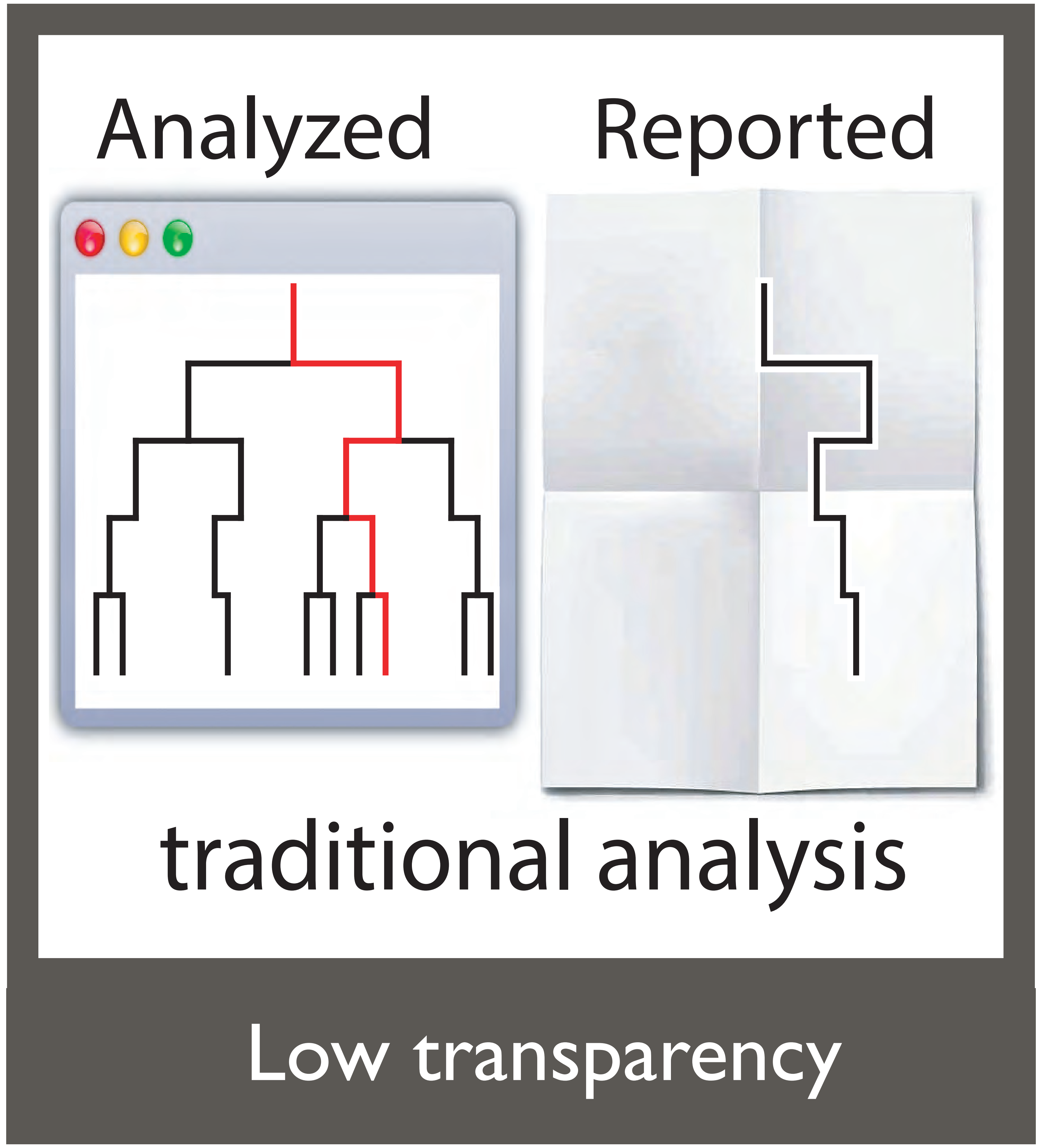
To summarize, the data do not contain evidence that feminine-named hurricanes cause more deaths than masculine-named hurricanes.

Björn Christensen[a] *and*
Sören Christensen[b,1]

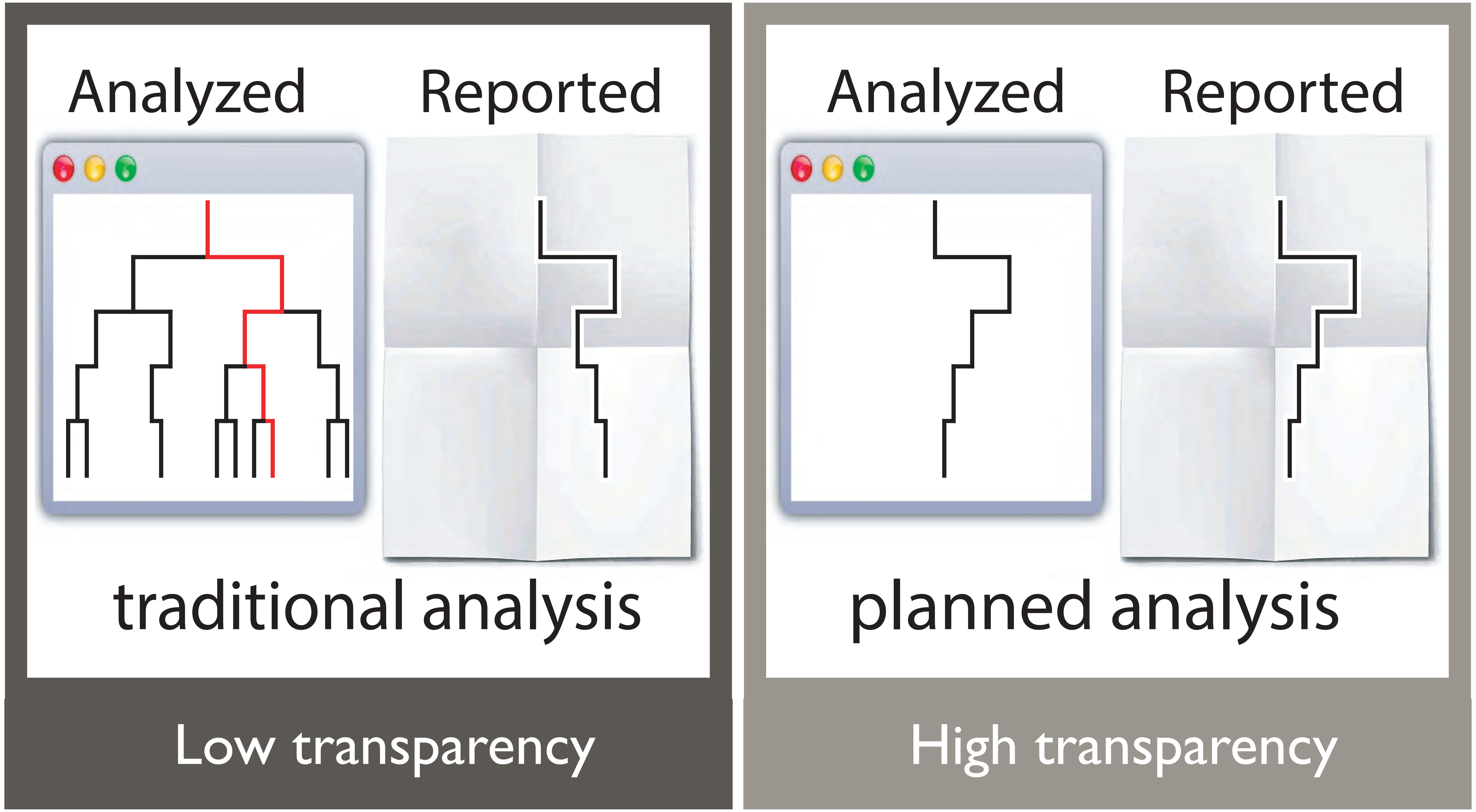
“If you torture the data long enough, it will confess.”

– Ronald Coase, 1988 (Nobel Laureate)

Statistical Analysis & Reporting



Statistical Analysis & Reporting




Do soccer referees give more red cards to dark-skinned players than light-skinned ones?
DE JONG
34
KOMPANY
4

Same Data, Many Possible Analyses

Zixin's plan

Skin colour:
3-levels: dark, medium, light

Co-variate use:
Control for positions (defensive players may commit more fouls)

Transformations:
Average skin-tone ratings (mean)

Exclusions:
Player-referee with no red card

Sophia's plan

Skin colour:
5 levels: (very) dark, medium, (very) light

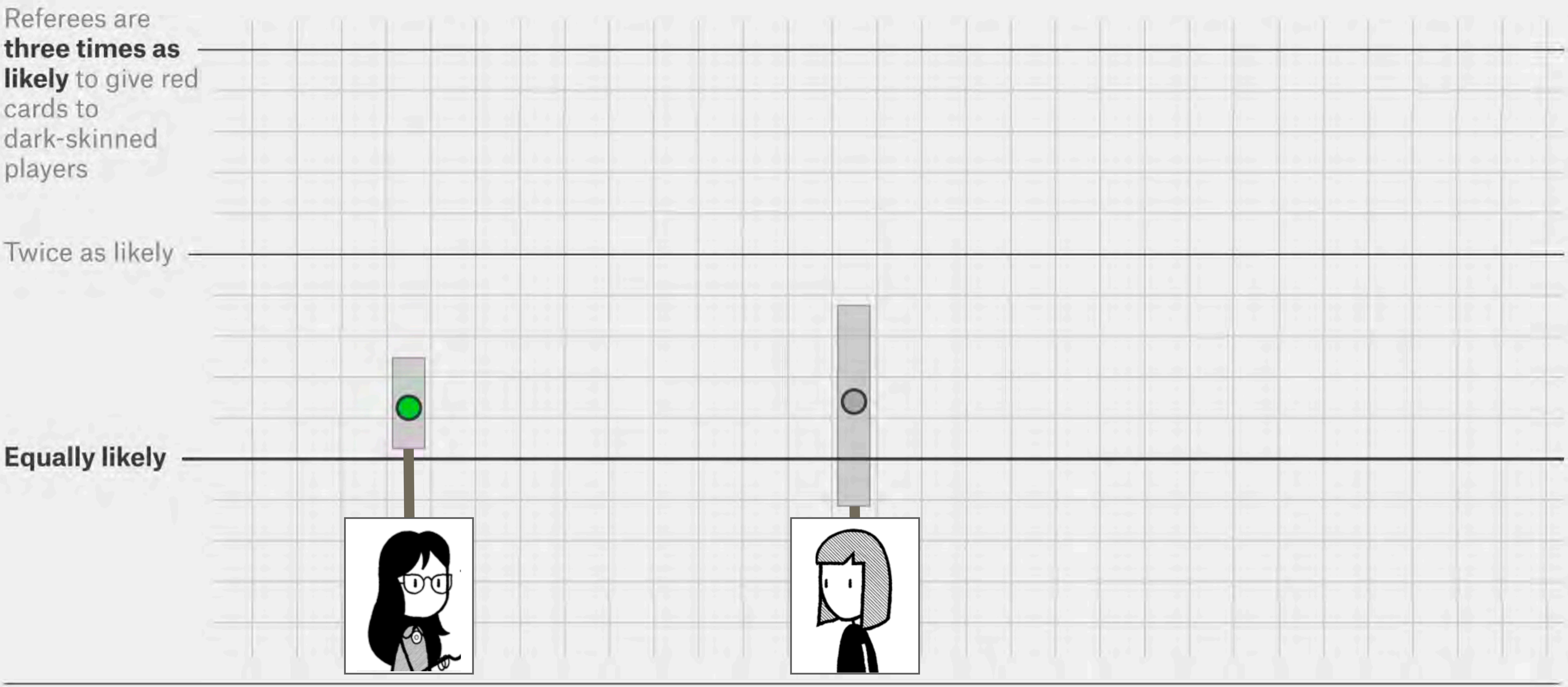
Co-variate use:
Control for referees' skin colour

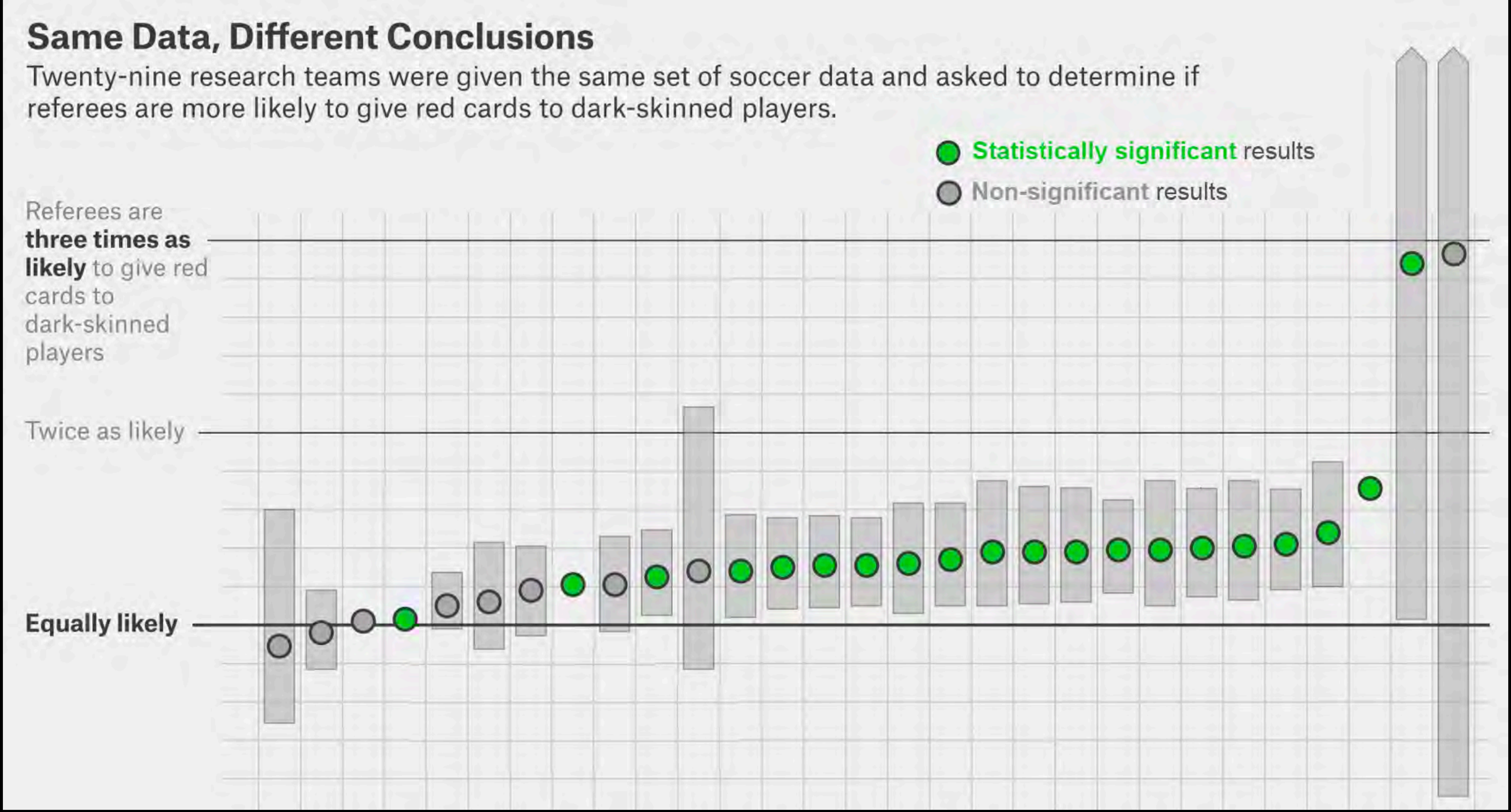
Transformations:
Max (darkest) skin-tone rating

