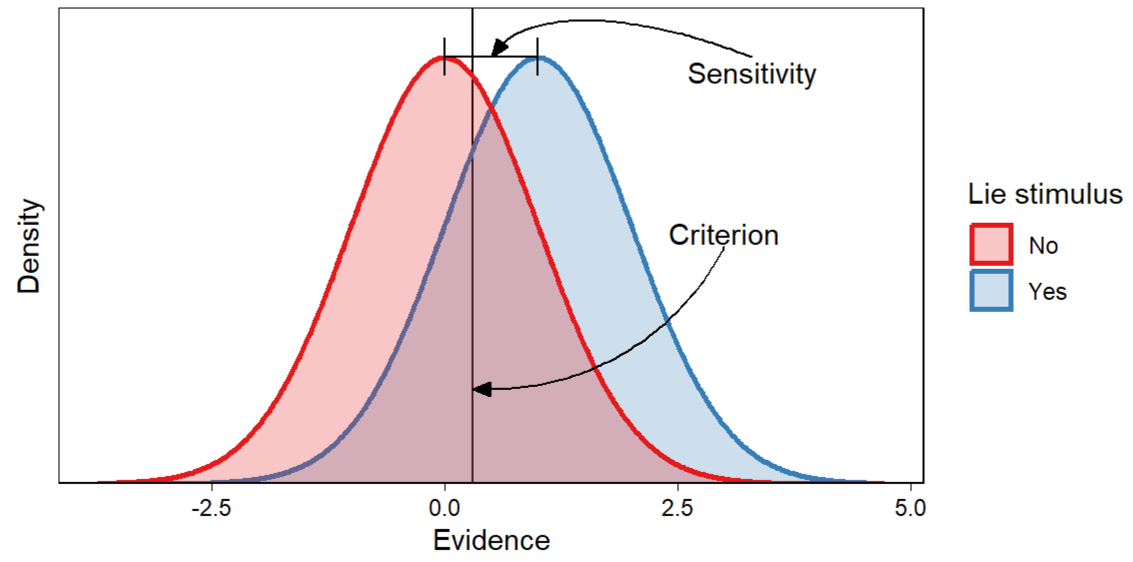
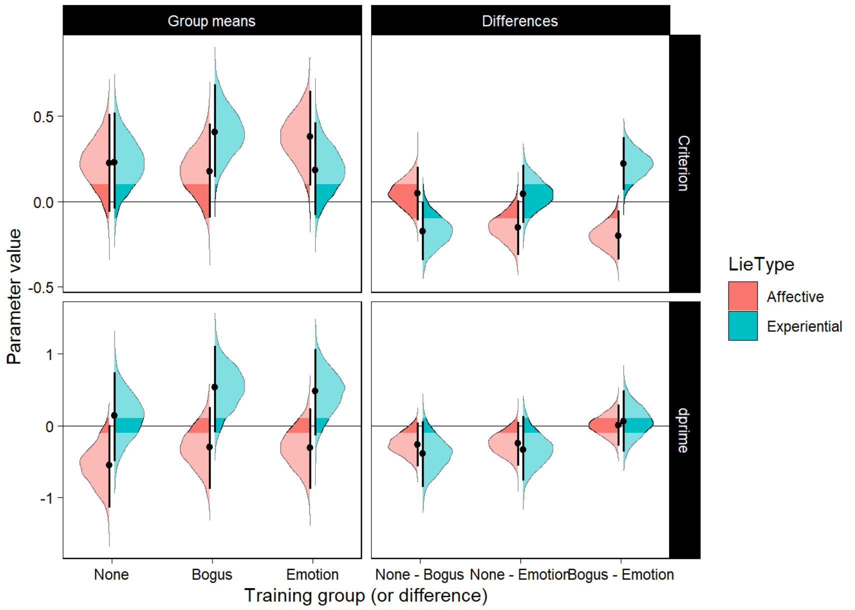
Veracity Judgements: Reinterpreting the processes behind "deception detection"

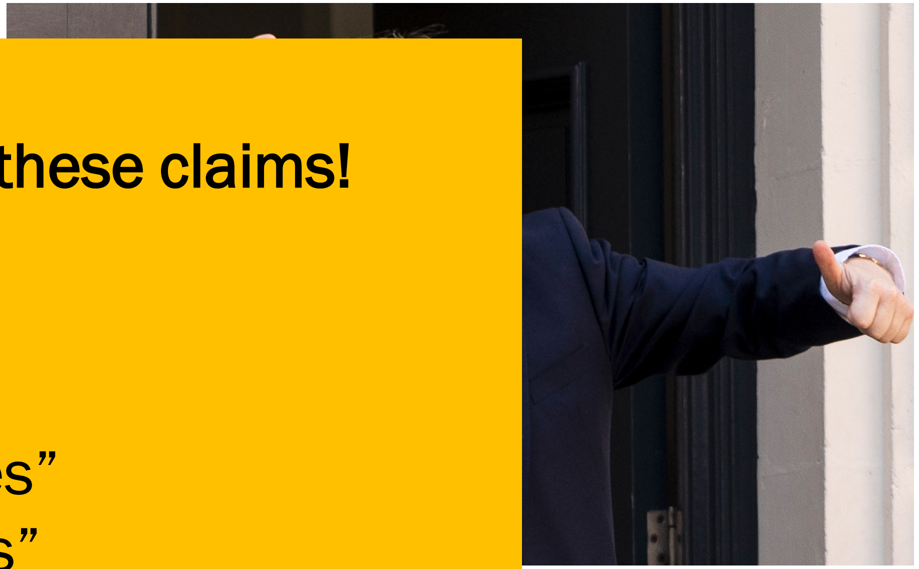
MIRCEA ZLOTEANU





Detecting Lies and Truths: Usual suspects

- People lie often and for various reasons
- However, de
  - Lies detec
- Bias toward



All deception papers reiterate these claims!

