Check Your Bias at the Door

Disclaimer

- The following presentation and its findings are **not** affiliated with the following organisations:
  - ARTD Consultants
  - Iris Ethics
  - Autism CRC
  - Social Impact Measurement Network of Australia
  - Australian Evaluation Society
  - The Research Society
- The presenter has **not** received financial or other payment to commission the research and has delivered this as an independent activity

Warning

- (Noting that Kizzy Gandy last year presented evidence that these aren't as effective as people think)
- This is going to provoke some hard conversations and (pardon the conference pun) burst some bubbles
  - And I'm not just talking about my use of Bayesian methods
- If you're not comfortable with that, probably best to find another session

Argumentum ad auctoritatem

Only when the evaluator has insight into the interests and motivations of other actors in the system, ***into the roles that [they are] consciously or inadvertently playing****, the obstacles and opportunities that impinge upon the evaluative effort, and the limitations and possibilities for putting the results of evaluation to work—only with sensitivity to the politics of evaluation research—can the evaluator be as creative and strategically useful as [they] should be.*

- Carol Weiss, “Where Politics and Evaluation Research Meet”, 1973

Contents

- The rationale
- What is bias anyway?
- The survey
- The findings
- The implications

The rationale

Argumentum ad auctoritatem (#2)

Today's polarised political and social spheres are increasingly shaped and dominated by ideological forces that often paint starkly different realities about critical issues we face. The future of human societies, human autonomy and our modern democracies will be determined by the ideological battles on critical issues...

…views, beliefs and ideologies play a major role in how we characterise and understand these challenges, and how we formulate solutions to them.

- Mohsen Javdani & Ha-Joon Chang, "Who said or what said? Estimating ideological bias in views among economists", 2023

What political position(s) do we hold?

Social science researchers tend to express left-liberal political attitudes. The ideological skew might influence research evaluations, but ***empirical evidence is limited****.*

- Henning Finseraas et al., “Ideological biases in research evaluations? The case of research on majority–minority relations”, 2022

Argumentum ad auctoritatem? (#3)

In God we trust; all others must use data.

- Mary Walton

Our world is politicised, populist, and polarised

…and Australia is not immune.

The Guardian Aus

Australian Bureau of Statistics to slash jobs despite census bungle

The ABS has had its budget cut under several governments. The ALP, under Kevin Rudd, cut its budget in 2008, resulting in the ABS halting its job vacancy survey just as the global financial crisis was about to hit the Australian economy.

The Abbott government further cut the ABS budget in 2014 and the Turnbull government continued the ongoing "efficiency dividends", which have been eroding the organisation's budget.

Donnelly said: "We are seriously concerned that the ABS is continually being forced to make decisions solely based on its meagre budget. The bureau needs adequate funding to provide robust, accurate data and that's clearly not currently the case.

Argumentum ad auctoritatem? (#4)

The operating assumption of social scientists, like that of journalists and gossips, is that exposure is good and more exposure, better; knowledge is always good and more knowledge, better: never tasteless, purposeless, or destructive.

.

- Harold Orlans, “Neutrality and Advocacy in Policy Research”, 1975

The rationale

Knowledge is the best defence…

…and transparency is an unexpectedly good offence.

What is bias anyway?

What is bias anyway?

Predicting political beliefs with polygenic scores for cognitive performance and educational attainment

Tobias Edwards, Alexandros Giannelis, Emily A. Willoughby, James J. Lee

Show more

+ Add to Mendeley Share Cite

https://doi.org/10.1016/j.intell.2024.101831 Get rights and content

Highlights

- Within-families, intelligence predict left-wing beliefs.
- DNA-based predictors of IQ also predict political beliefs within families.
- Our results imply that being genetically predisposed to be smarter causes left-wing beliefs.

Abstract

Intelligence is correlated with a range of left-wing and liberal political beliefs. This may suggest intelligence directly alters our political views. Alternatively, the association may be confounded or mediated by socioeconomic and environmental factors. We studied the effect of intelligence within a sample of over 300 biological and adoptive families, using both measured IQ and polygenic scores for cognitive performance and educational attainment. We found both IQ and polygenic scores significantly predicted all six of our political scales. Polygenic scores predicted social liberalism and lower authoritarianism, within-families. Intelligence was able to significantly predict social liberalism and lower authoritarianism, within families, even after controlling for socioeconomic variables. Our findings may provide the strongest causal inference to date of intelligence directly affecting political beliefs.

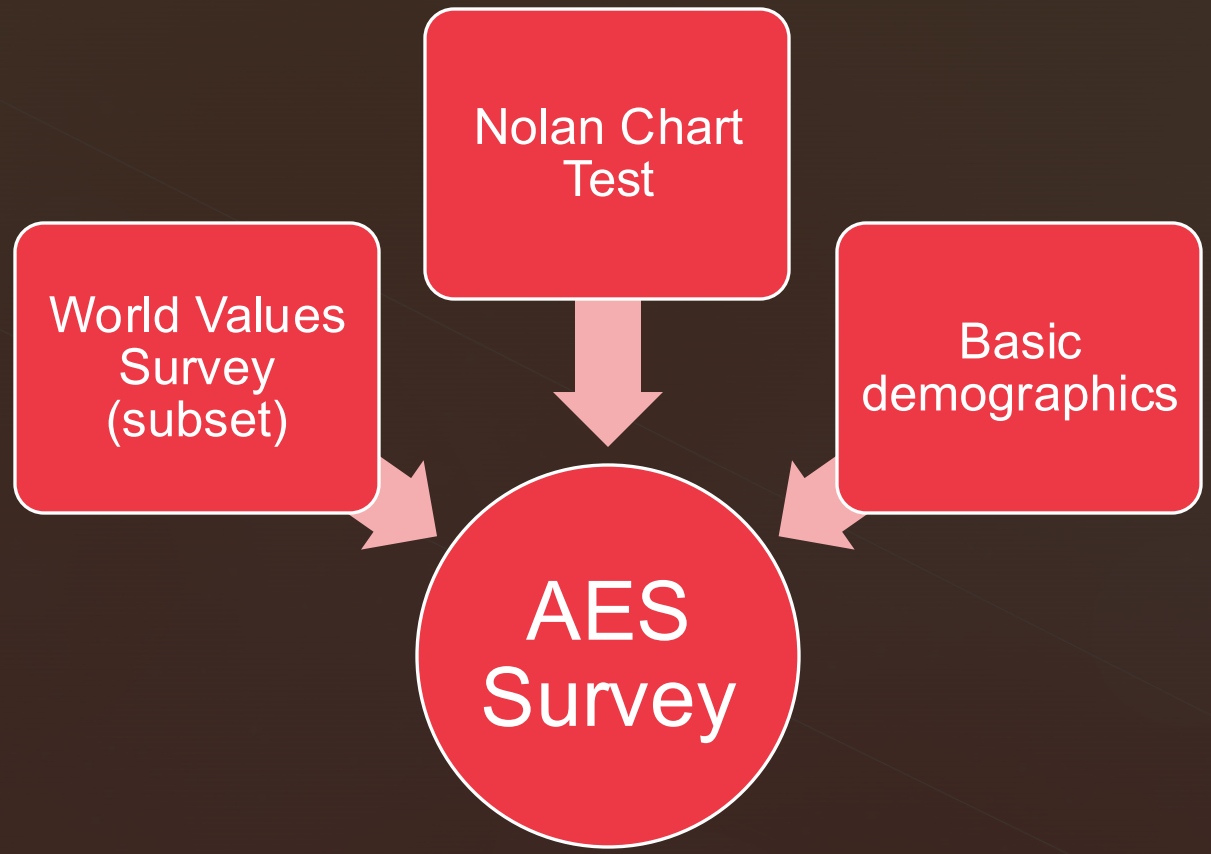
Confounding factors

The survey

The survey

- Online survey
- Advertised through LinkedIn
- ~7.5 minutes in length

The survey


Nolan Chart Test
World Values Survey (subset)
Basic demographics
AES Survey

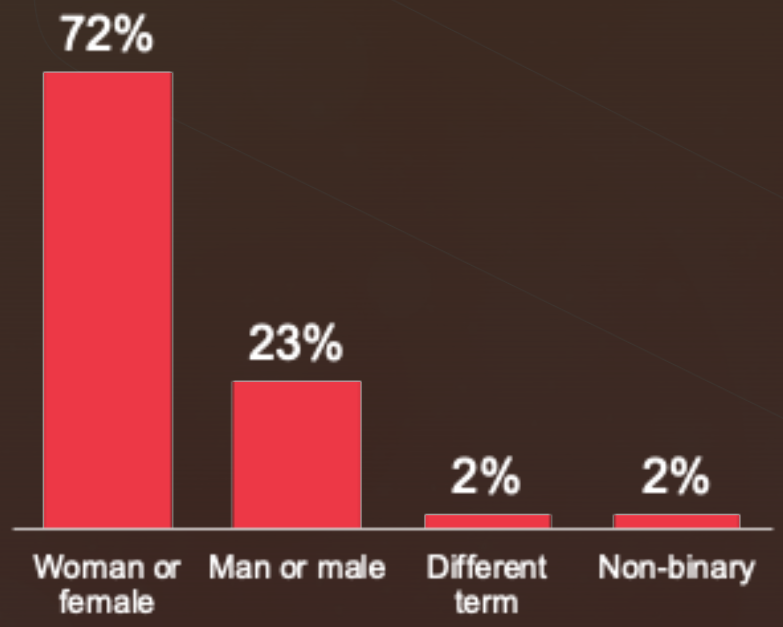
Double consent process

The survey (limitations)

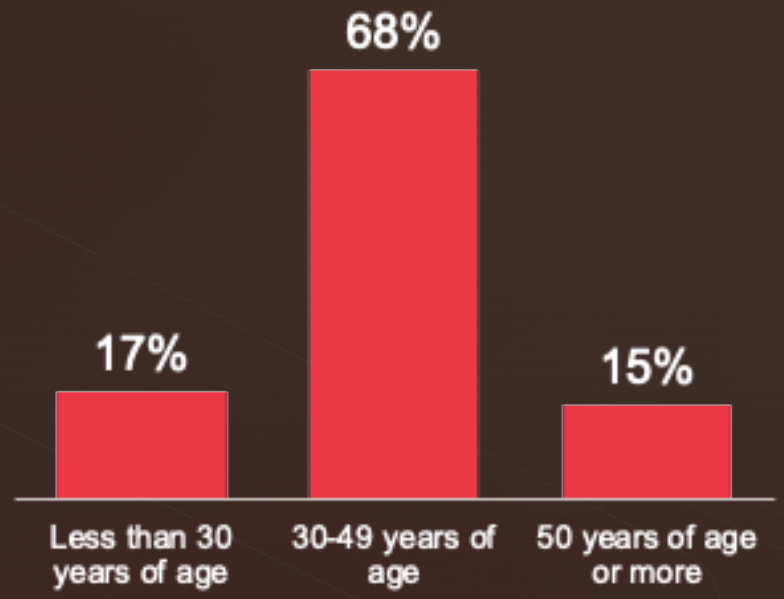
- Small sample (n = 47)
- LinkedIn only
- Anonymous
- Demographic skew
- Instrument limitations