Exclusions:
Games with no red card

Analytic approach:

Same Data, Different Conclusions



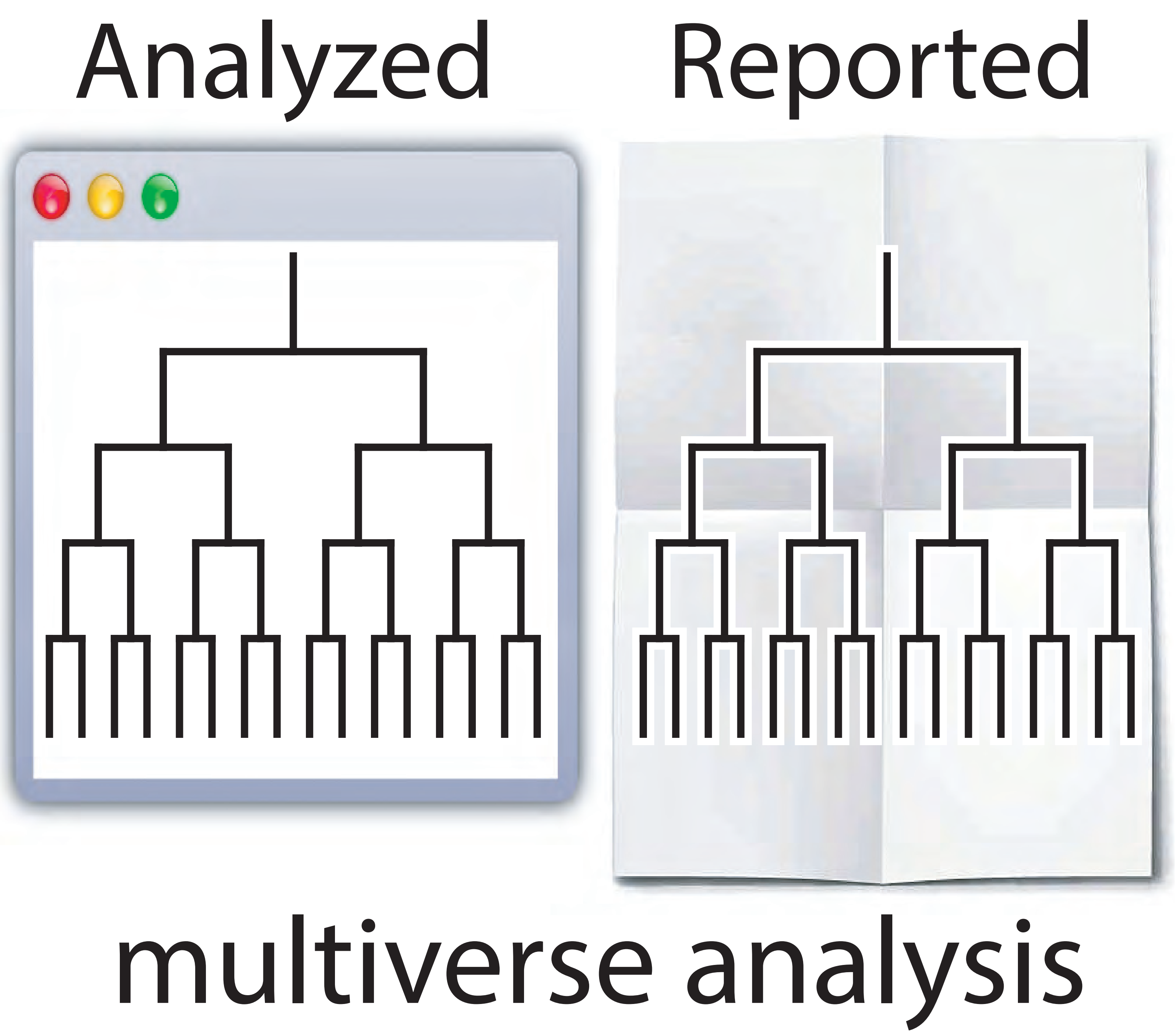


Brian A. Nosek et al. (2015) Many analysts, one dataset: Making transparent how variations in analytical choices affect results

Same Data, Many Possible Analyses

Multiverse Analysis

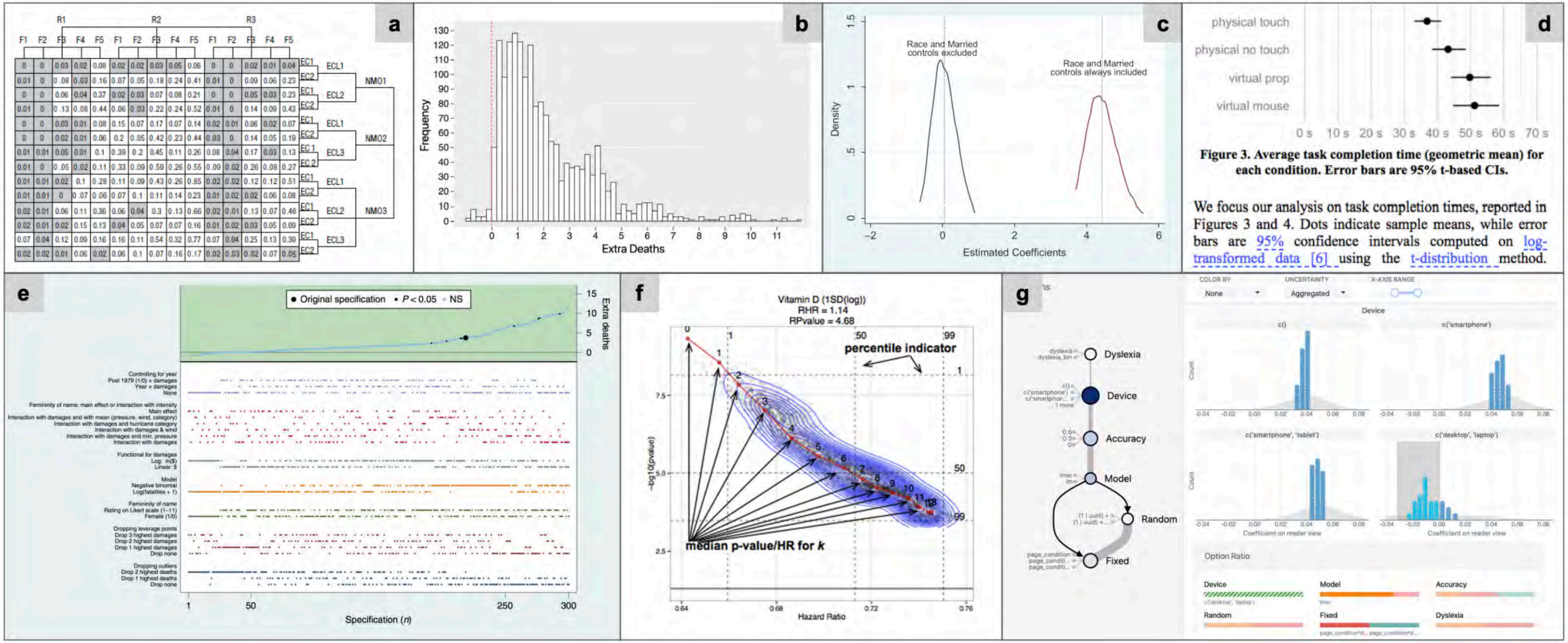
Performing and reporting all analyses corresponding to a large set of reasonable analysis scenarios.



Steegen et al. (2016) **Increasing Transparency Through a Multiverse Analysis**
Simonsohn et al. (2015) **Specification Curves: Descriptive and Inferential Statistics on All Reasonable Specifications**

How to report all of these analyses?

Tasks and Visualizations



Brian D. Hall, Yang Liu, Yvonne Jansen, Pierre Dragicevic, Fanny Chevalier, Matthew Kay. **A Survey of Tasks and Visualizations in Multiverse Analysis Reports** COMPUTER GRAPHICS forum. 2021

Tasks and Visualizations

	Name	Section	Icon	Composition ▸ Process	Composition ▸ Parameters	Outcome ▸ Range	Outcome ▸ Frequency	Connect ▸ Outcome Range	Connect ▸ Outcome Frequency	Connect ▸ Specific Outcomes	Connect Combo ▸ Outcome Range	Connect Combo ▸ Outcome Frequency	Connect Combo ▸ Idiosyncratic	Validate ▸ Metrics	Validate ▸ Details
Archetypes	Outcome Histogram	6.1		0	0	3	3	0	0	0	0	0	0	0	0
	Outcome Curve	6.2.1		0	0	3	2	0	0	0	0	0	0	0	0
	Universe Specification Panel	6.2.2		0	2	0	0	0	0	0	0	0	0	0	0
	Descriptive Specification Curve	6.2.3		0	2	3	2	3	2	3	2	1	3	0	0
	Outcome Density Plot	6.3		0	1	3	3	2	2	1	2	2	3	0	0
	Vibration of Effects Plot	6.4		0	0	3	2	2	2	1	1	1	3	0	0
	Outcome Matrix	6.5		0	1	3	2	2	2	3	2	2	3	0	0
	Multiverse Computation Schematic	6.6		3	3	0	0	0	0	0	0	0	0	0	0
Systems	Explorable Multiverse Analysis Reports	6.7.1		0	2	1	1	1	1	1	1	1	0	0	3
	Boba	6.7.2		3	3	3	3	3	3	3	1	1	3	3	0

Category	Task
Composition	**Composition ▷ Process**: understand the process that defines and creates the universes being considered. **Composition ▷ Parameters**: understand the definition and composition of universe parameters and parameter values.
Outcome	**Outcome ▷ Range**: assess range or spread of outcome values across all universes. **Outcome ▷ Frequency**: assess overall frequency of outcome values across all universes.
Connect	**Connect ▷ OutcomeRange**: connect parameters to outcomes by comparing similarity or range of outcome values across a subset of universes defined by a specific parameter value. **Connect ▷ OutcomeFrequency**: connect parameters to outcomes by comparing frequency of outcome values across a subset of universes defined by a specific parameter value. **Connect ▷ SpecificOutcomes**: connect parameters to outcomes by examining specific outcome values of interest and identifying parameter values that lead to those outcomes.
Connect Combinations	**ConnectCombo ▷ OutcomeRange**: connect combinations of parameters to outcomes by comparing range of outcome values across subsets of universes defined by parameter values. **ConnectCombo ▷ OutcomeFrequency**: connect combinations of parameters to outcomes by comparing frequency of outcome values across subsets of universes defined by parameter values. **ConnectCombo ▷ Idiosyncratic**: connect combinations of parameters to outcomes according to idiosyncratic patterns particular to a given visualization or analysis.
Validate	**Validate ▷ Metrics**: assess validity metrics of universes or com-pare metrics across parameter values. **Validate ▷ Details**: assess validity of universes by examining the underlying details of analyses in each universe to interrogate their validity.

Brian D. Hall, Yang Liu, Yvonne Jansen, Pierre Dragicevic, Fanny Chevalier, Matthew Kay. **A Survey of Tasks and Visualizations in Multiverse Analysis Reports** COMPUTER GRAPHICS forum. 2021

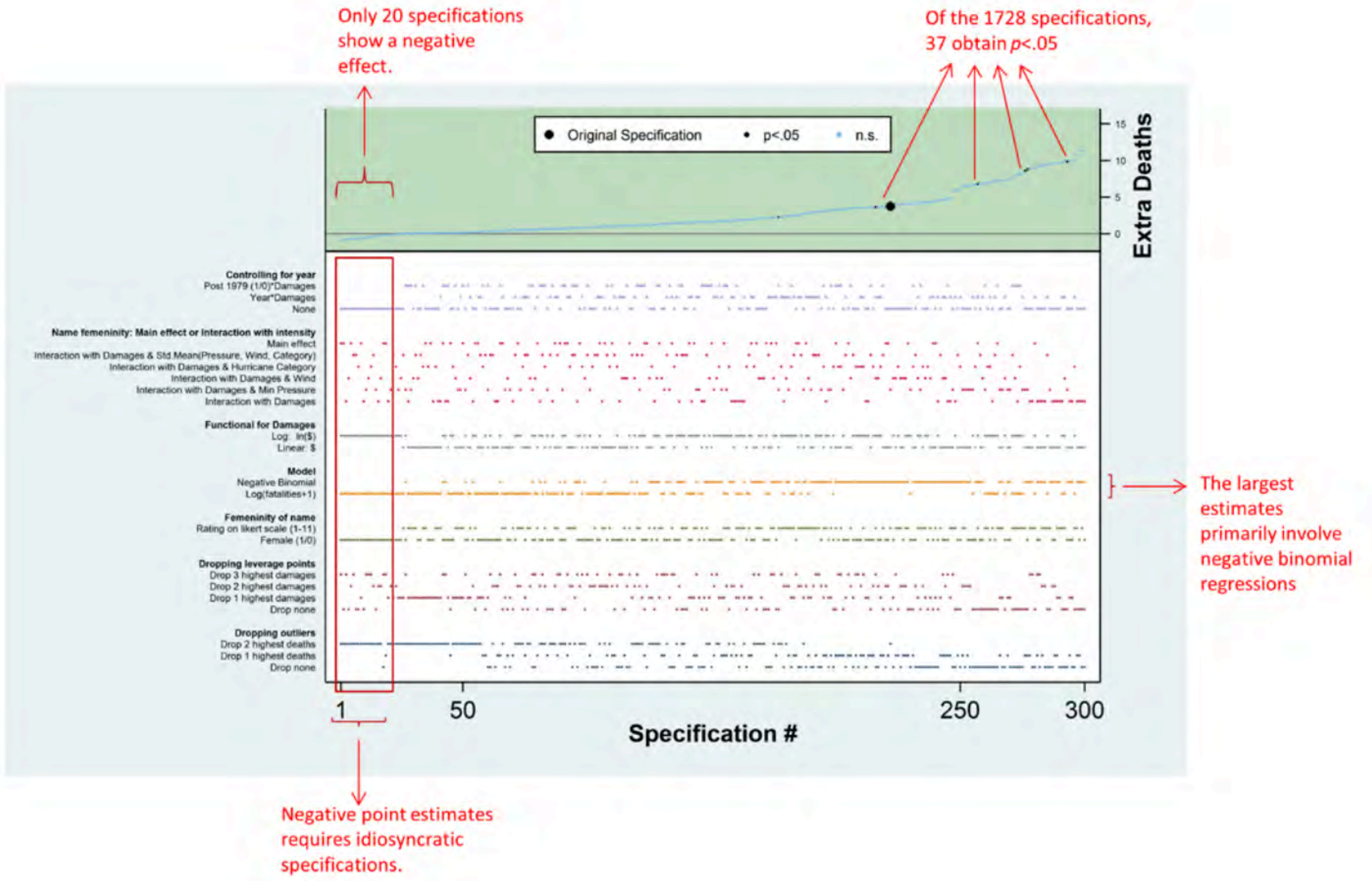
Category	Task
Composition	**Composition ▷ Process**: understand the process that defines and creates the universes being considered. **Composition ▷ Parameters**: understand the definition and composition of universe parameters and parameter values.
Outcome	**Outcome ▷ Range**: assess range or spread of outcome values across all universes. **Outcome ▷ Frequency**: assess overall frequency of outcome values across all universes.
Connect	**Connect ▷ OutcomeRange**: connect parameters to outcomes by comparing similarity or range of outcome values across a subset of universes defined by a specific parameter value. **Connect ▷ OutcomeFrequency**: connect parameters to outcomes by comparing frequency of outcome values across a subset of universes defined by a specific parameter value. **Connect ▷ SpecificOutcomes**: connect parameters to outcomes by examining specific outcome values of interest and identifying parameter values that lead to those outcomes.
Connect Combinations	**ConnectCombo ▷ OutcomeRange**: connect combinations of parameters to outcomes by comparing range of outcome values across subsets of universes defined by parameter values. **ConnectCombo ▷ OutcomeFrequency**: connect combinations of parameters to outcomes by comparing frequency of outcome values across subsets of universes defined by parameter values. **ConnectCombo ▷ Idiosyncratic**: connect combinations of parameters to outcomes according to idiosyncratic patterns particular to a given visualization or analysis.
Validate	**Validate ▷ Metrics**: assess validity metrics of universes or com-pare metrics across parameter values. **Validate ▷ Details**: assess validity of universes by examining the underlying details of analyses in each universe to interrogate their validity.

Brian D. Hall, Yang Liu, Yvonne Jansen, Pierre Dragicevic, Fanny Chevalier, Matthew Kay. **A Survey of Tasks and Visualizations in Multiverse Analysis Reports** COMPUTER GRAPHICS forum. 2021

Visual Summaries

			R1					R2					R3				
			F1	F2	F3	F4	F5	F1	F2	F3	F4	F5	F1	F2	F3	F4	F5
NMO1	ECL1	EC1	0	0	0	0	0	0.18	0.19	0.3	0.07	0.21	0.01	0.01	0.04	0	0.03
		EC2	0.01	0.01	0.02	0	0.02	0.93	0.76	0.99	0.52	0.78	0.24	0.15	0.37	0.06	0.23
	ECL2	EC1	0	0	0	0	0	0.17	0.26	0.3	0.13	0.27	0.02	0.02	0.06	0.01	0.05
		EC2	0.02	0.01	0.04	0	0.01	0.89	0.86	0.98	0.47	0.66	0.24	0.2	0.42	0.05	0.19
NMO2	ECL1	EC1	0.01	0.01	0.01	0	0	0.47	0.45	0.21	0.06	0.07	0.09	0.11	0.04	0.01	0.01
		EC2	0.06	0.05	0.02	0.01	0	0.94	0.93	0.4	0.24	0.13	0.4	0.41	0.13	0.06	0.02
	ECL3	EC1	0.01	0.01	0	0	0	0.49	0.45	0.22	0.08	0.1	0.1	0.11	0.04	0.01	0.02
		EC2	0.1	0.11	0.02	0.02	0	0.87	0.91	0.32	0.27	0.18	0.42	0.48	0.12	0.09	0.04
NMO3	ECL1	EC1	0	0	0.01	0	0.02	0.19	0.22	0.1	0.07	0.29	0.02	0.03	0.02	0.01	0.07
		EC2	0.01	0	0.01	0	0.01	0.46	0.62	0.08	0.08	0.1	0.1	0.14	0.02	0.01	0.02
	ECL2	EC1	0.01	0	0.01	0	0.01	0.37	0.27	0.18	0.09	0.23	0.07	0.05	0.03	0.01	0.04
		EC2	0.03	0.02	0.01	0.02	0.01	0.76	0.84	0.13	0.16	0.1	0.24	0.29	0.04	0.04	0.02
	ECL3	EC1	0.01	0.01	0.02	0	0	0.37	0.35	0.18	0.09	0.24	0.08	0.08	0.05	0.01	0.04
		EC2	0.05	0.04	0.03	0.01	0	0.9	0.97	0.22	0.18	0.16	0.36	0.45	0.09	0.04	0.02

Steegen et al. (2016) **Increasing Transparency Through a Multiverse Analysis**



Simonsohn et al. (2015) **Specification Curves: Descriptive and Inferential Statistics on All Reasonable Specifications**

Can we do better?

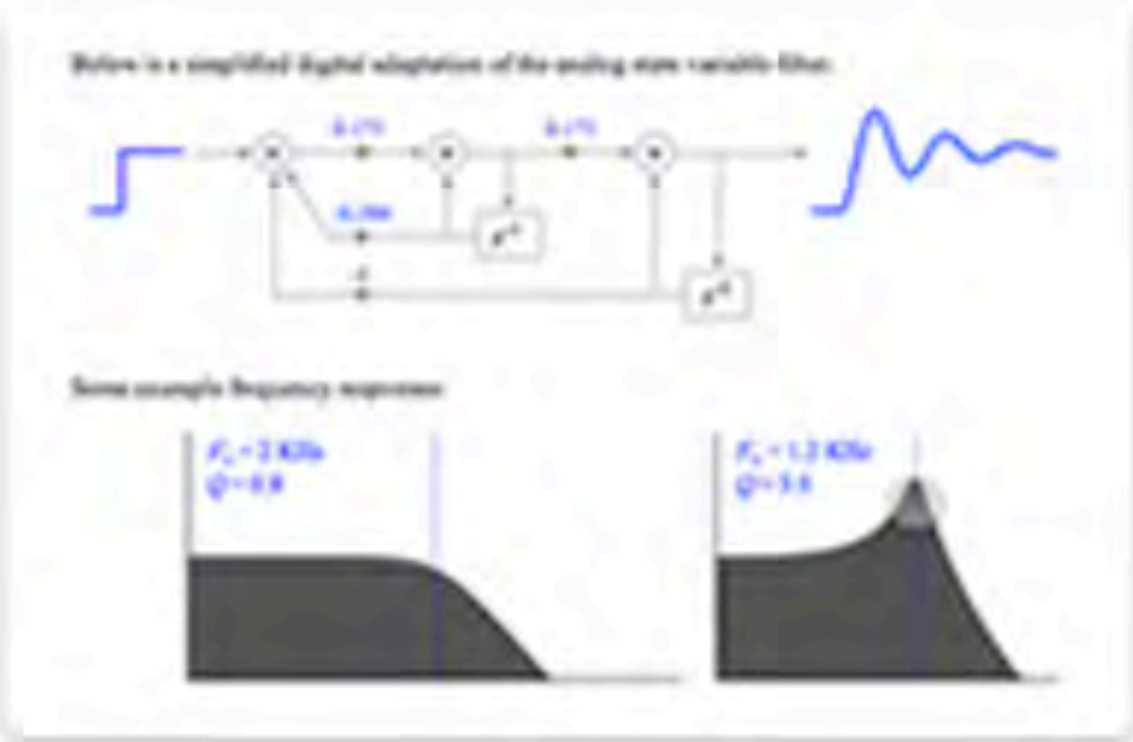
Explorable Explanations

Bret Victor / March 10, 2011

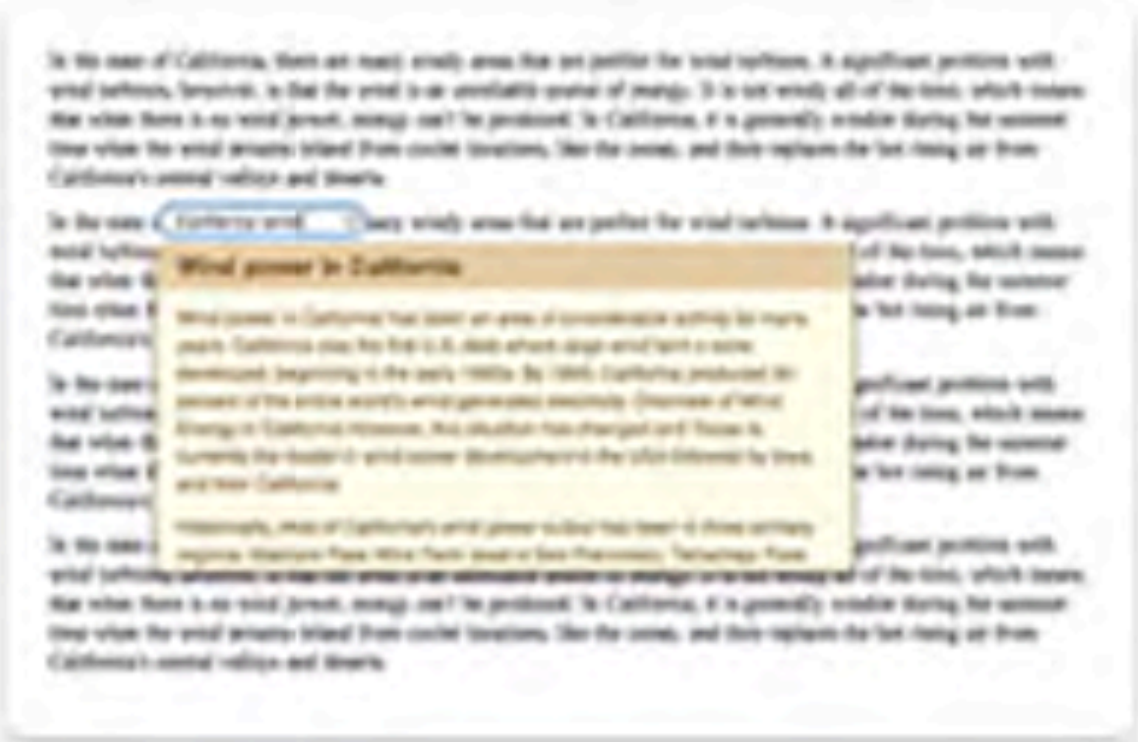
What does it mean to be an **active reader**?

A **reactive document** allows the reader to play with the author's assumptions and analyses, and see the consquences.

An **explorable example** makes the abstract concrete, and allows the reader to develop an intuition for how a system works.



Contextual information allows the reader to learn related material just-in-time, and cross-check the author's claims.

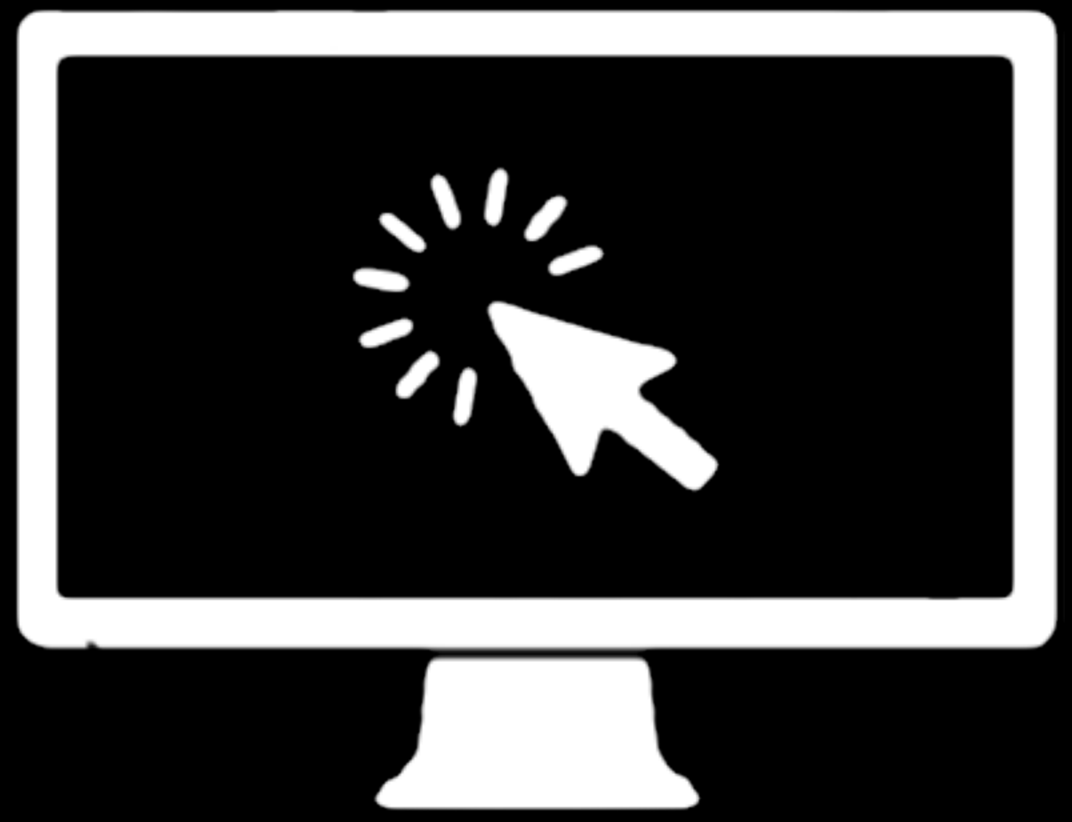


http://worrydream.com/ExplorableExplanations/

Explorable Multiverse Analysis Report (**EMAR**)