Survey respondents

Gender



Age

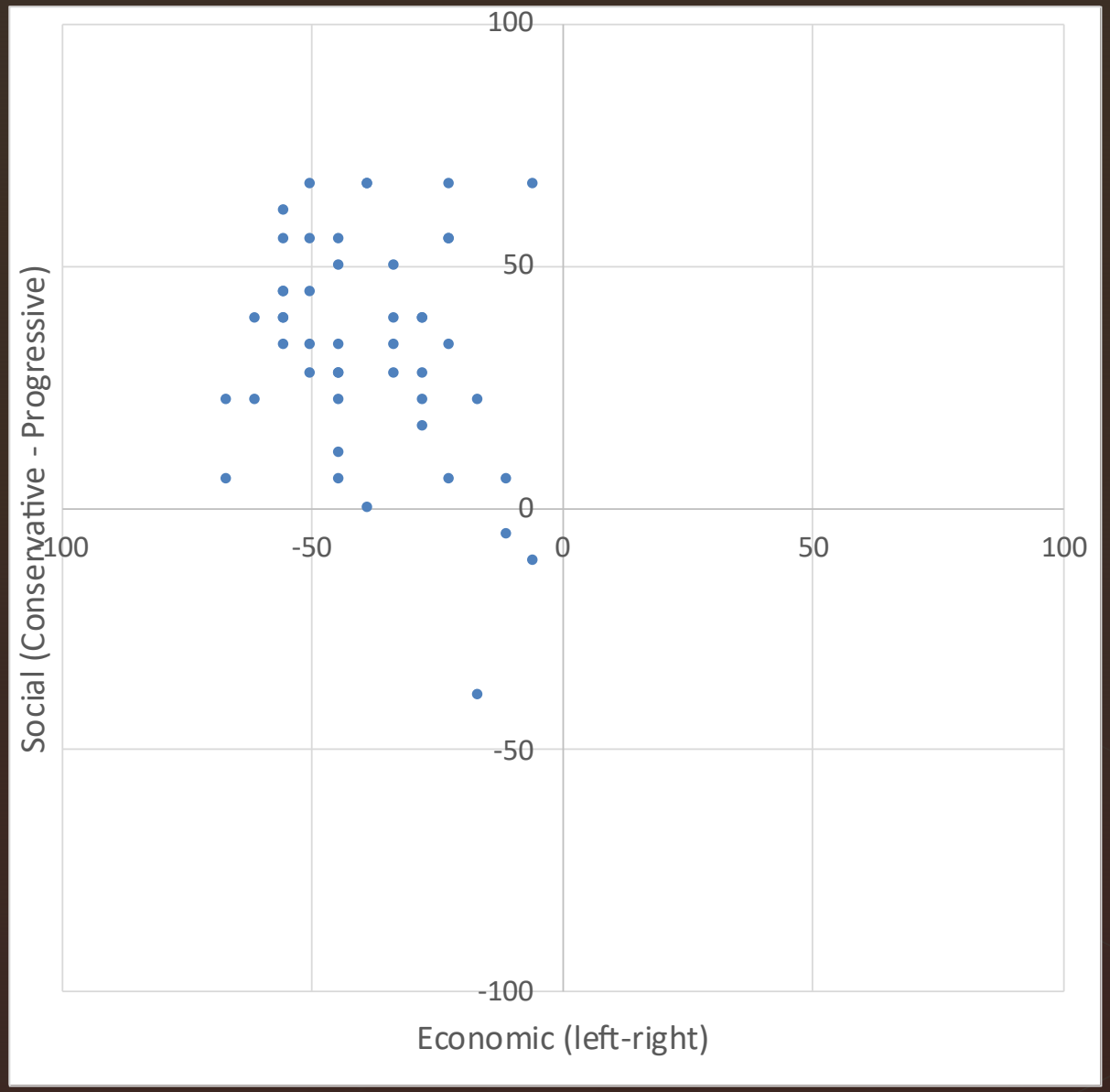


Reweighting

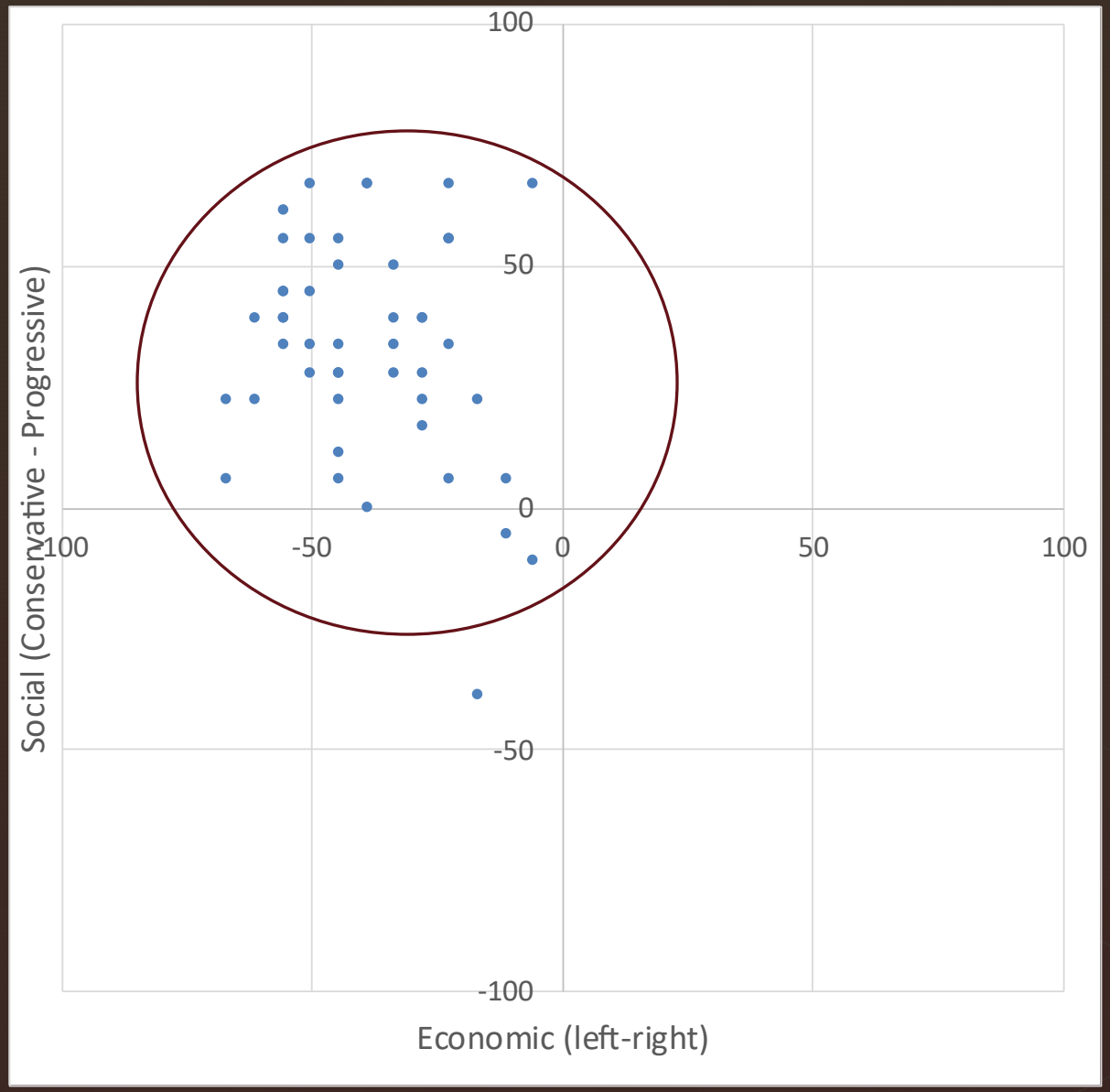
The findings

The findings – Nolan Chart Test

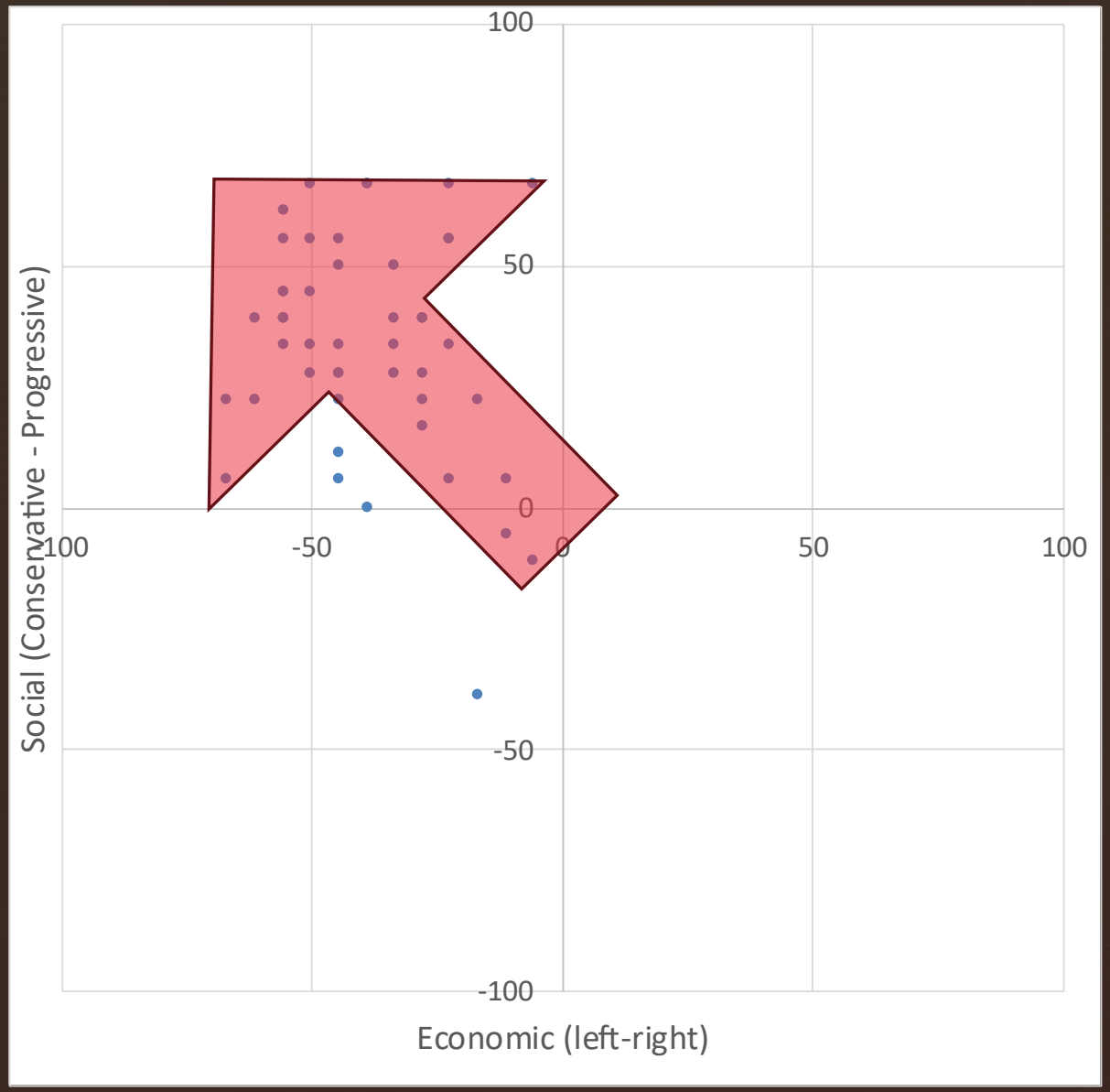
Nolan Test


100
50
0
-50
-100
-100
-50
0
50
100
Social (Conservative - Progressive)
Economic (left-right)