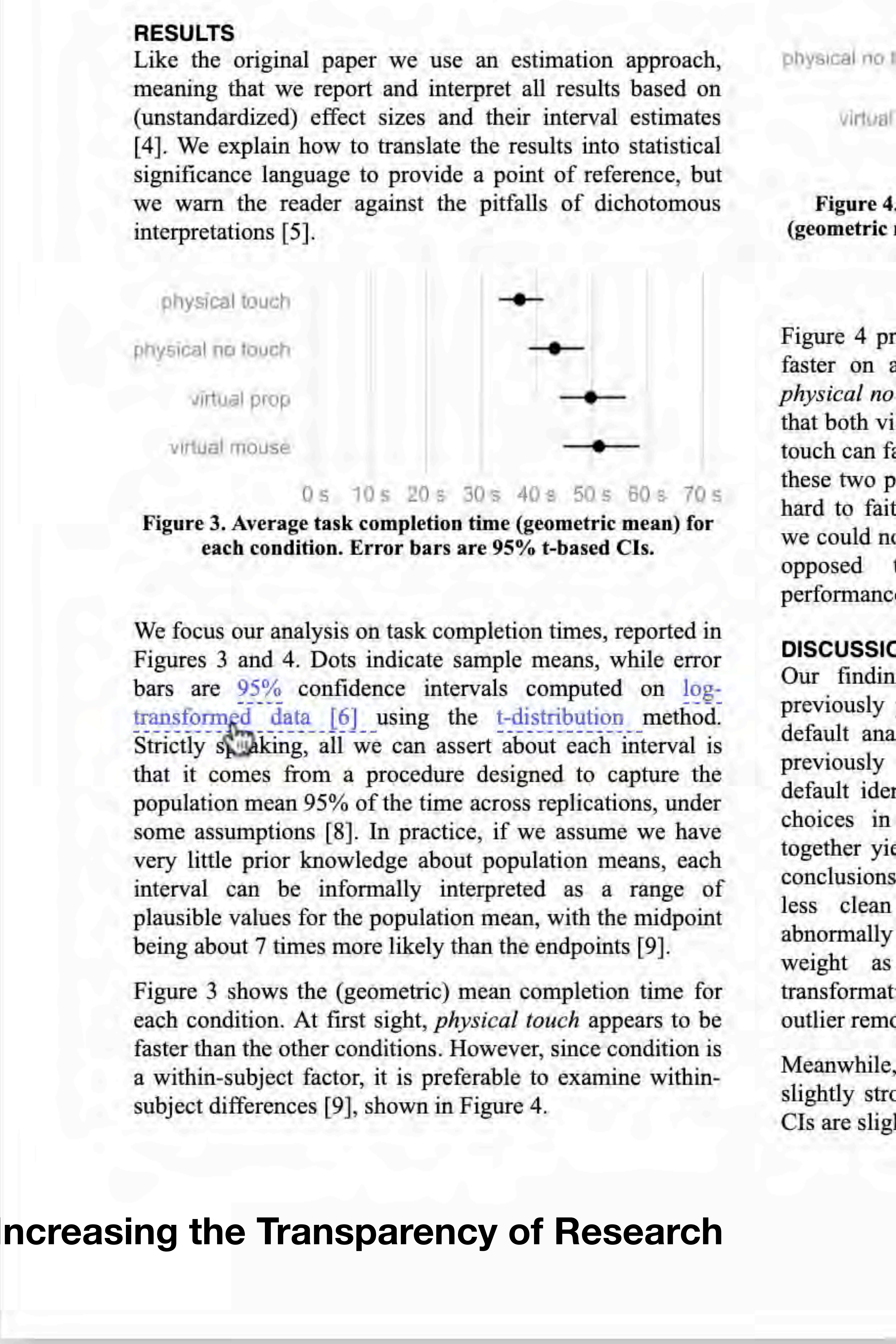
Pierre Dragicevic, Yvonne Jansen, Abhraneel Sarma, Matthew Kay, Fanny Chevalier **Increasing the Transparency of Research Papers with Explorable Multiverse Analyses.** ACM CHI. 2019.

https://explorablemultiverse.github.io/



Promote and *support* transparent statistical reporting

- More trustworthy
- More interpretable
- Easier to verify
- Easier to replicate



RESULTS

Like the original paper we use an estimation approach, meaning that we report and interpret all results based on (unstandardized) effect sizes and their interval estimates [4]. We explain how to translate the results into statistical significance language to provide a point of reference, but we warn the reader against the pitfalls of dichotomous interpretations [5].

Figure 3. Average task completion time (geometric mean) for each condition. Error bars are 95% t-based CIs.

We focus our analysis on task completion times, reported in Figures 3 and 4. Dots indicate sample means, while error bars are 95% confidence intervals computed on log-transformed data [6] using the t-distribution method. Strictly speaking, all we can assert about each interval is that it comes from a procedure designed to capture the population mean 95% of the time across replications, under some assumptions [8]. In practice, if we assume we have very little prior knowledge about population means, each interval can be informally interpreted as a range of plausible values for the population mean, with the midpoint being about 7 times more likely than the endpoints [9].

Figure 3 shows the (geometric) mean completion time for each condition. At first sight, *physical touch* appears to be faster than the other conditions. However, since condition is a within-subject factor, it is preferable to examine within-subject differences [9], shown in Figure 4.

Pierre Dragicevic, Yvonne Jansen, Abhraneel Sarma, Matthew Kay, Fanny Chevalier **Increasing the Transparency of Research Papers with Explorable Multiverse Analyses.** ACM CHI. 2019.

multiverse: R library

A. There are two other reasonable alternatives for removing outliers based on the value of death. In `multiverse`, we can do this through local modifications to the original analysis.

```r
df ← hurricane %>%
  filter(name != "Katrina" & name != "Audrey")
```

B. Decisions, also referred to as ***parameters***, are declared using `branch()`. First argument is the name of the ***parameter***. Alternate analysis are passed as arguments in the form `options_name ~ option_value`

```r
df ← hurricane %>%
  filter(branch(death_outliers,
    "no_exclusion" ~ TRUE,
    "most_extreme" ~ name != "Katrina",
    "two_most_extreme" ~ !(name %in% c("Katrina", "Audrey"))
))
```

C. `multiverse` compiles this declaration into three separate expressions

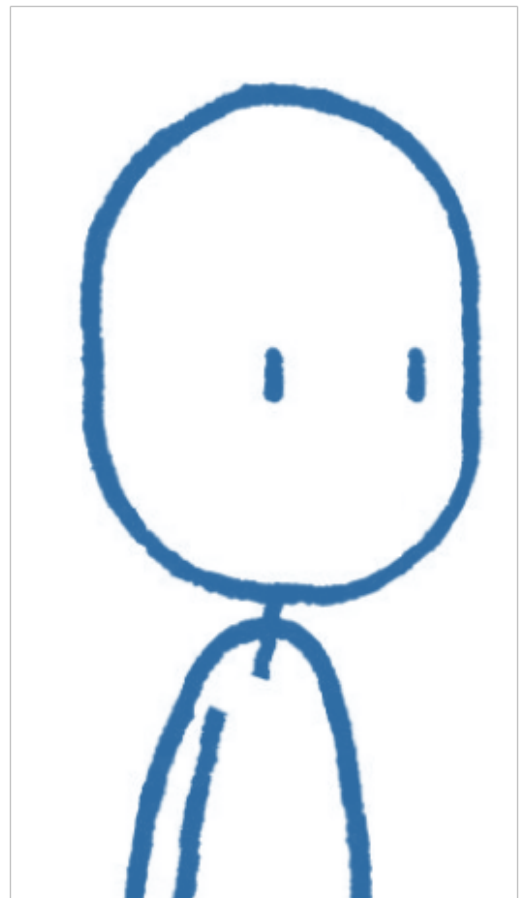
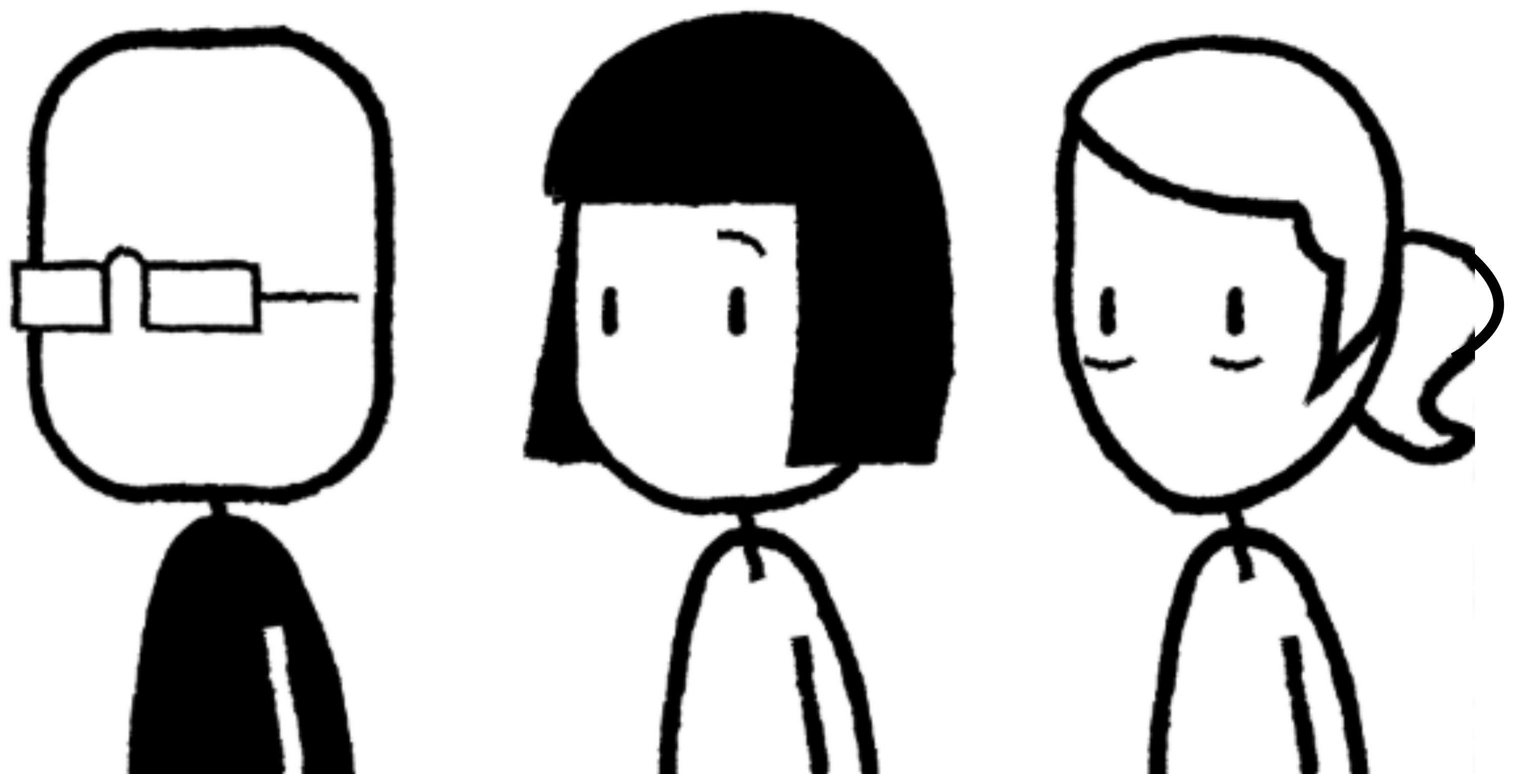
```r
df ← hurricane %>%
  filter(TRUE)
```

```r
df ← hurricane %>%
  filter(name != "Katrina)
```

```r
df ← hurricane %>%
  filter(name != "Katrina" & name != "Audrey")
```

Abhraneel Sarma, Alex Kale, Michael Moon, Nathan Taback, Fanny Chevalier, Jessica Hullman and Matthew Kay **multiverse: Multiplexing Alternative Data Analyses in R Notebooks** ACM CHI. 2023.

Scientists make sense of data
... communicate to other scientists
... and to lay audiences.

Scientists | … to other scientists | … to lay audiences

Weighted Graph Comparison Techniques for Brain Connectivity Analysis

Basak Alper
UCSB
basakalper@gmail.com

Benjamin Bach
INRIA
benjamin.bach@inria.fr

Nathalie Henry Riche
Microsoft Research
nath@microsoft.com

Tobias Isenberg
INRIA
tobias.isenberg@inria.fr

Jean-Daniel Fekete
INRIA
jean-daniel.fekete@inria.fr

ABSTRACT

The analysis of brain connectivity is a vast field in neuroscience with a frequent use of visual representations and an increasing need for visual analysis tools. Based on an in-depth literature review and interviews with neuroscientists, we explore high-level brain connectivity analysis tasks that need to be supported by dedicated visual analysis tools. A significant example of such a task is the comparison of different connectivity data in the form of weighted graphs. Several approaches have been suggested for graph comparison within information visualization, but the comparison of *weighted* graphs has not been addressed. We explored the design space of applicable visual representations and present augmented adjacency matrix and node-link visualizations. To assess which representation best support weighted graph comparison tasks, we performed a controlled experiment. Our findings suggest that matrices support these tasks well, outperforming node-link diagrams. These results have significant implications for the design of brain connectivity analysis tools that require weighted graph

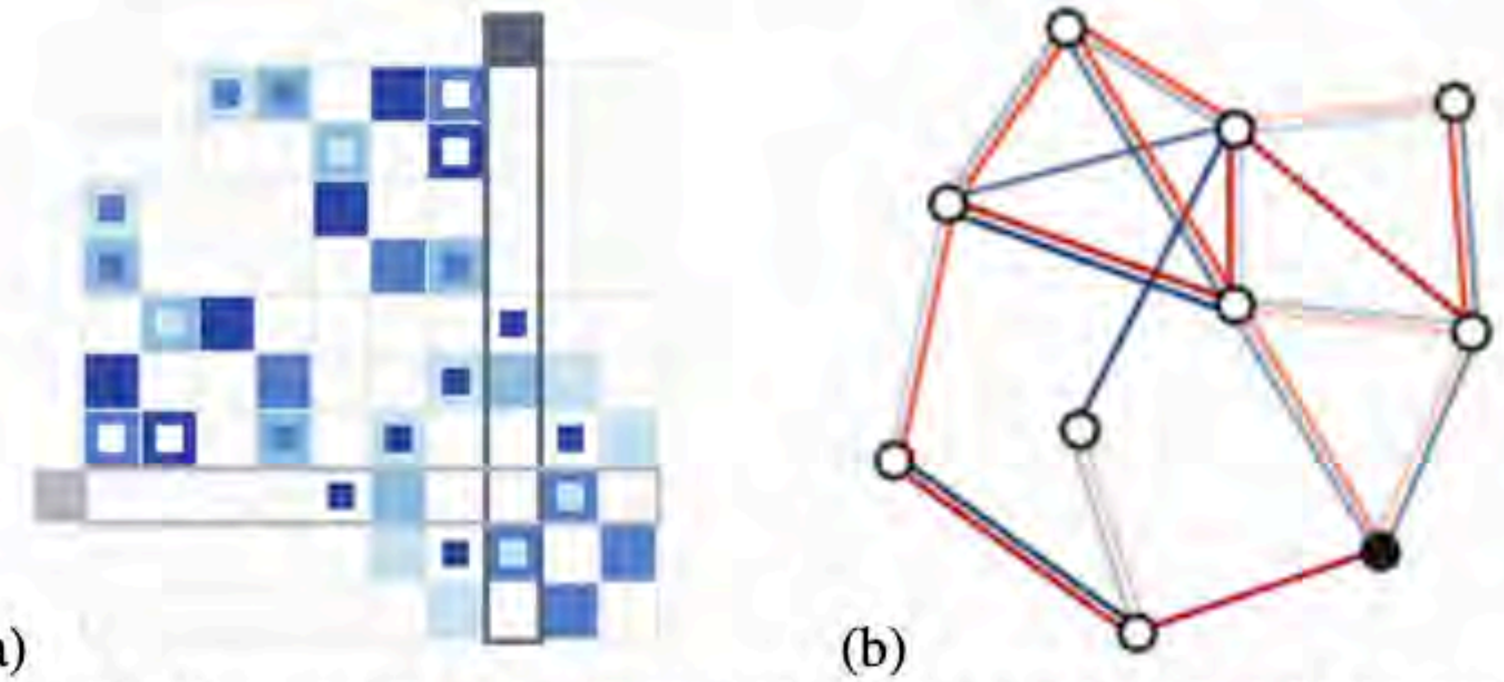

(a) (b)

Figure 1. Alternative superimposed (a) matrix and (b) node-link visualizations supporting weighted graph comparisons.

ral groupings or anatomically segregated brain regions [26]. Although, the processes for acquiring the two types of connectivity data differ, they can both be expressed as weighted graphs in which nodes represent ROIs and edges can encode the strength of their correlation (*functional*) or the density of fibers connecting them (*anatomical*).

Scientists | … to other scientists | … to lay audiences

NEW YORK POST

LIFESTYLE

Hair dyes could raise risk of breast cancer: study

By Tamar Lapin

Published Oct. 14, 2017, 8:17 p.m. ET

The Washington Post

Democracy Dies in Darkness

Are hair dyes safe? Health worries are increasing interest in the go-gray style trend.

January 26, 2020 More than **5 years ago**

FAST COMPANY

12-04-2019 | NEWS

A harrowing study of 46,000 women shows hair dyes are heavily associated with cancer

Scientists | … to other scientists | … to lay audiences

NEW YORK POST

LIFESTYLE

Hair dyes could raise risk of breast cancer: study

By Tamar Lapin
Published Oct. 14, 2017, 8:17 p.m. ET

The Washington Post
Democracy Dies in Darkness

Are hair dyes safe? Health worries are increasing interest in the go-gray style trend.

January 26, 2020 More than **5 years ago**

FAST COMPANY

12-04-2019 | NEWS

A harrowing study of 46,000 women shows hair dyes are heavily associated with cancer

Does the Use of Hair Dyes Increase the Risk of Developing Breast Cancer? A Meta-analysis and Review of the Literature

Ritika Gera [1, 2], Ramia Mokbel [3], Ivanna Igor [1], Kefah Mokbel [4]

Affiliations + expand

PMID: 29374694 DOI: 10.21873/anticanres.12276

Abstract

Background/aim: Hair dye may contain mutagenic compounds which could be associated with an increased incidence of breast cancer in women who use it. The aim of this study was to examine the association between the personal use of hair dyes and the risk of breast cancer.

Materials and methods: We conducted a literature review of epidemiological studies reporting breast cancer-specific risks among hair dye users versus non-users. The data for the incidence of breast cancer following the 'ever' use of hair dye in studies which met the inclusion criteria was analysed using a meta-analysis. The relative risk ratio (RR) and 95% confidence intervals (CI) were determined.

Results: A total of eight case-control studies published between 1980 and 2017 met the selection criteria and were included in the meta-analysis. Compared to non-users, using a random effects model and the Duval and Tweedie's trim and fill procedure to adjust for publication bias in the presence of between studies heterogeneity, the adjusted RR for women using hair dyes was 1.1885 (95% CI=1.03228-1.36835). This indicates an 18.8% increased risk of future development of breast cancer among hair dye users.

Conclusion: Although further work is required to confirm our results and clarify potential mechanisms, our findings suggest that exposure to hair dyes may contribute to an increased breast cancer risk.

Scientists | … to other scientists | … to lay audiences

NEW YORK POST

LIFESTYLE

Hair dyes could raise risk of breast cancer: study

By Tamar Lapin

Published Oct. 14, 2017, 8:17 p.m. ET

Women who frequently dye their hair may be at greater risk of contracting breast cancer, a new study found.

The Washington Post

Democracy Dies in Darkness

Are hair dyes safe? Health worries are increasing interest in the go-gray style trend.

January 26, 2020 More than **5 years ago**

When Keanu Reeves walked into a Los Angeles gala holding hands with artist Alexandra Grant, fans applauded the 55-year-old actor for choosing an "age appropriate" romantic partner. Most striking about Grant, 46, was her steel-gray hair.

Why wasn't she coloring it? In an Instagram post, she explained: In her 20s, she began graying, and she covered it with various shades of dye until she could no longer tolerate the chemicals.

FAST COMPANY

12-04-2019 | NEWS

A harrowing study of 46,000 women shows hair dyes are heavily associated with cancer

Scientists | … to other scientists | … to lay audiences

NEW YORK POST

LIFESTYLE

Hair dyes could raise risk of breast cancer: study

Women who frequently dye their hair may be at greater risk of contracting breast cancer, a new study found.

“Headlines should only be used to decide whether to read the article or not. They’re written to grab eyeballs and they’re often inflammatory and not scientific.”

– Fred Hutch biostatistician Dr. Ruth Etzioni

January 26, 2020 More than 5 years ago

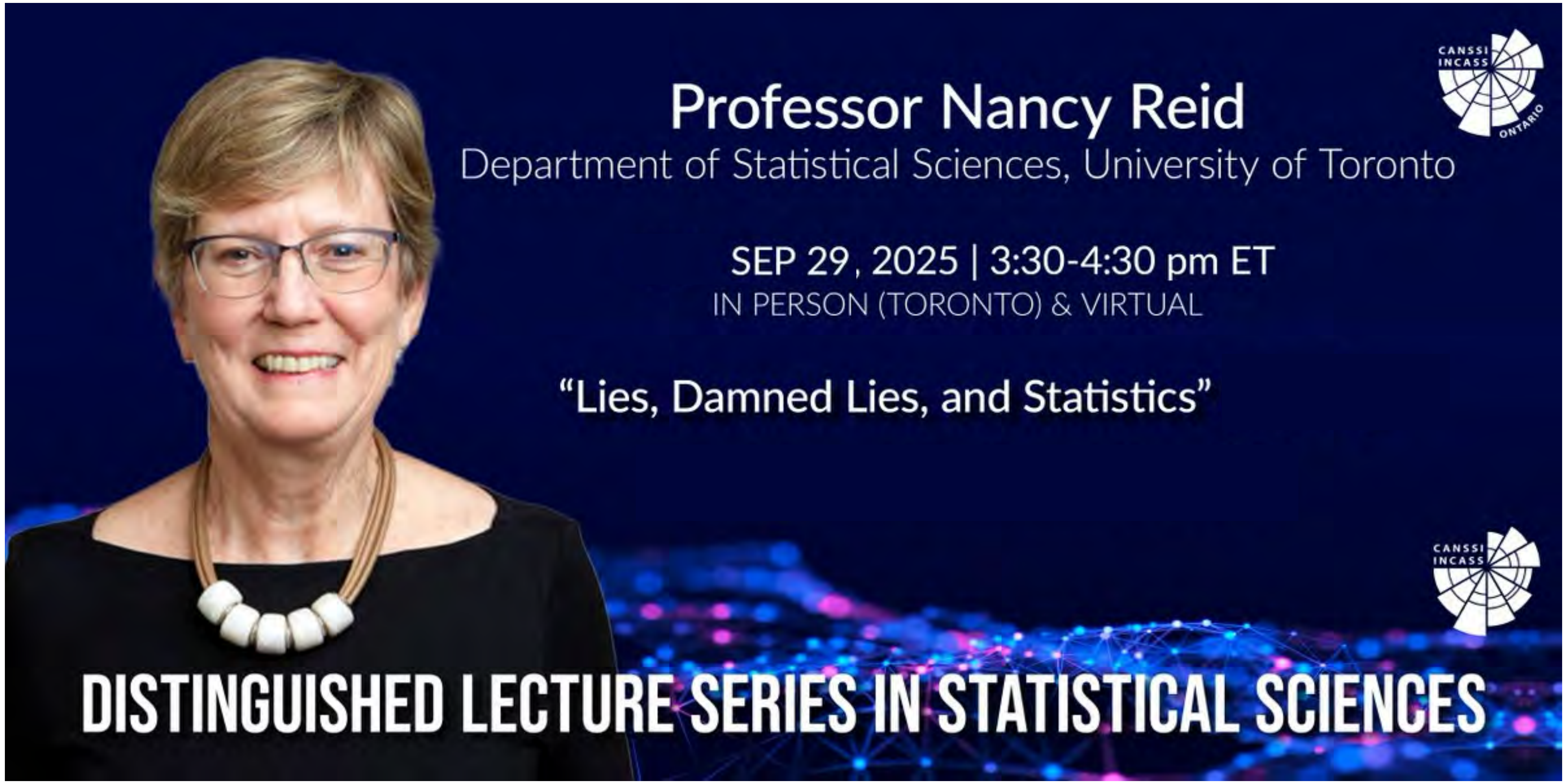
In her 20s, she began graying, and she covered it with various shades of dye until she could no longer tolerate the chemicals.

FAST COMPANY

12-04-2019 | NEWS

A harrowing study of 46,000 women shows hair dyes are heavily associated with cancer

Scientists | … to other scientists | … to lay audiences



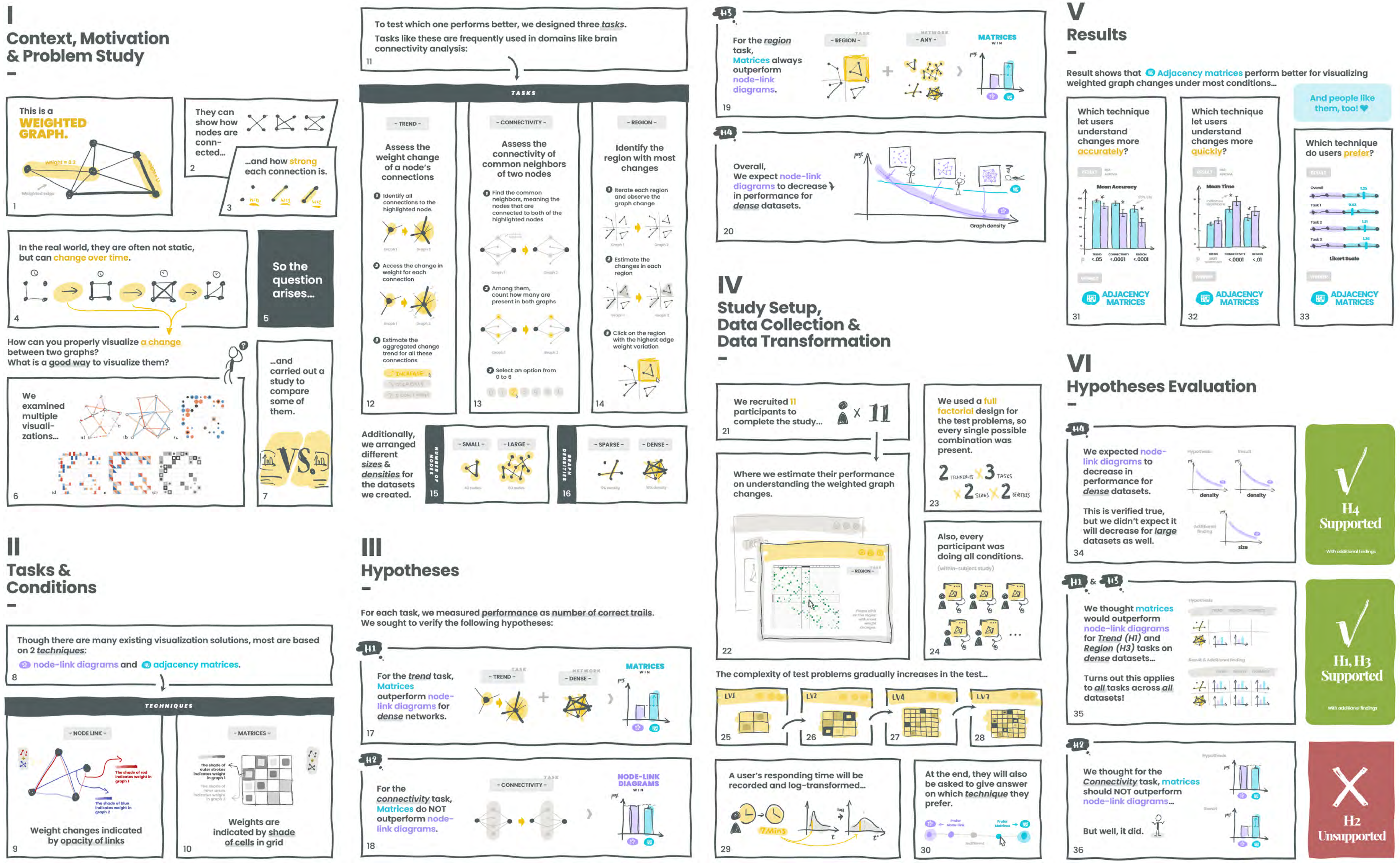
YouTube

Can we do better?