Nolan Test

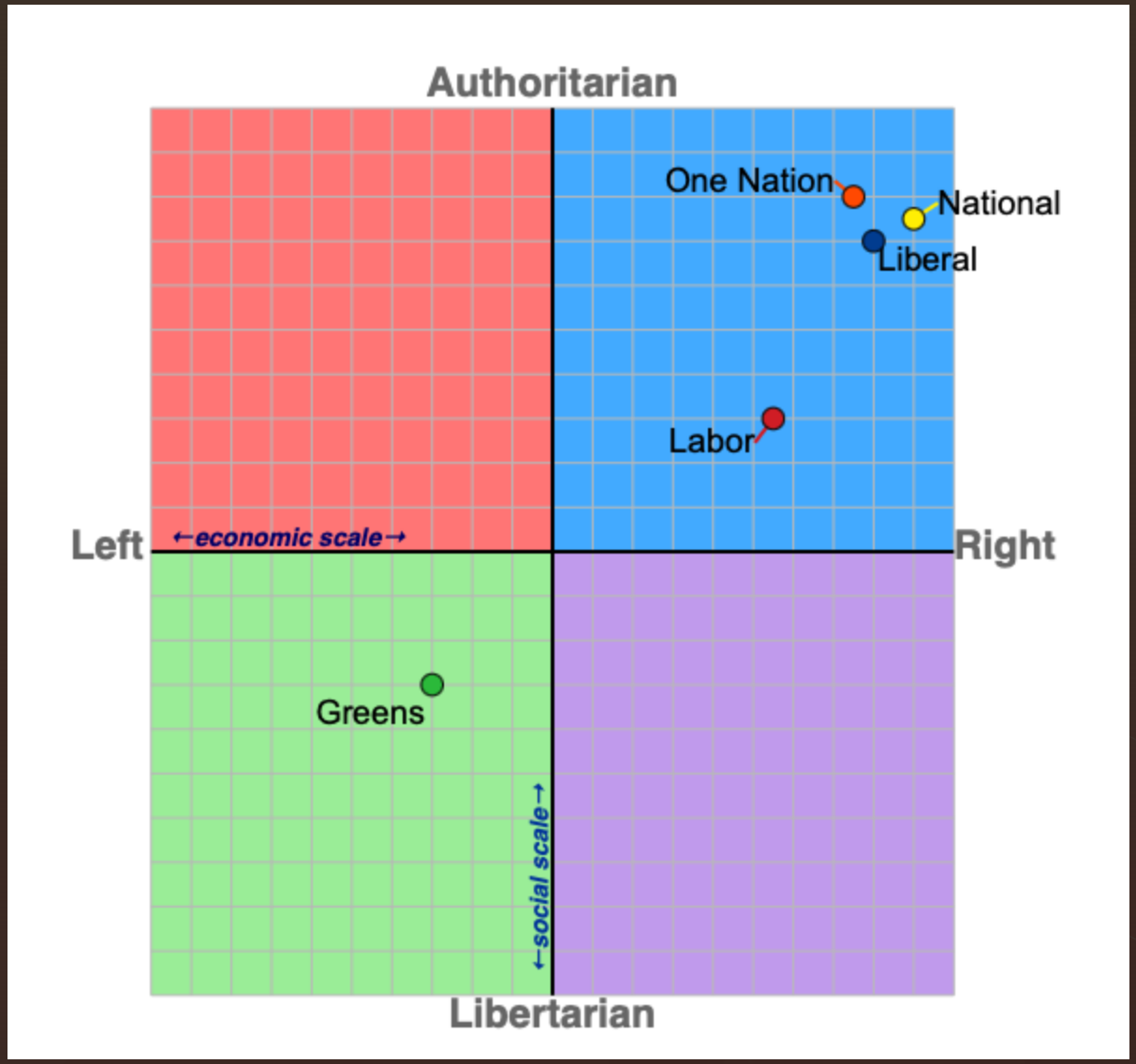

100
50
0
-50
-100
-100
-50
0
50
100
Social (Conservative - Progressive)
Economic (left-right)

Overfitting



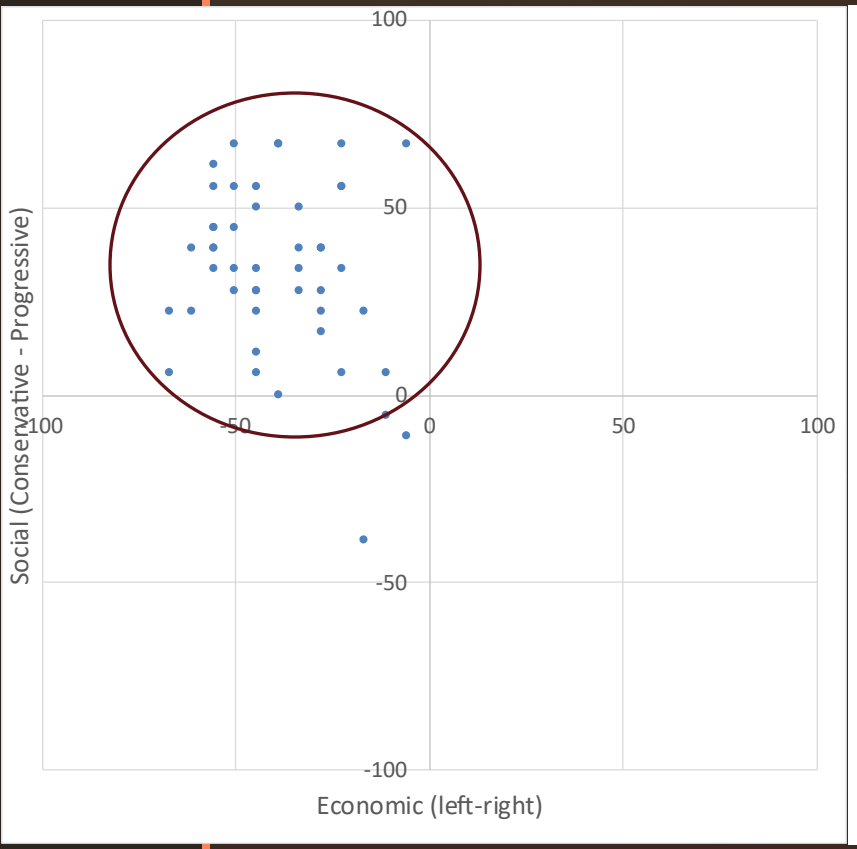
Vote Compass

Political Compass

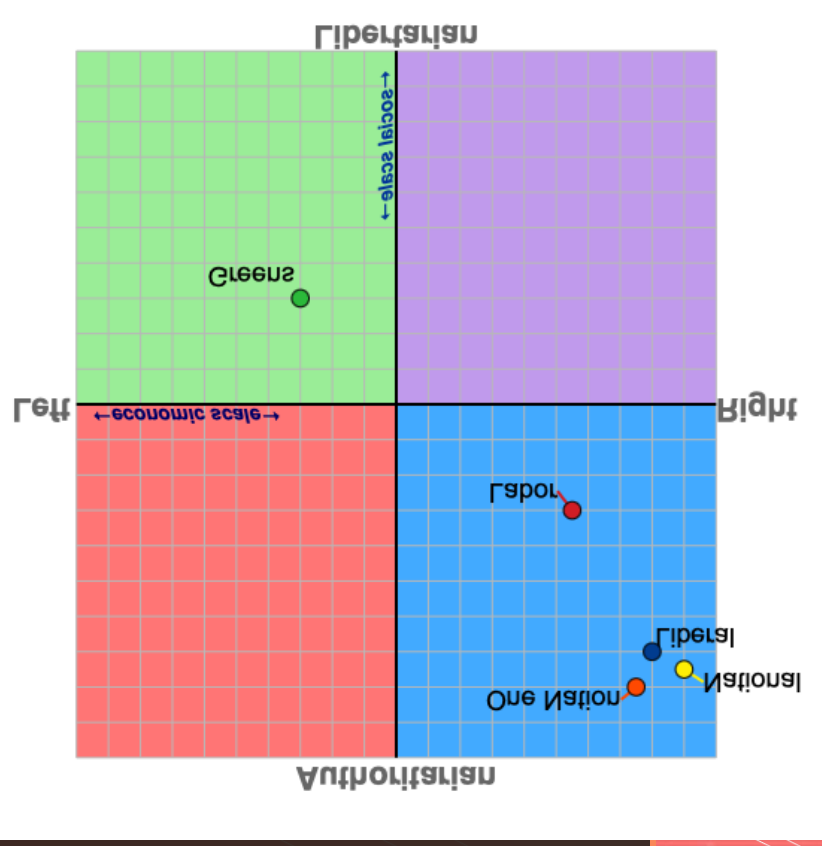


Political Compass

Side by side


100
50
0
-50
-100
-100
-50
0
50
100
Social (Conservative - Progressive)
Economic (left-right)


Progressive →
GRN
YOU
ALP
LNP
ONP
SOCIAL
← Conservative
← Left
ECONOMIC
Right →


Libertarian
←social issues→
Greens
Left
←economic scale→
Right
Labor
Liberal
National
One Nation
Authoritarian

The findings – WVS data

WVS Module subsets:	
	Political participation
	Social values
	Economic values

Comparing distributions

- Most of the scales in the WVS are:
  - Interval (0-10 scales)
  - Ordinal (Likert & Word Scale)
- For comparison, “mean” values are used, with ordinal scales converted to interval.*
- Two-tailed t-test applied, significance at $p < 0.05$
- This is a statistical shortcut, taken for two reasons…

**The more sophisticated Wilcoxon rank-sum test also uses a mean, it’s just better hidden in the mathematics.*

Comparing distributions

- Reason 1: The pattern is what's important
- Reason 2: Bayesian analysis is better

How I snuck a presentation on applied Bayesian statistics into #AES25

Meaningful Comparisons With Ordinal-Scale Items

Martin Schnuerch[1] [a], Julia M. Haaf[2], Alexandra Sarafoglou[2], Jeffrey N. Rouder[3]

[1] Department of Psychology, University of Mannheim, Mannheim, Germany, [2] University of Amsterdam, Amsterdam, Netherlands, [3] University of California, Irvine, CA, US; University of Mannheim, Mannheim, Germany

Keywords: stochastic dominance, bayes factors, ordinal scales, meaningful comparisons, Likert items

https://doi.org/10.1525/collabra.38594

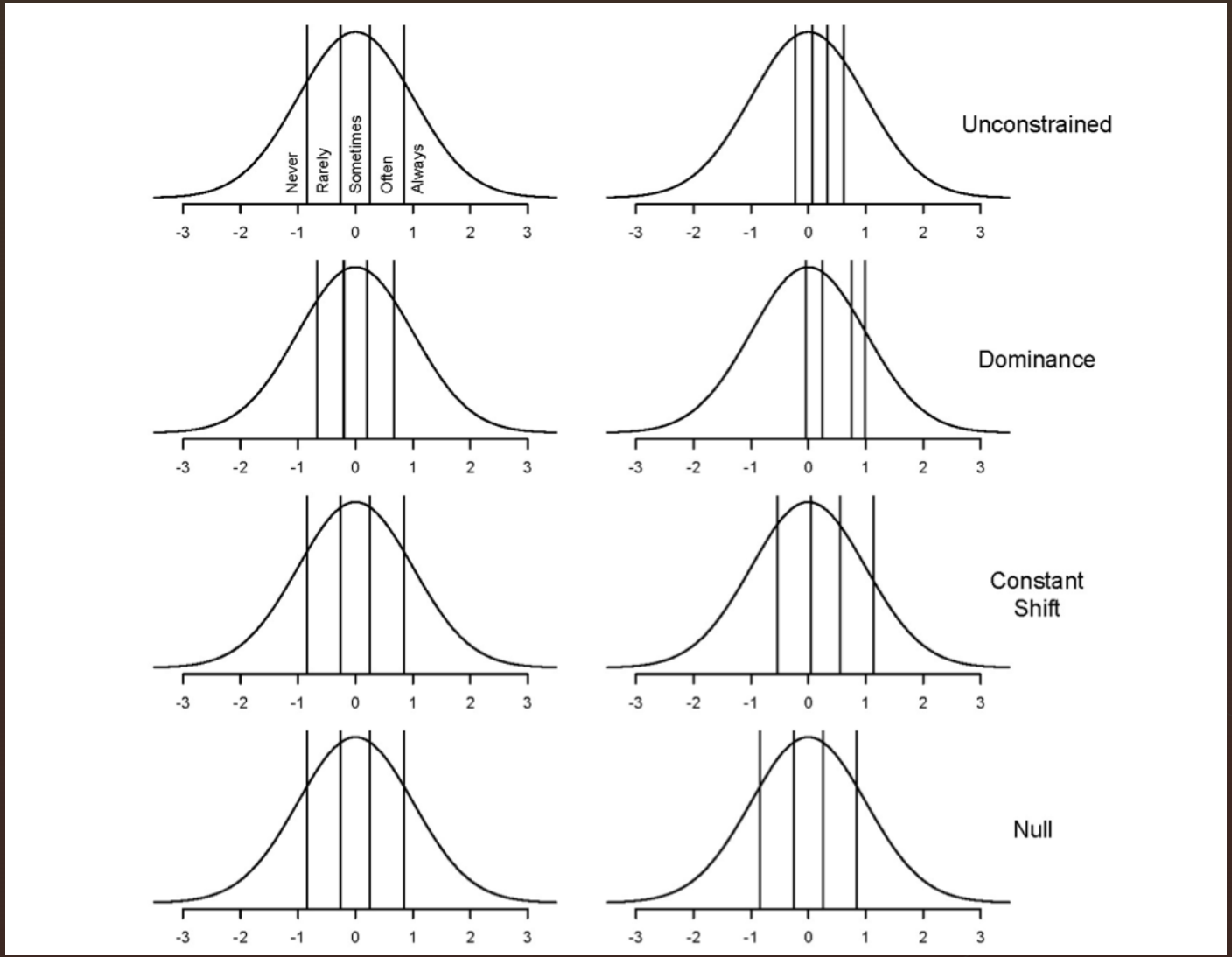
Collabra: Psychology

Vol. 8, Issue 1, 2022

How it works

- Instead of just testing for a difference, it tests four hypotheses for the nature of the difference:
  - Unconstrained (difference with no systematic pattern)
  - Dominance (systematic, with no overlap in category populations)
  - Constant shift (systematic and constant difference)
  - Null (no difference)

It’s easier to explain with a picture

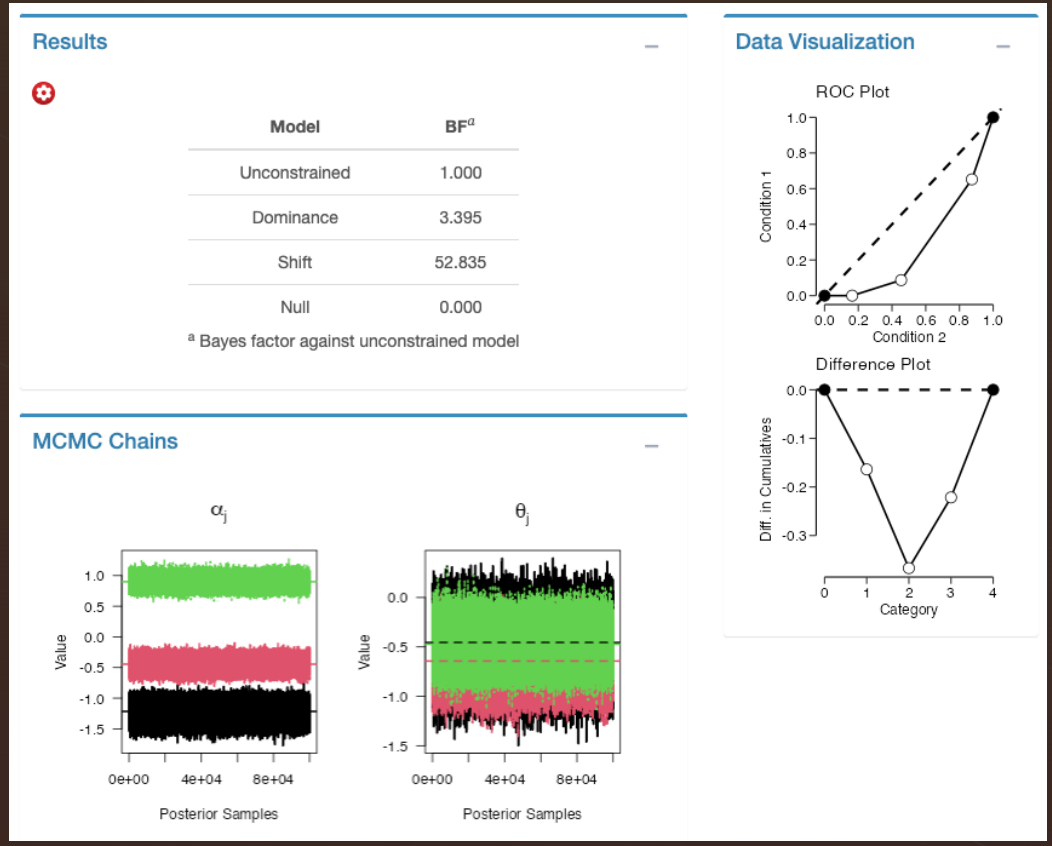


The technical details

- Markov Chain Monte Carlo simulation
- Outputs in terms of Bayes factors => calculate Posterior Odds Ratios (PORs)
- PORs are presented relative to unconstrained model for ease of interpretation
- Posterior-Prior Odds Ratios are functionally equivalent to PORs due to use of naïve priors (prior odds are equal across all models)
- POR > 10 is treated as “strong” evidence*