Scientists

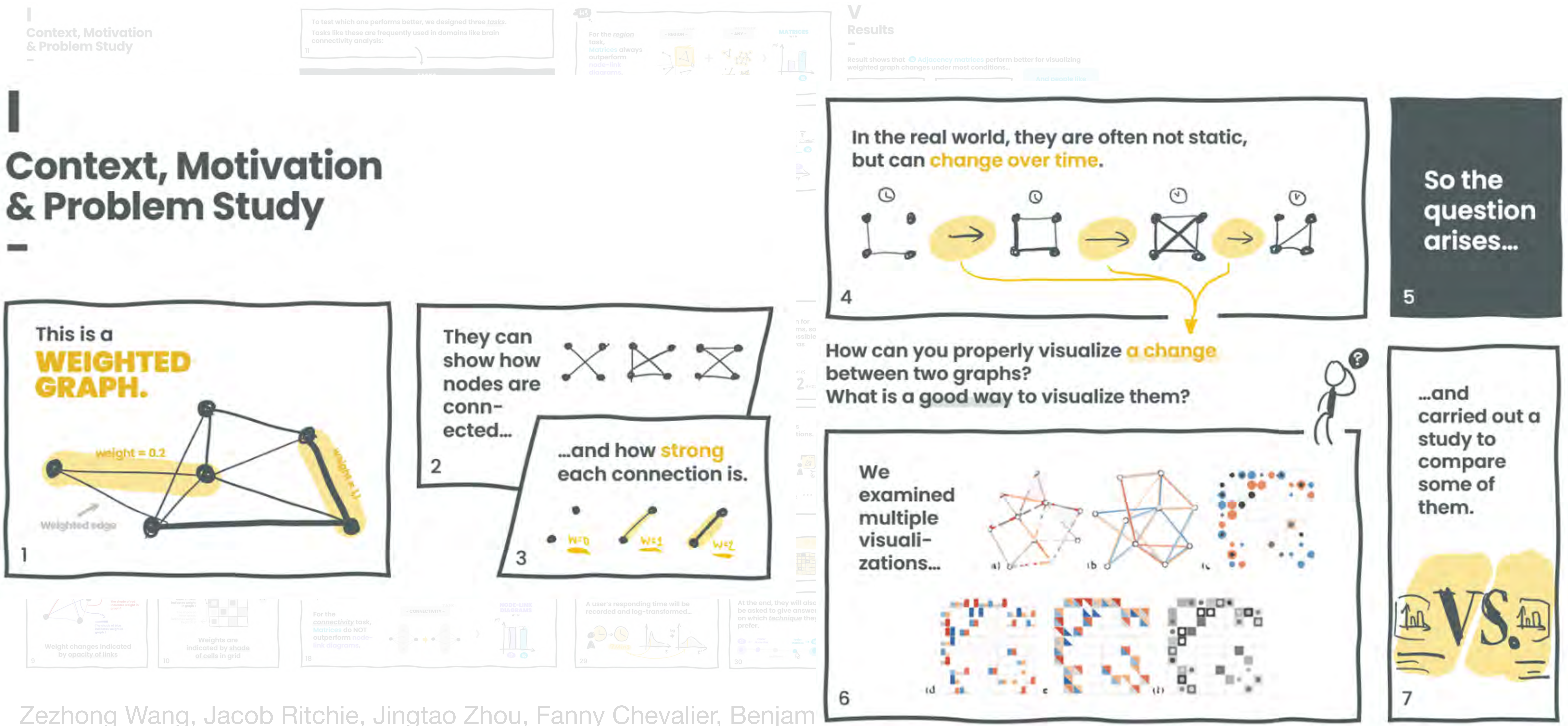
… to other scientists

… to lay audiences



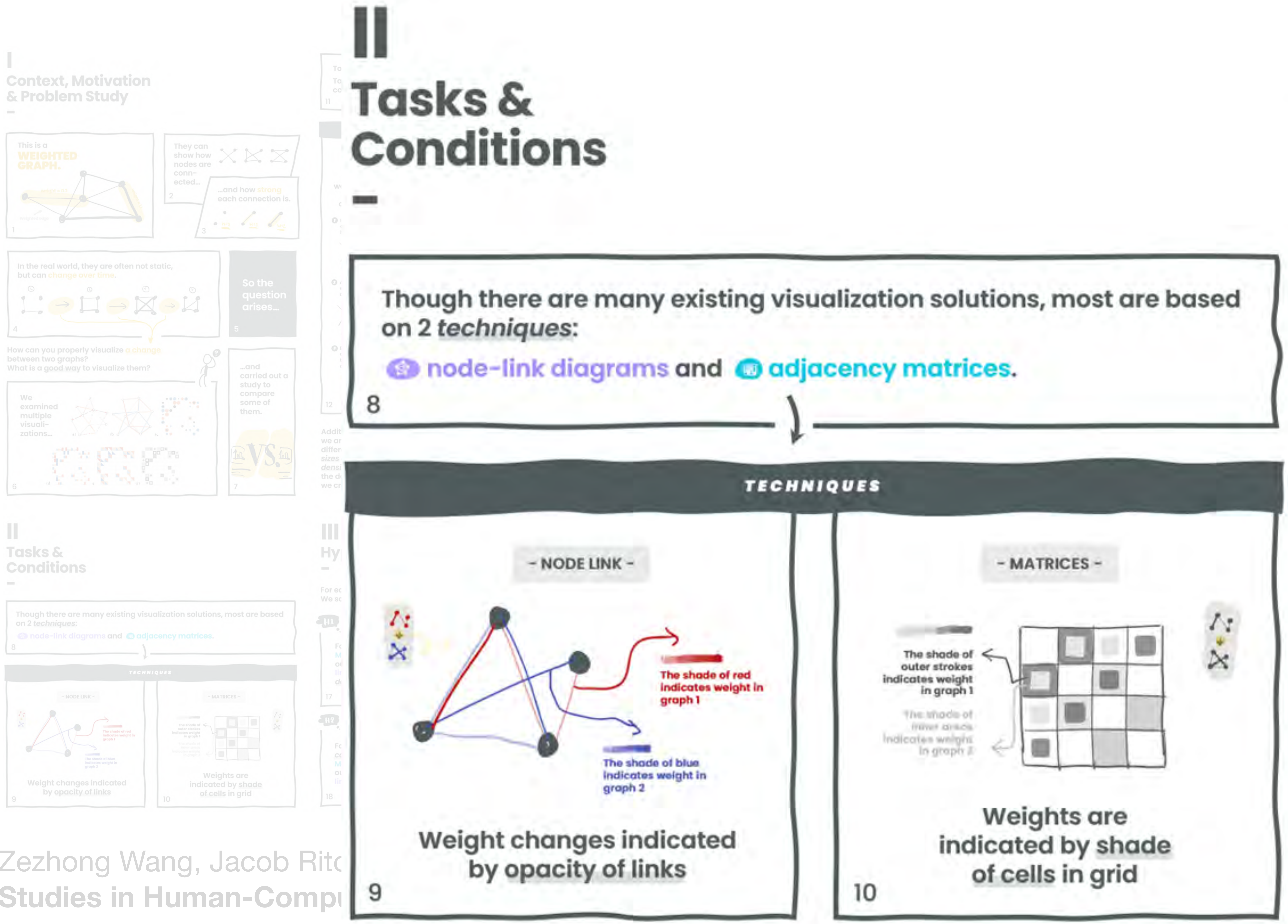
Zezhong Wang, Jacob Ritchie, Jingtao Zhou, Fanny Chevalier, Benjamin Bach. **Data Comics for Reporting Controlled User Studies in Human-Computer Interaction**. IEEE TVCG. 2021.

Scientists | … to other scientists | … to lay audiences



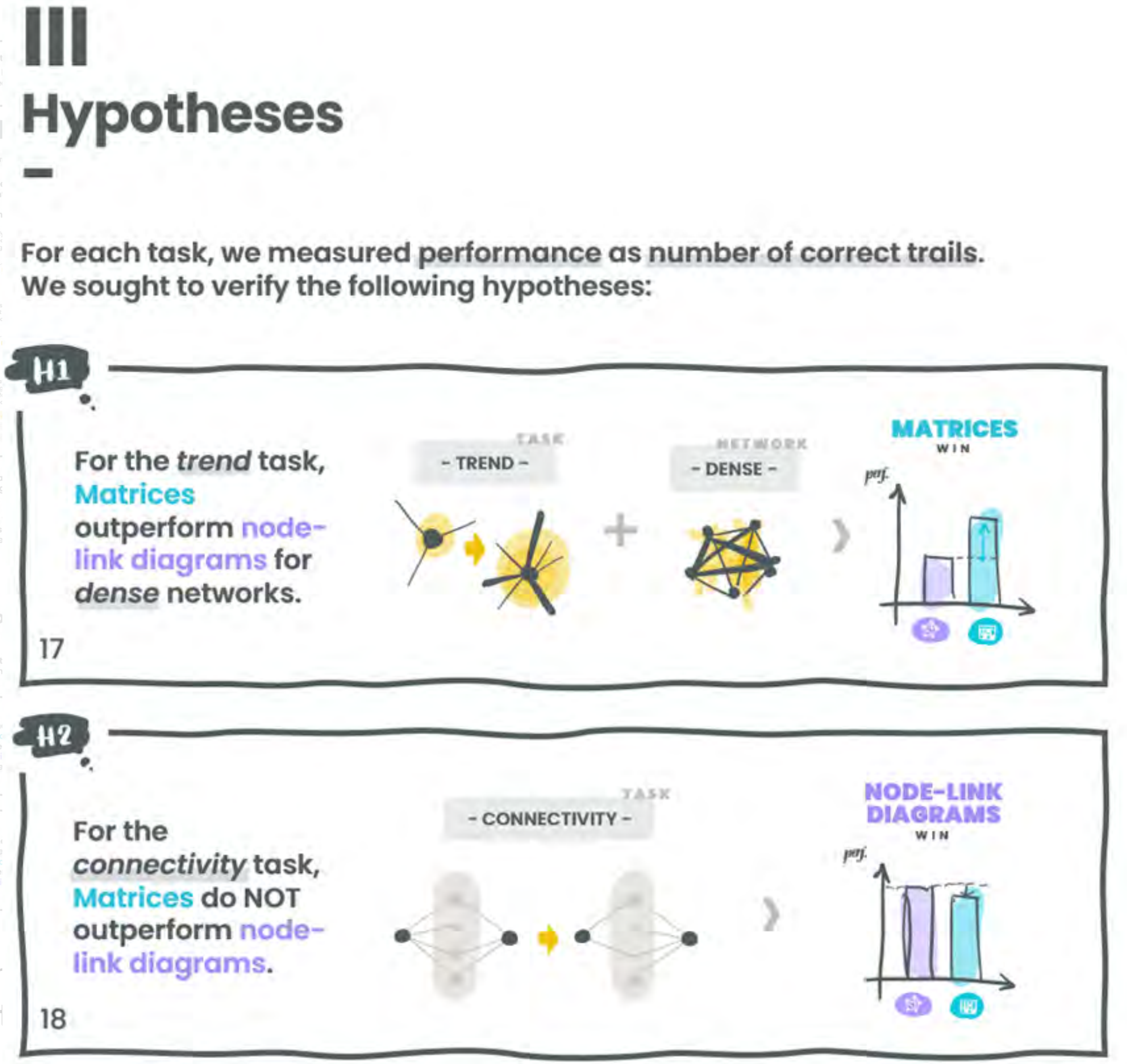
Zezhong Wang, Jacob Ritchie, Jingtao Zhou, Fanny Chevalier, Benjam
Studies in Human-Computer Interaction. IEEE TVCG. 2021.