**R. Kass & A. Raftery (1995). "Bayes Factors". Journal of the American Statistical Association. 90 (430): 791*

There’s an app for that

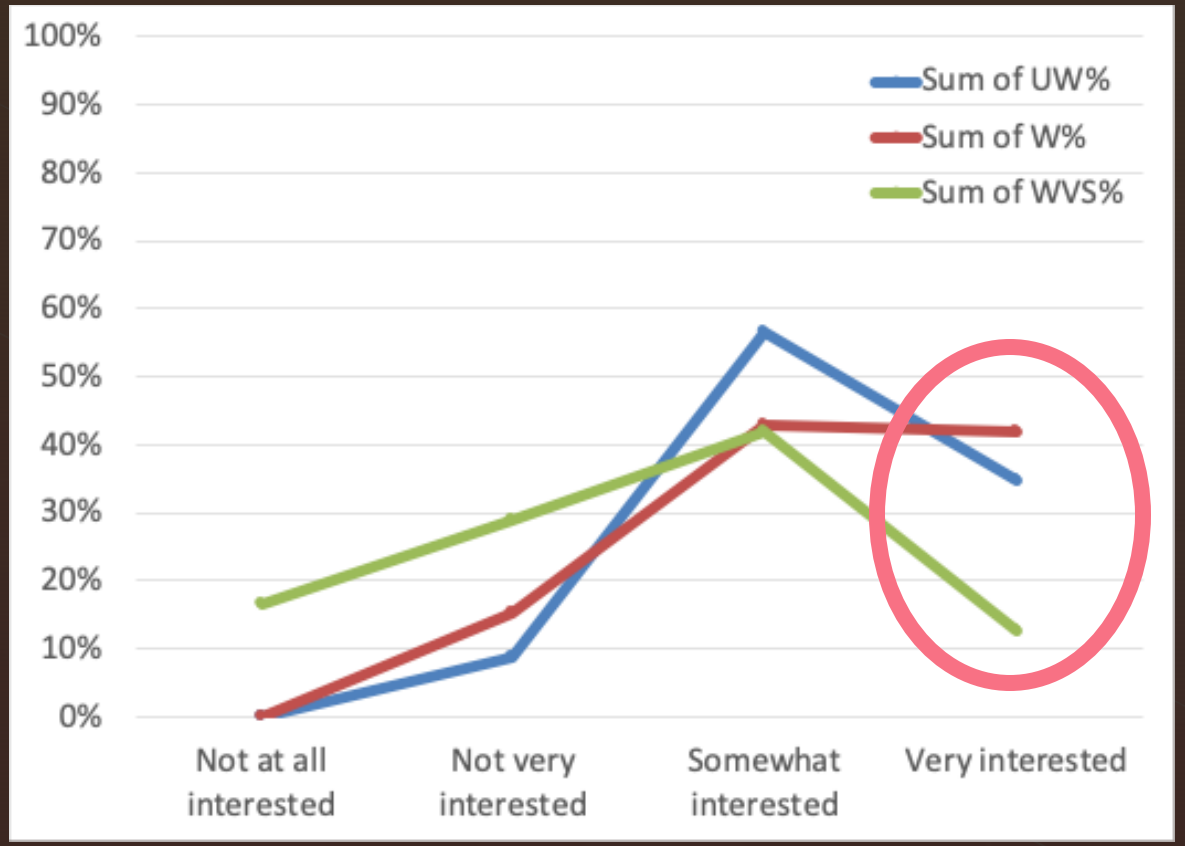


Political Participation

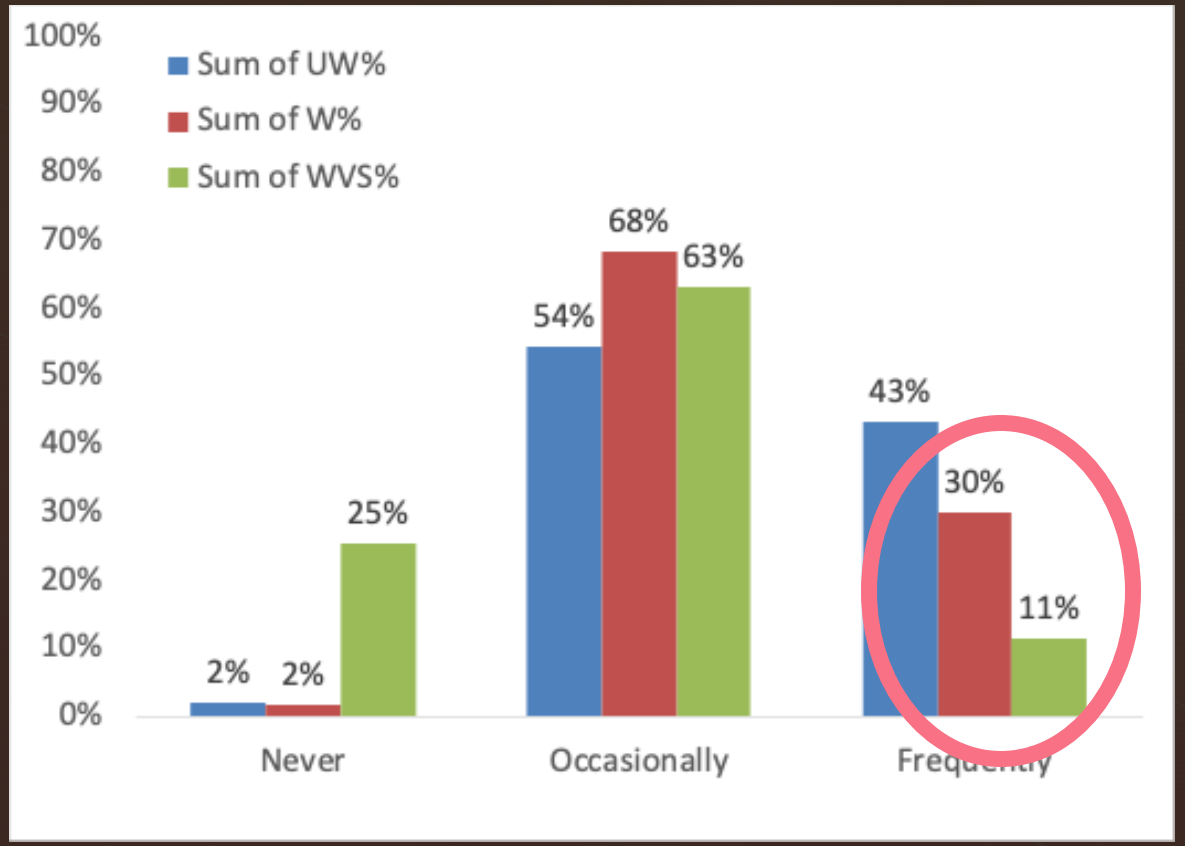
Overview

Question	P-UW	P-W	POR-UW	POR-W
How interested would you say you are in politics?	**0.00**	**0.00**	**Shift – 53**	**Shift – 107**
When you get together with your friends, would you say you discuss political matters frequently, occasionally or never?	**0.00**	**0.00**	**Shift – 19**	Shift – 5
I'm going to describe various types of political systems and ask what you think about each as a way of governing this country?				
—Having a democratic political system	0.08	**0.00**	Dom – 1.5	Shift – 7
—Having a strong leader who does not have to bother with parliament and elections	**0.01**	**0.00**	**Shift – 24**	**Shift – 44**
—Having a system governed by religious law in which there are no political parties or elections	**0.00**	**0.00**	Dom – 1.1	Dom – 1.1
—Having experts, not government, make decisions according to what they think is best for the country	0.14	0.41	Null – 2.6	Dom – 1.8
—Having the army rule	**0.00**	**0.00**	Dom – 1.2	Dom – 1.1

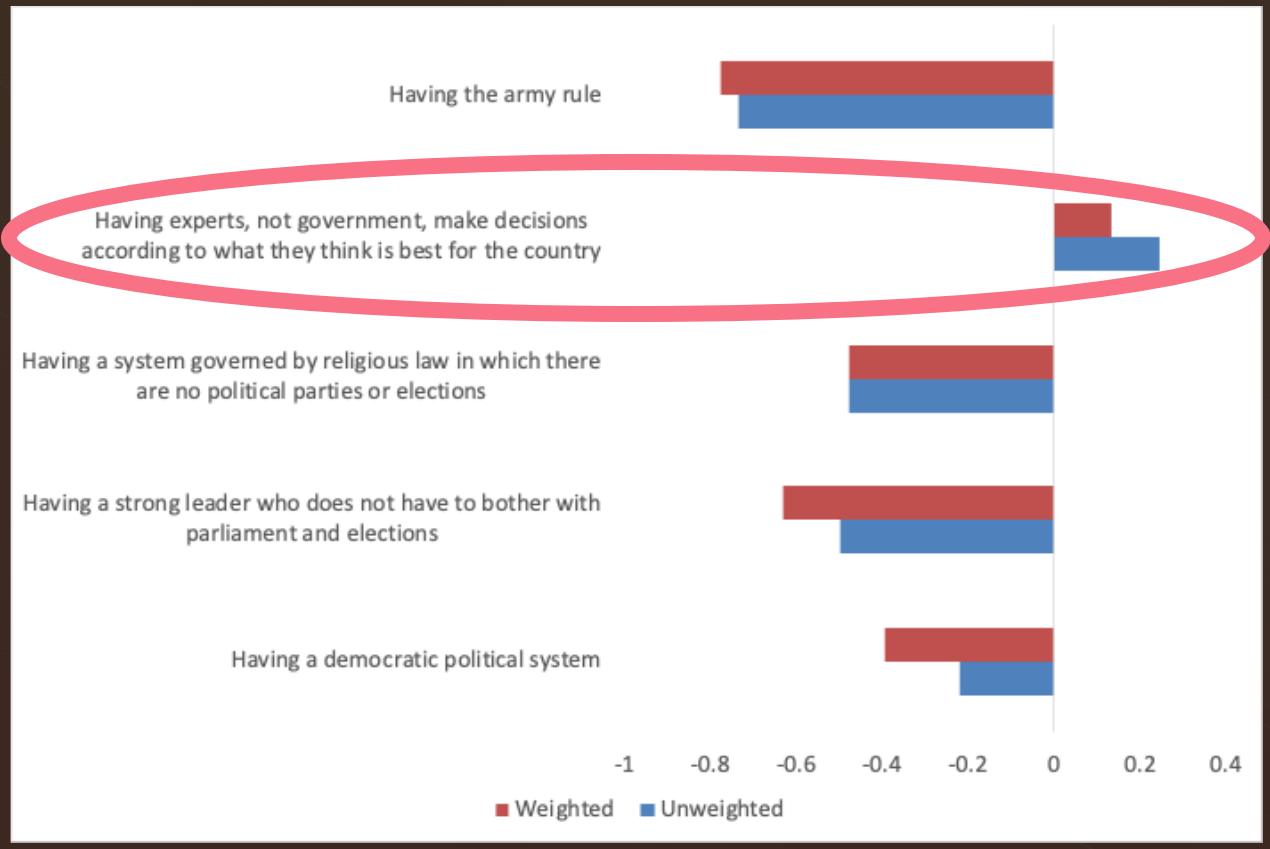
Evaluators are more interested in politics



And evaluators like to talk about it with family and friends



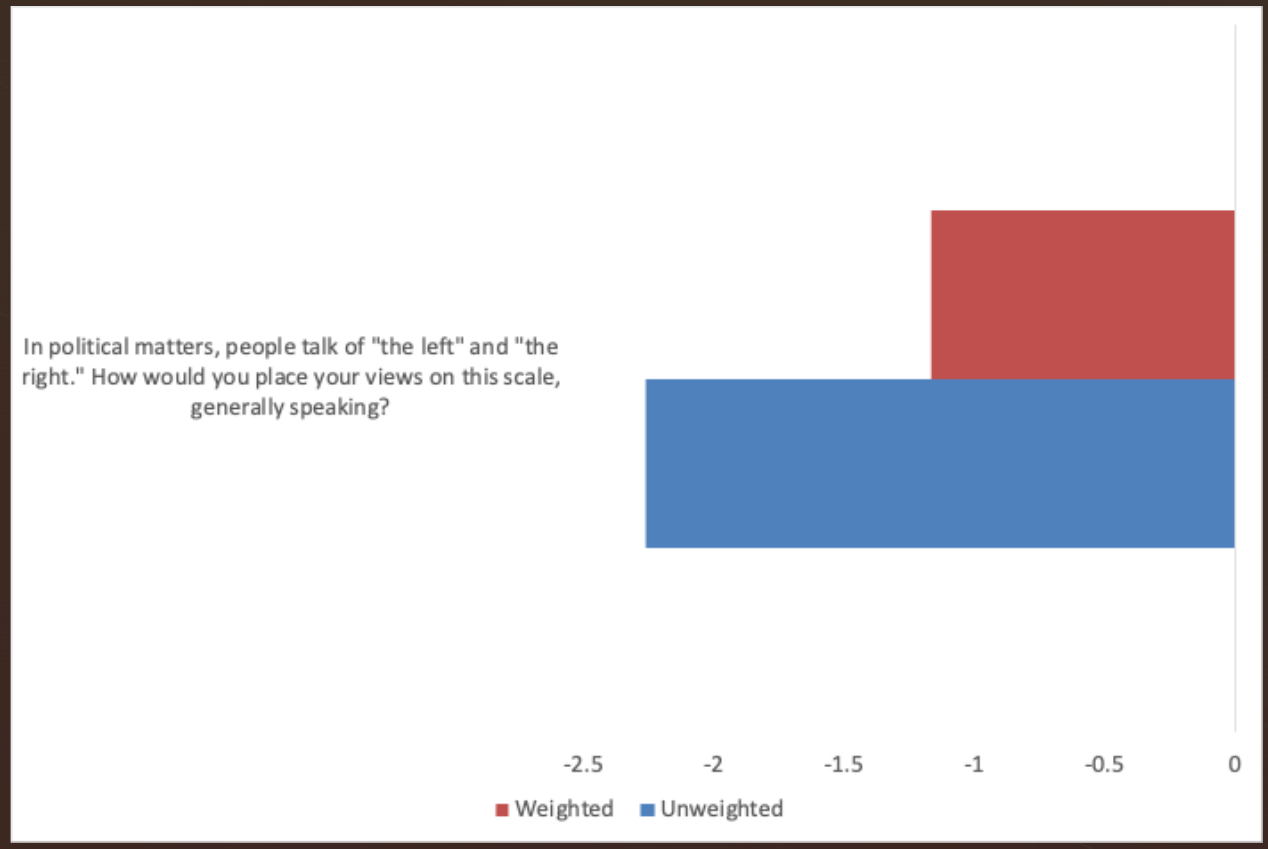
Evaluators may be a little more technocratic



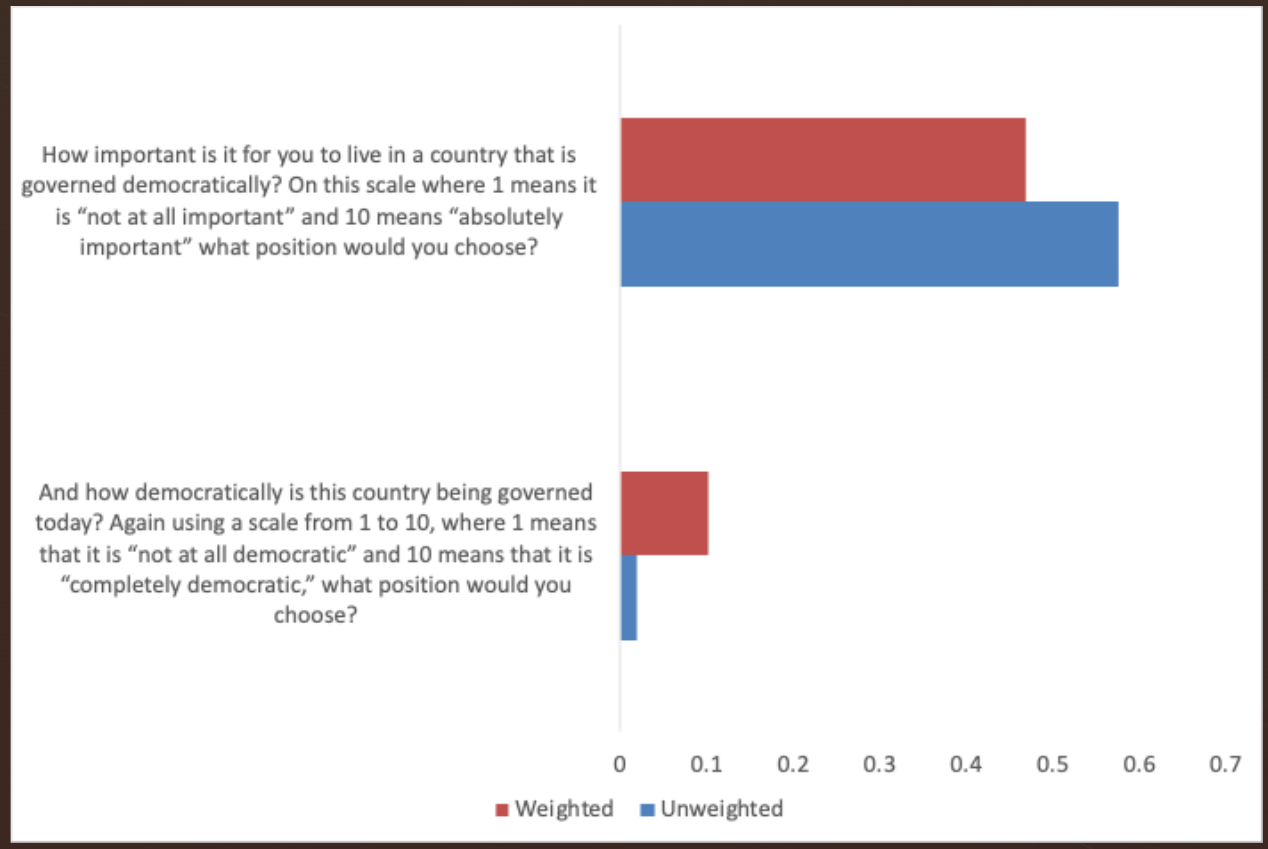
Overview

Question	P-UW	P-W	POR-UW	POR-W
In political matters, people talk of "the left" and "the right." How would you place your views on this scale, generally speaking?	**0.00**	**0.00**	Dom – 5.5	**Unc – 1**
How important is it for you to live in a country that is governed democratically? On this scale where 1 means it is “not at all important” and 10 means “absolutely important” what position would you choose?	0.05	0.10	**Null – 11**	Null – 1.2
And how democratically is this country being governed today? Again using a scale from 1 to 10, where 1 means that it is “not at all democratic” and 10 means that it is “completely democratic,” what position would you choose?	0.95	0.73	**Unc – 1**	Dom – 2
On a scale from 1 to 10 where “1” is “not satisfied at all” and “10” is “completely satisfied”, how satisfied are you with how the political system is functioning in your country these days?	0.29	0.10	Null – 1.9	**Null – 16** **Shift – 10**

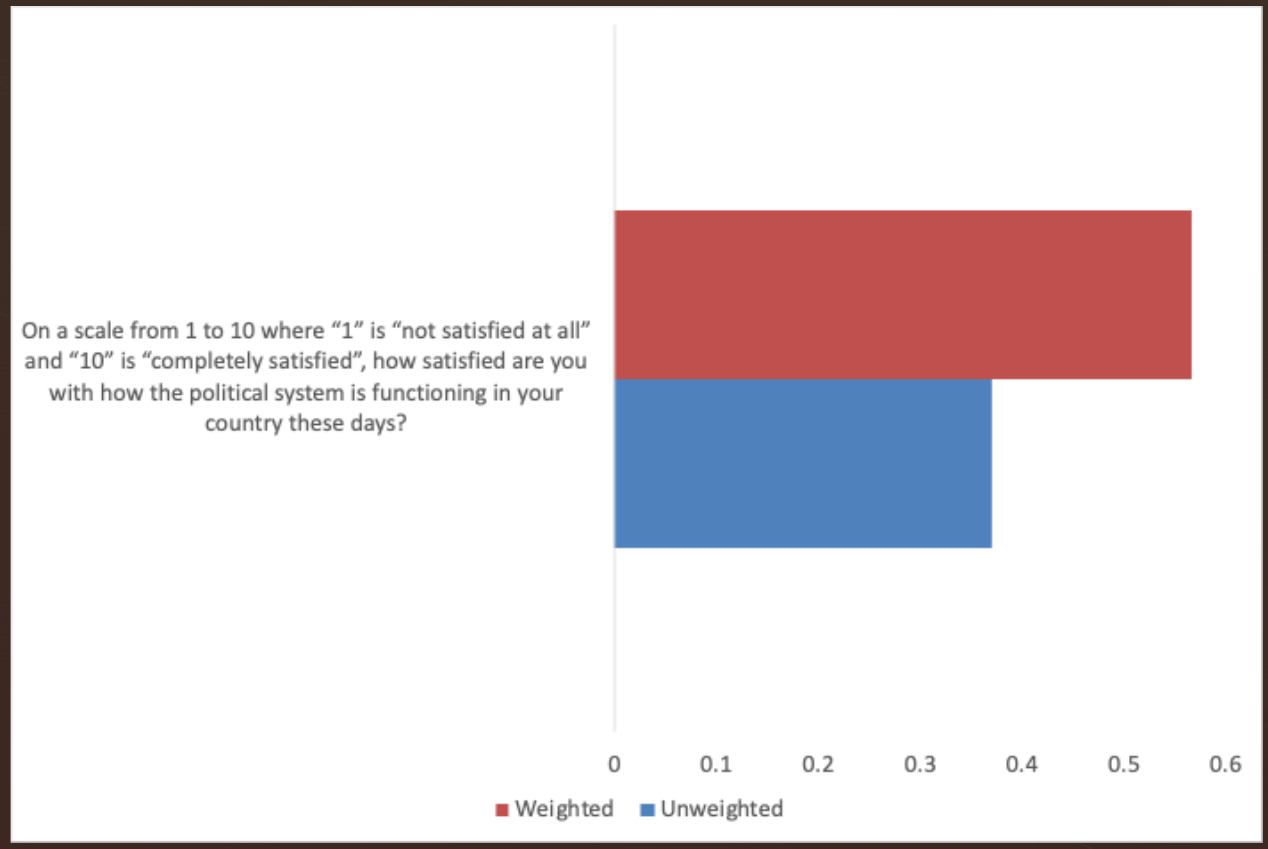
The Nolan Chart data agree.