Scientists | … to other scientists | … to lay audiences



Zezhong Wang, Jacob Rit… …porting Controlled User Studies in Human-Compu…

Scientists | … to other scientists | … to lay audiences

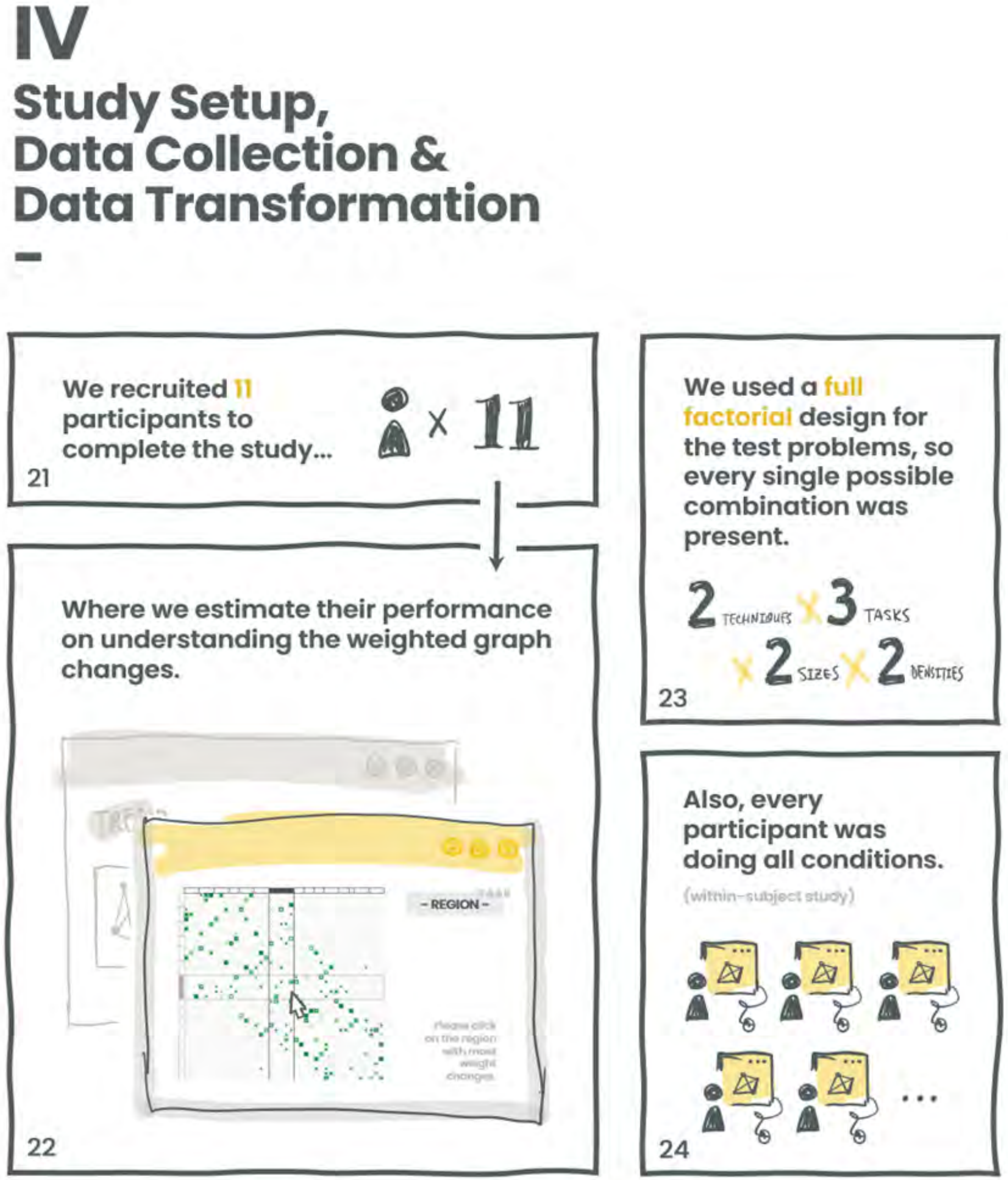


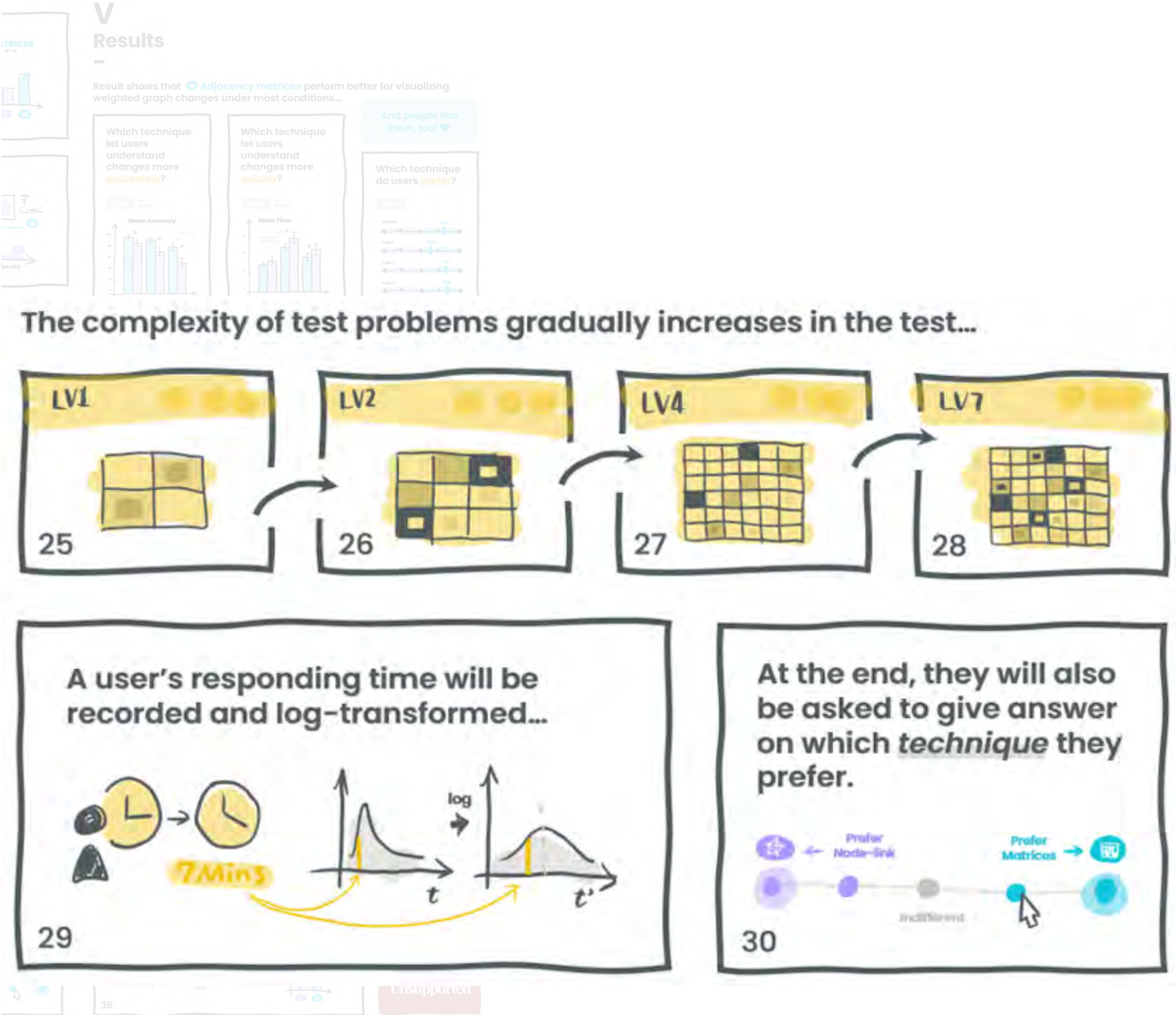
Zezhong Wang, Jacob Ritch … ting Controlled User
Studies in Human-Comput…

Scientists

… to other scientists

… to lay audiences





amin Bach. **Data Comics for Reporting Controlled User**

Scientists | … to other scientists | … to lay audiences

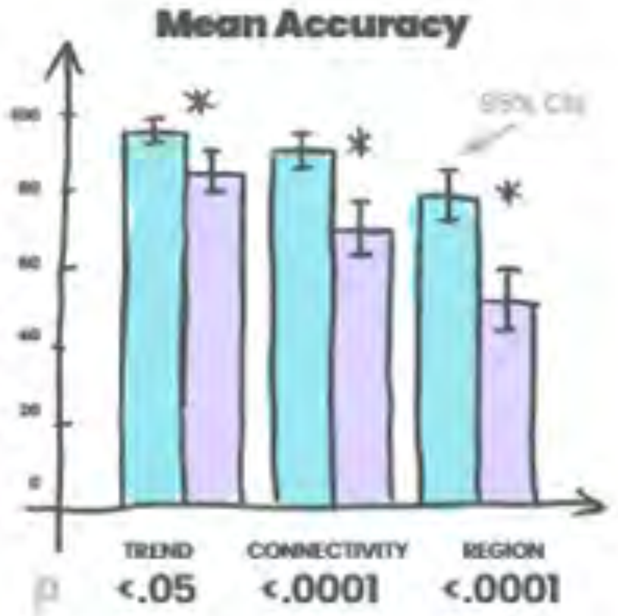
V

Results

—

Result shows that Adjacency matrices perform better for visualizing weighted graph changes under most conditions…

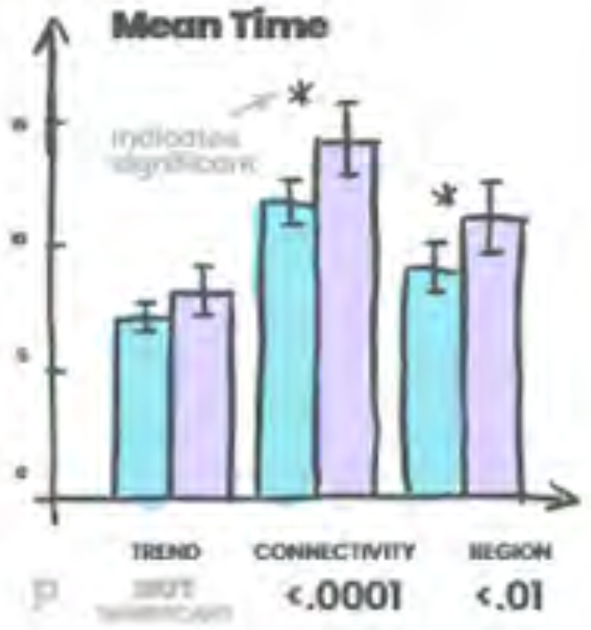
Which technique let users understand changes more **accurately**?



ADJACENCY MATRICES

31

Which technique let users understand changes more **quickly**?

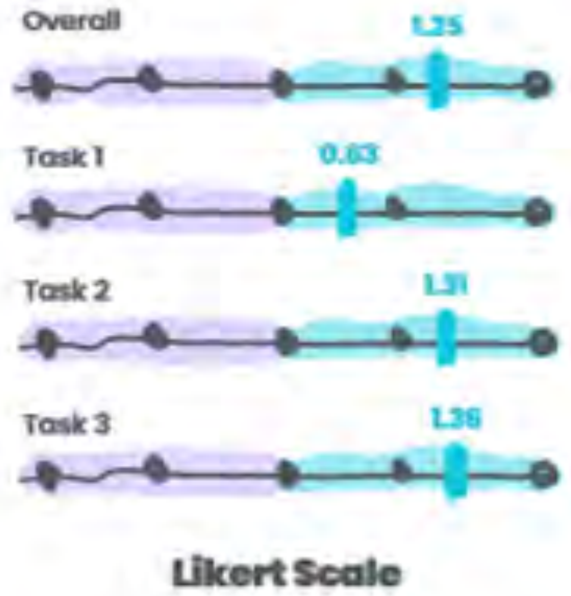


ADJACENCY MATRICES

32

And people like them, too! ♥

Which technique do users **prefer**?