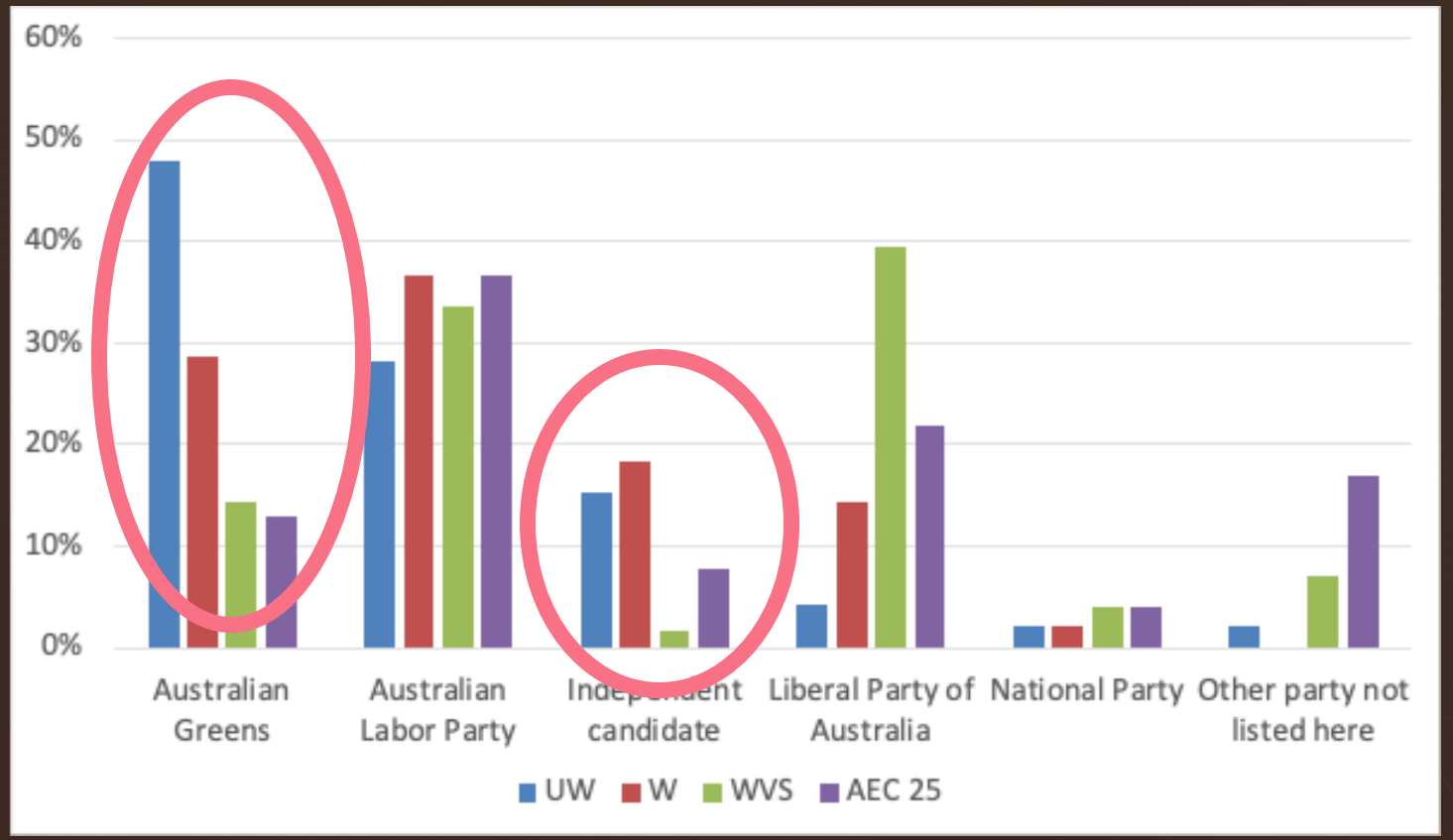
Evaluators prefer to live in democracies



And are more satisfied with the current political system



There is a strong cohort of Greens and independent voters

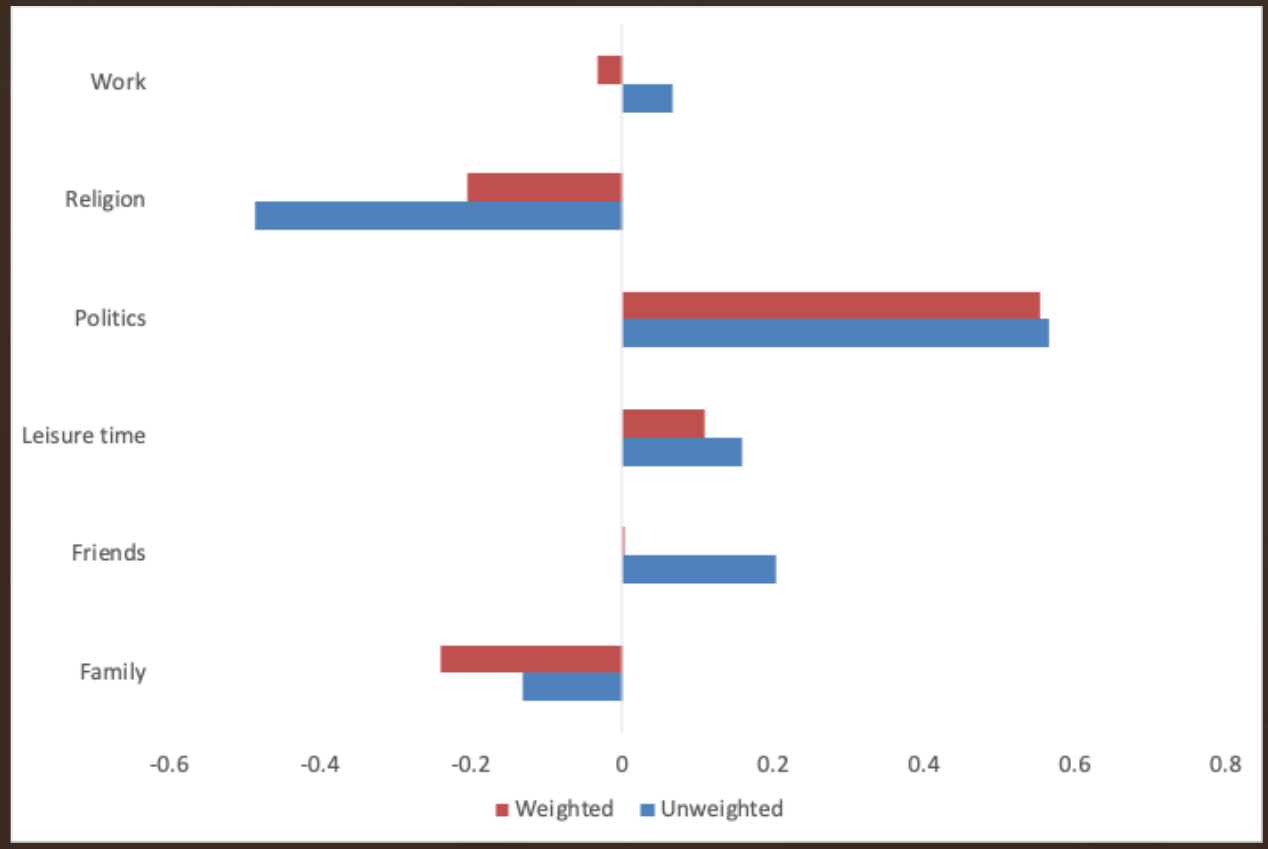


Social Values

Overview

Question	P-UW	P-W	POR-UW	POR-W
For each of the following, indicate how important it is in your life:				
—Family	**0.03**	**0.00**	Dom – 1.2	Dom – 1.2
—Friends	**0.02**	0.98	Shift – 1.1	Unc – 1
—Leisure time	0.08	0.23	Unc – 1	Unc – 1
—Politics	**0.00**	**0.00**	**Shift – 15**	**Shift – 22**
—Religion	**0.00**	0.19	Shift – 7.9	Unc – 1
—Work	0.59	0.79	Unc – 1	**Unc – 1**

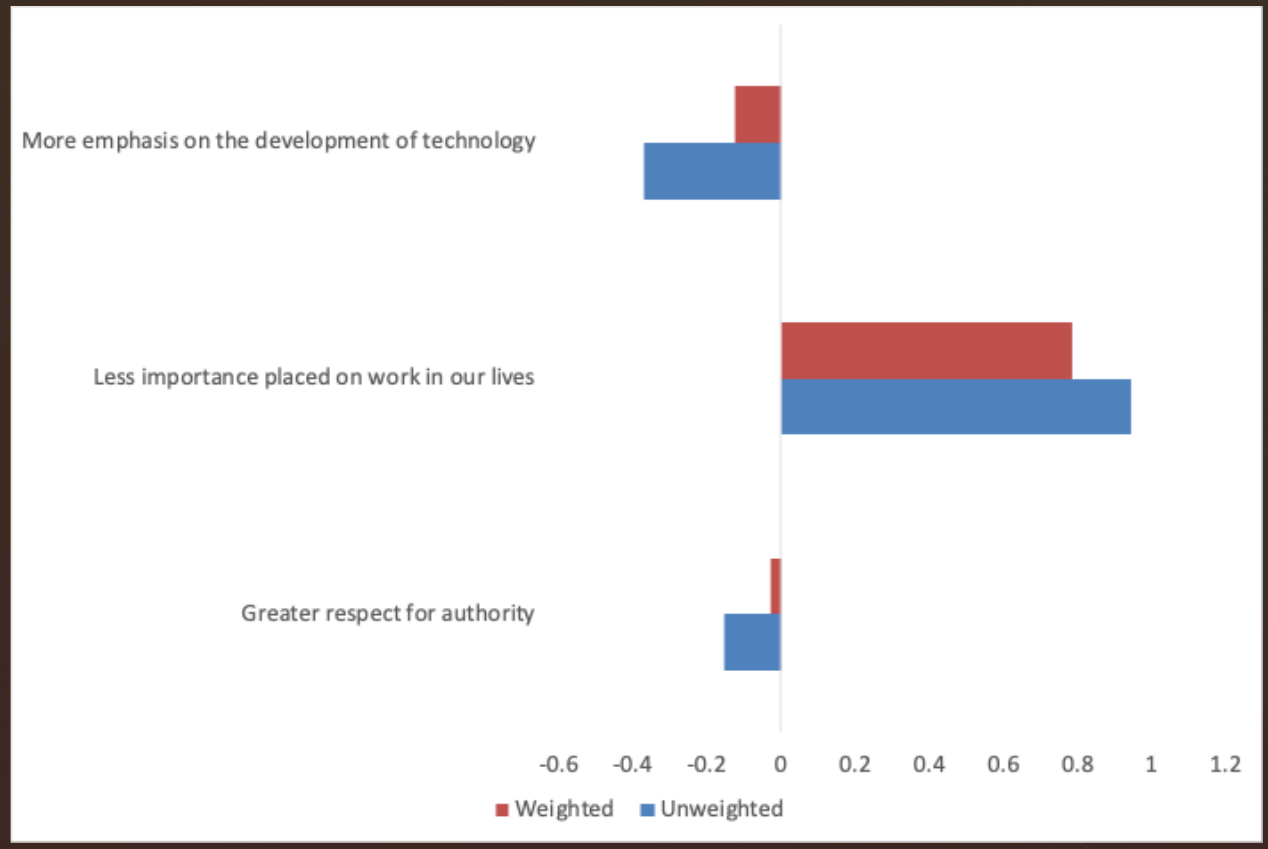
Politics matter for evaluators



Overview

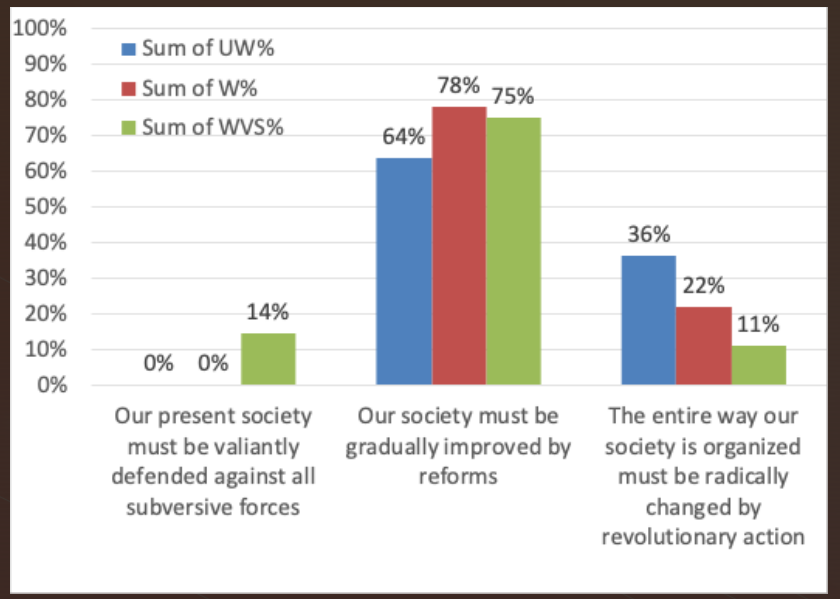
Question	P-UW	P-W	POR-UW	POR-W
Here is a list of various changes in our way of life that might take place in the near future. Please tell me for each one, if it were to happen, whether you think it would be a good thing, a bad thing, or don't you mind?				
—Greater respect for authority	0.12	0.77	Dom – 1.3	Unc – 1
—Less importance placed on work in our lives	**0.00**	**0.00**	Shift – 5.3	Dom – 1.8
—More emphasis on the development of technology	**0.00**	0.20	Dom – 1.5	Unc – 1