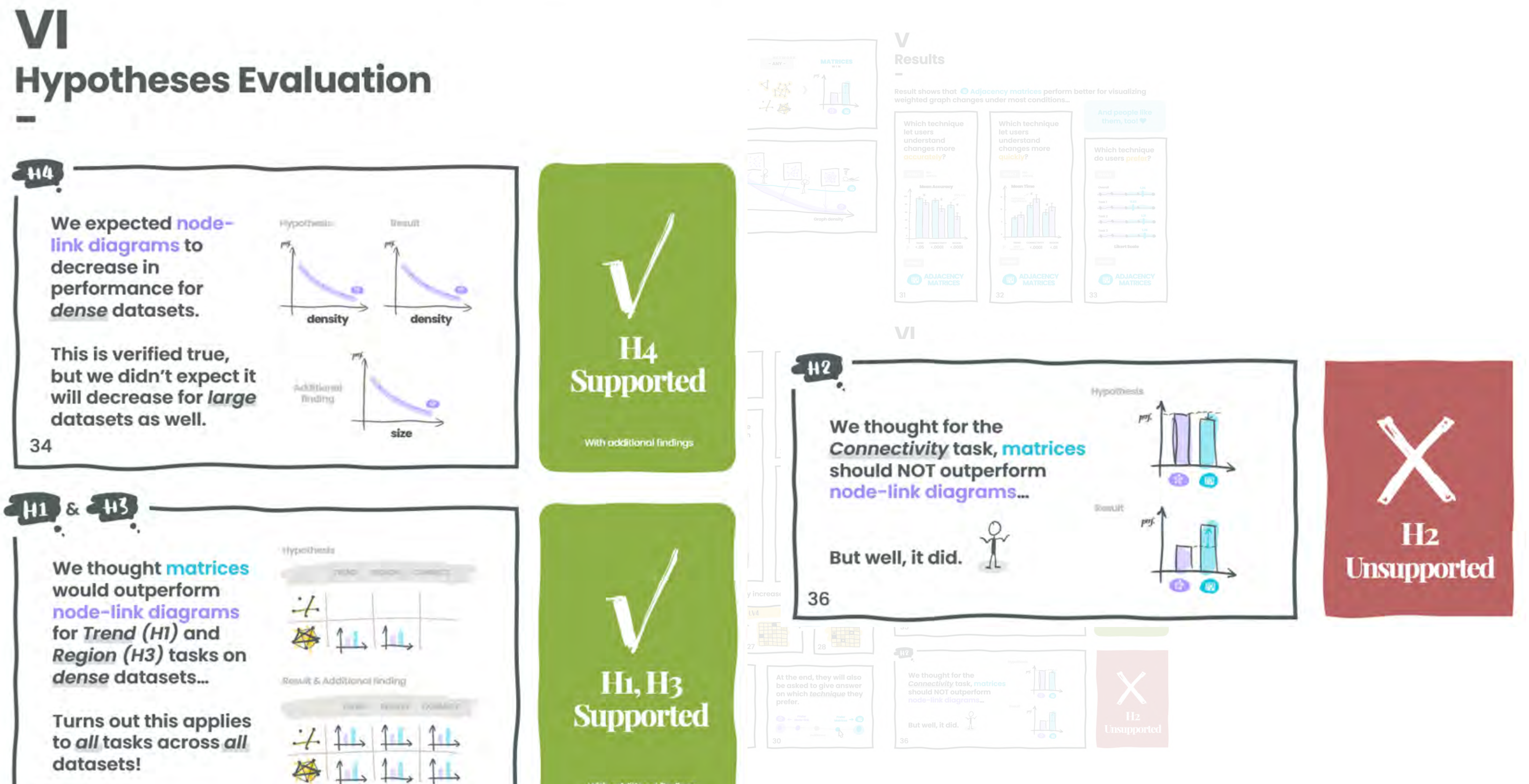


ADJACENCY MATRICES

33

Zezhong Wang, Jacob Rit… …orting Controlled User Studies in Human-Comp…

Scientists | … to other scientists | … to lay audiences

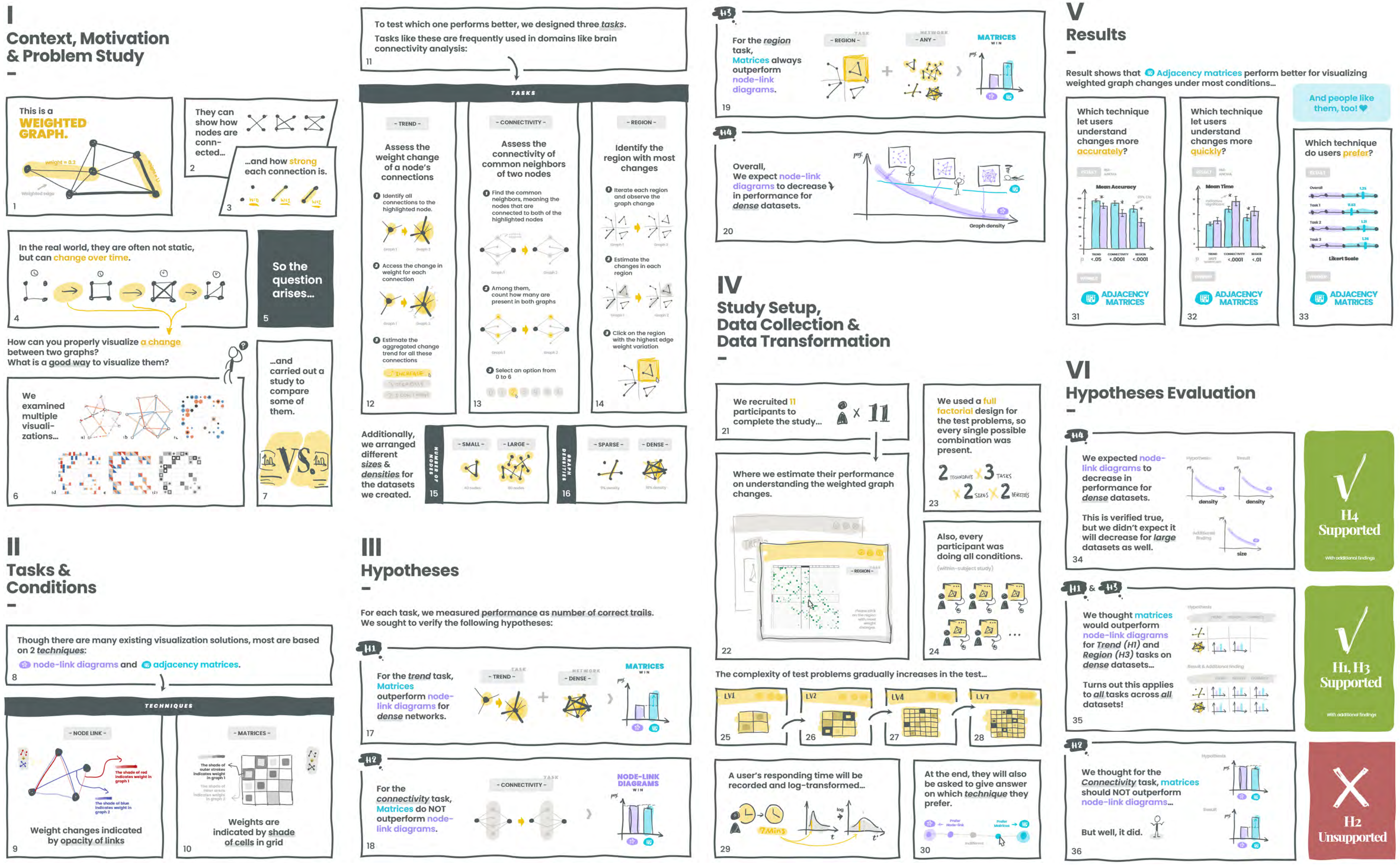


r, Benjamin Bach. **Data Comics for Reporting Controlled User**

Scientists

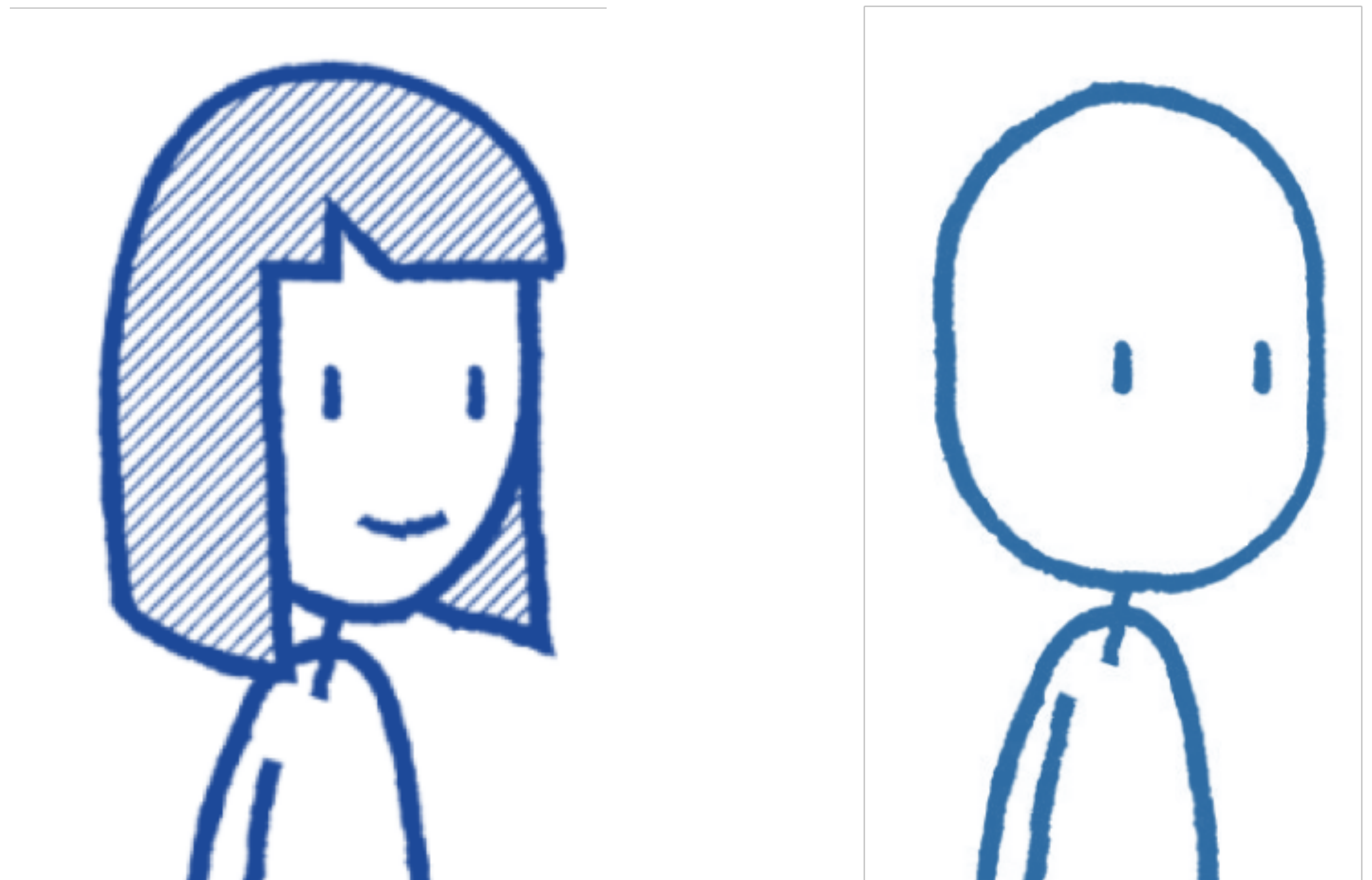
… to other scientists

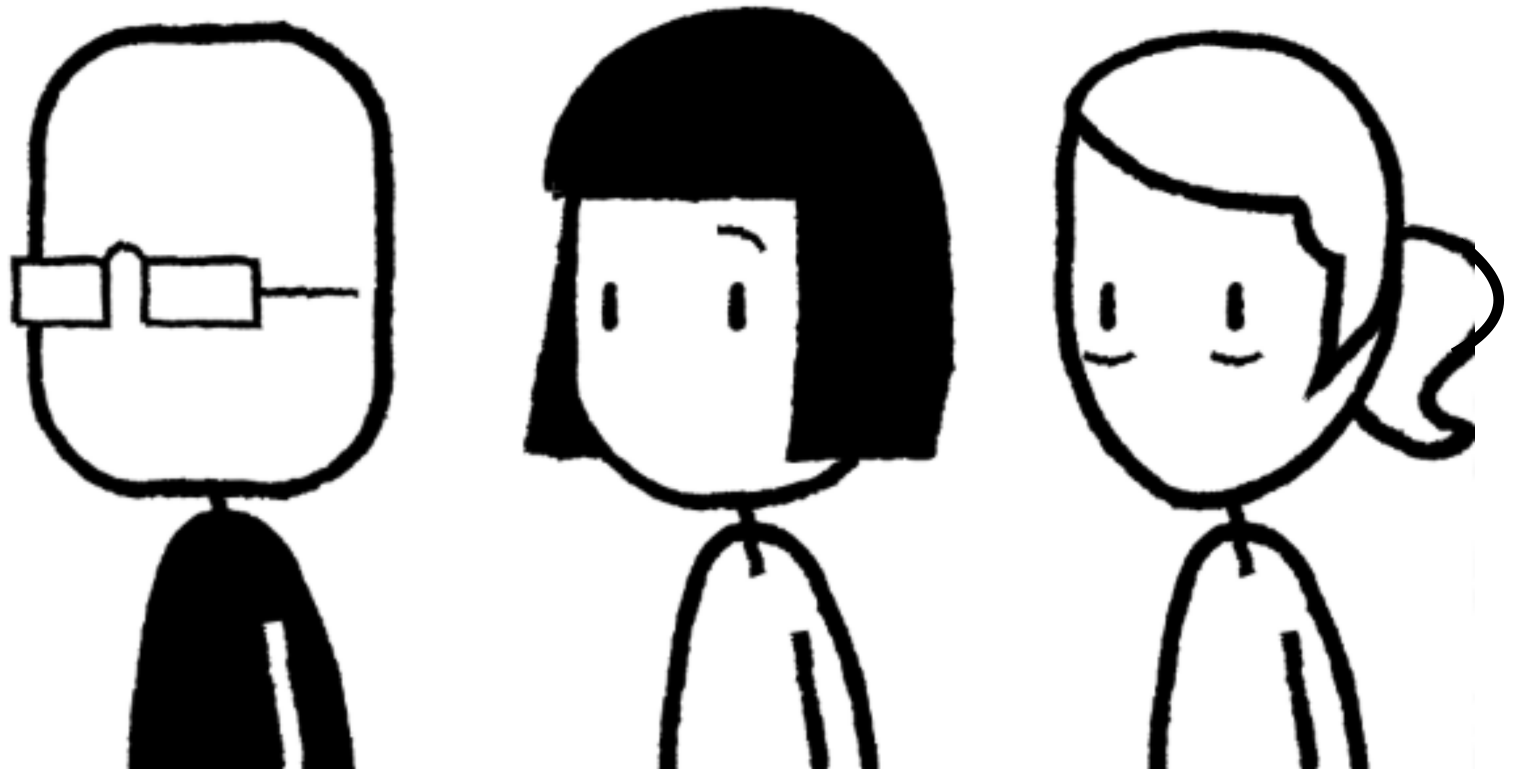
… to lay audiences



Zezhong Wang, Jacob Ritchie, Jingtao Zhou, Fanny Chevalier, Benjamin Bach. **Data Comics for Reporting Controlled User Studies in Human-Computer Interaction**. IEEE TVCG. 2021.

Scientists make sense of data
... communicate to other scientists
... and to lay audiences.

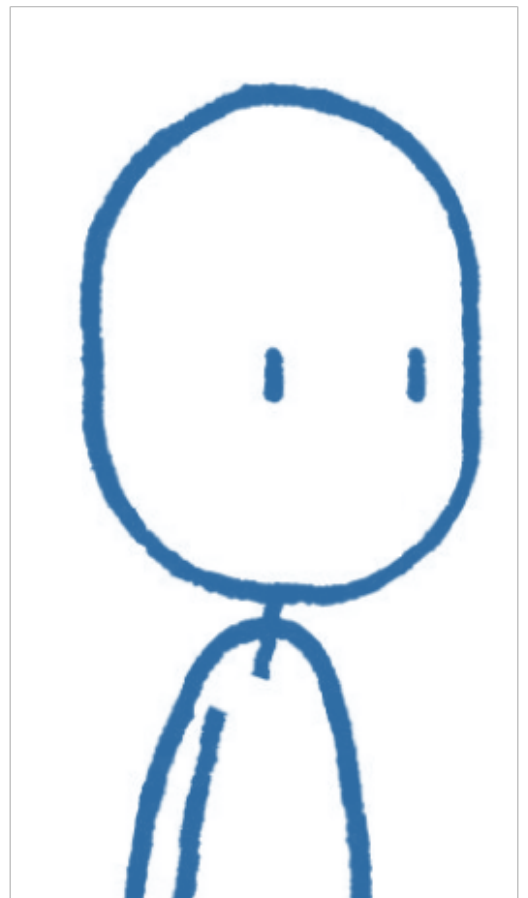


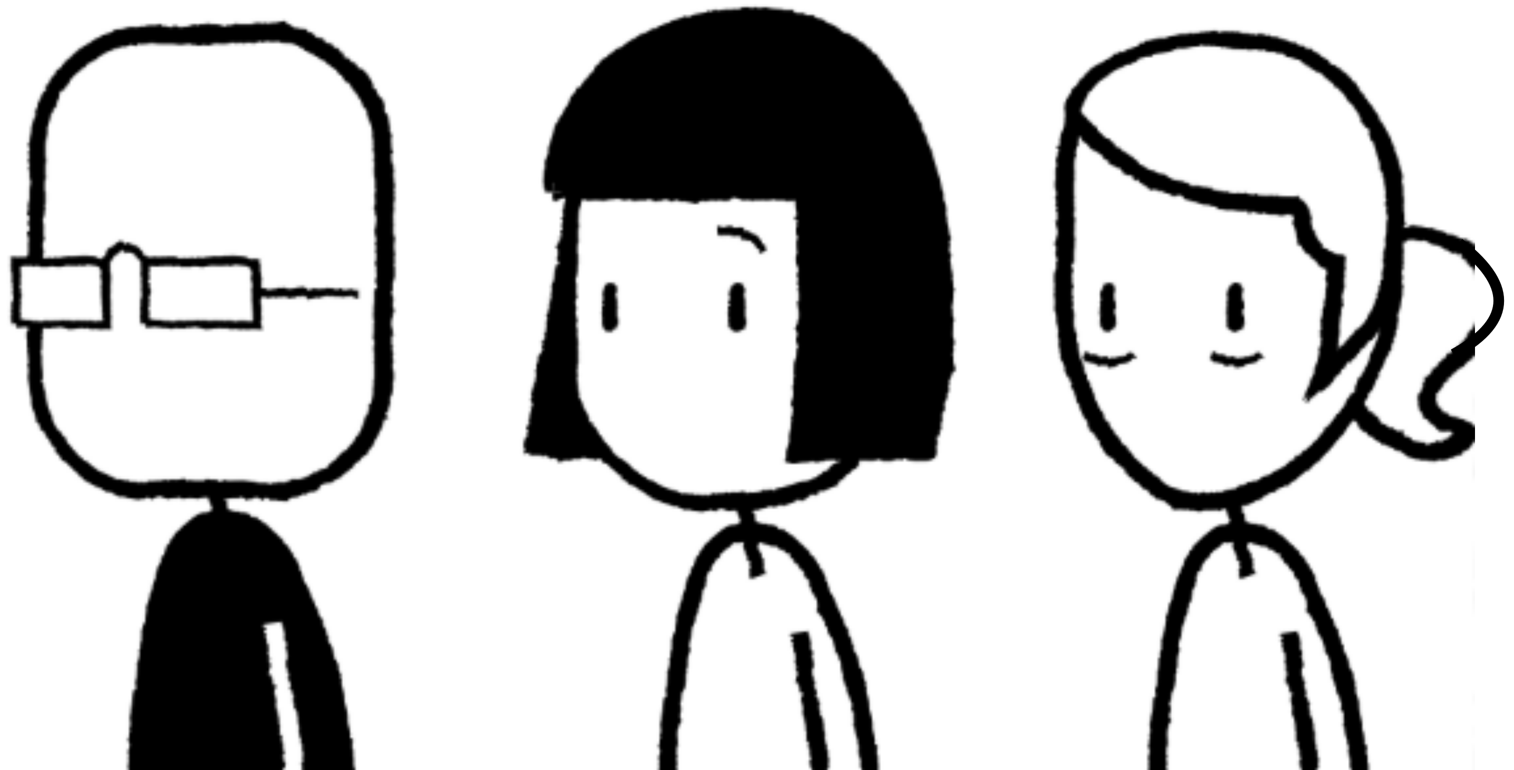


Visualizations tailored to researchers' goals & tasks

Transparent research practice

Innovative media for public understanding





Fanny CHEVALIER

http://fannychevalier.net