Evaluators might like less work



Evaluators are possibly bolshie!

- **Unweighted:** The chi-square statistic is 30.4905. The *p*-value is < 0.00001. The result is significant at $p < .05$.
- **Weighted:** The chi-square statistic is 10.6902. The *p*-value is .004772. The result is significant at $p < .05$.

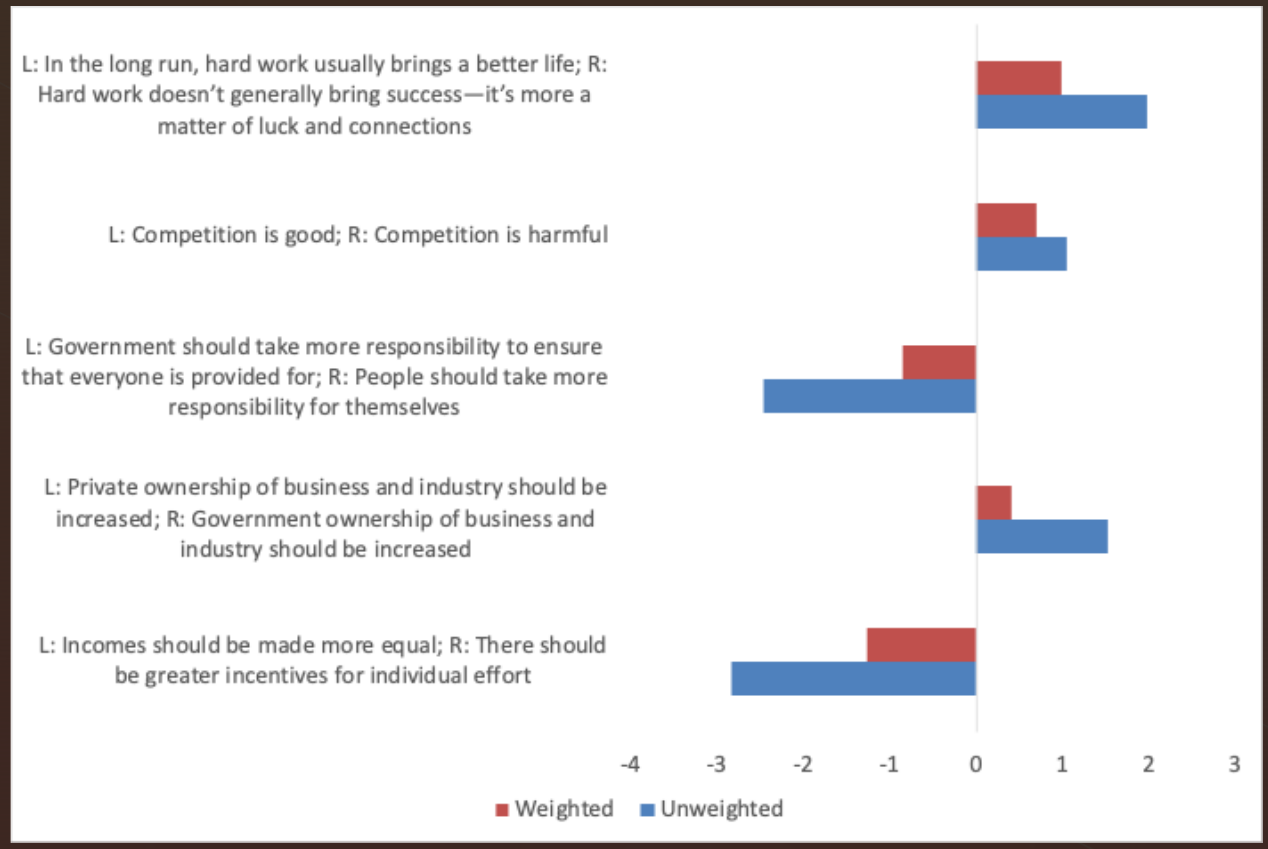


Economic Values

Overview

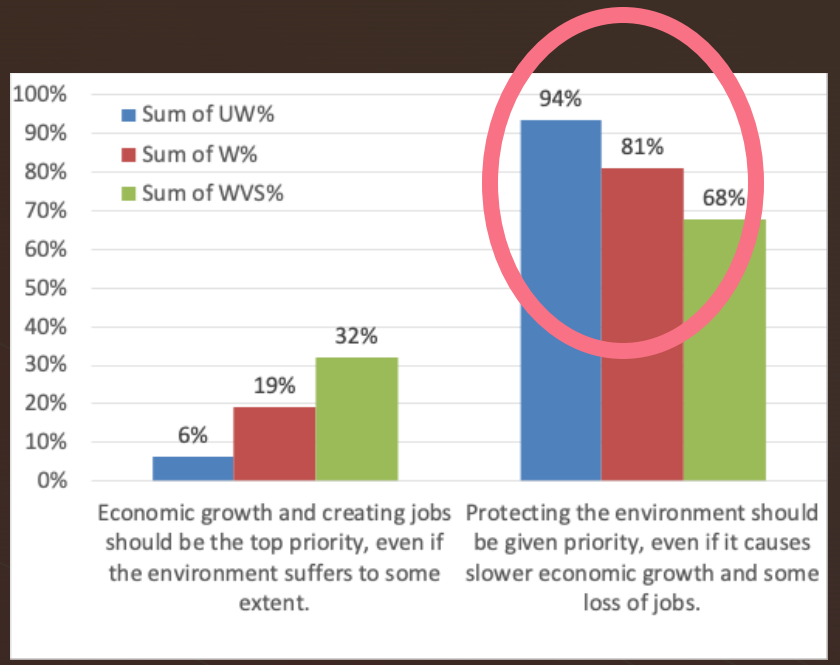
Question	P-UW	P-W	POR-UW	POR-W
How would you place your views on this scale? 1 means you agree completely with the statement on the left; 10 means you agree completely with the statement on the right; and if your views fall somewhere in between, you can choose any number in between.				
L: Incomes should be made more equal R: There should be greater incentives for individual effort	**0.03**	**0.00**	**Shift – 52,875**	Dom – 6
L: Private ownership of business and industry should be increased R: Government ownership of business and industry should be increased	**0.02**	0.98	**Shift – 6579**	Null – 5.7
L: Government should take more responsibility to ensure that everyone is provided for R: People should take more responsibility for themselves	0.08	0.23	**Shift – 109041**	Shift – 5.46
L: Competition is good R: Competition is harmful	**0.00**	**0.00**	Dom – 5.8	Dom – 3.4
L: In the long run, hard work usually brings a better life R: Hard work doesn't generally bring success—it's more a matter of luck and connections	**0.00**	0.19	**Shift – 5701**	Dom – 3

Evaluators are less divergent on economic policy



But broadly pro-environment

- Unweighted: The chi-square statistic is 14.1276. The p-value is .000171. Significant at $p < .05$.
- Weighted: The chi-square statistic is 3.8913. The p-value is .048538. Significant at $p < .05$.



The implications

We’ve got a (potential) bias problem

How do we reckon with this?

- In our work
- In how we present our field to commissioners and stakeholders
- In how we engage with the public

Interventions designed to reduce implicit prejudices and implicit stereotypes in real world contexts: a systematic review

Chloë FitzGerald ✉, Angela Martin, Delphine Berner & Samia Hurst

BMC Psychology **7**, Article number: 29 (2019) | Cite this article

Abstract

Background

Implicit biases are present in the general population and among professionals in various domains, where they can lead to discrimination. Many interventions are used to reduce implicit bias. However, uncertainties remain as to their effectiveness.

Methods

We conducted a systematic review by searching ERIC, PUBMED and PSYCHINFO for peer-reviewed studies conducted on adults between May 2005 and April 2015, testing interventions designed to reduce implicit bias, with results measured using the Implicit Association Test (IAT) or sufficiently similar methods.

Results

30 articles were identified as eligible. Some techniques, such as engaging with others' perspective, appear unfruitful, at least in short term implicit bias reduction, while other techniques, such as exposure to counterstereotypical exemplars, are more promising. Robust data is lacking for many of these interventions.

Conclusions

Caution is thus advised when it comes to programs aiming at reducing biases. This does not weaken the case for implementing widespread structural and institutional changes that are multiply justified.

Addressing implicit bias is hard

Checking our bias at the door

Make our implicit biases explicit

Then address them

…is that even possible or desirable?

Check vs check vs check

Check vs check vs check

check 1 of 3 **verb**

'chek

checked; checking; checks

Synonyms of *check* >

transitive verb

1 a : to inspect, examine, or look at appraisingly or appreciatively → usually used with *out* or *over*

3 : to leave or accept for safekeeping

check a coat

5 a : to restrain or diminish the action or force of

Argumentum ad auctoritatem (#5)
(I promise it's the last one)

Evaluations which neglect to consider values are at risk of becoming irrelevant and unused.

- Indran Naidoo et al., "Value-based evaluation and why it matters", 2023

Yes, and…

Evaluators who neglect to consider their values are at risk of becoming irrelevant and unused.

Discussion

Thanks

- The World Values Survey for providing access to their instruments and reporting
- All the cited sources who open access published their studies
- Martin Schnuerch and Jeff Rouder for their research and tools on Bayesian analysis of ordinal scale data (also open access and open source published)