- 54% accuracy
- veracity effect
- truth bias
- (mysterious) “cues”
- astute “decoders”

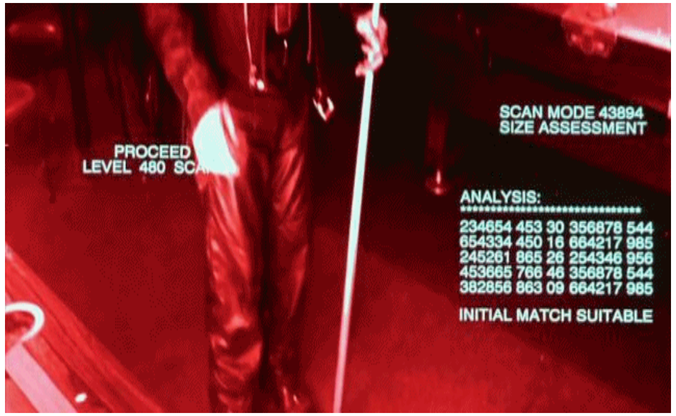
- No one definitive cue of lying (DePaulo et al., 2003)
- Meta-analyses find (inconsistent) “cues”

Theory: How do people judge veracity?

Deception Detection

- Default position
- Reliable dif. b/w Liars and Truth-tellers
- Assumes diagnostic cues
- "Decoders" perceive behavioural cues
- Accuracy **exists** > properly interpreting cues
- (universal, cross-cultural, involuntary)

Evidence? weak



Veracity Judgements

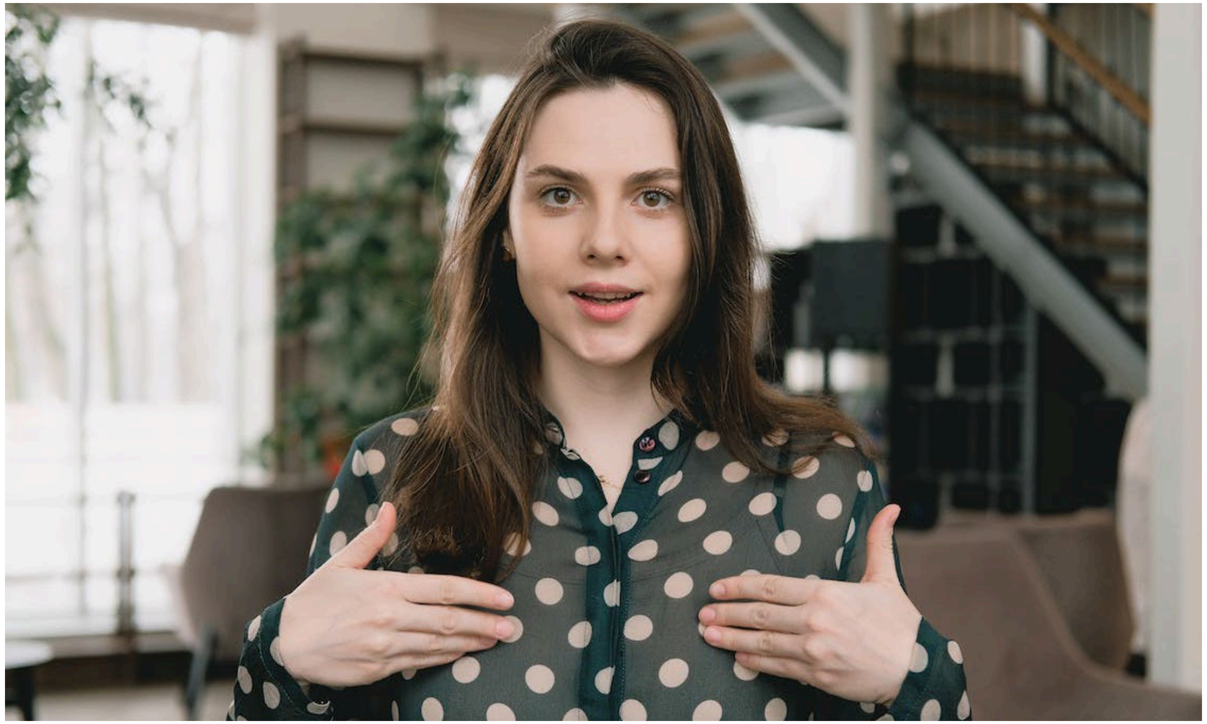
- "Judges" make inferences about others' veracity
- No need for "deception cues"
- Does not assume (a priori) accuracy is possible (no diagnosticity)
- Measure shifts in judgement (e.g., bias, confidence)
- Focus on situation, contexts, & individuals

Compatible with most findings

Solutions?

- **paradigm shift (hard)**
- **open science (moderate)**
- **better analysis plan (easy!)**

Typical study



Dishonest	Somewhat Dishonest	Unsure	Somewhat Honest	Honest
Not at all Confident	2	3	4	Very Confident

Typical Data Processing

- This is where the magic happens (read: QRPs)

Input Responses

	A	B	C	D	E	F
1	**Participant**	**Condition**	**AnswerV1**	**AnswerV2**	**AnswerV3**	**AnswerV4**
2	1	Control	Lie	Lie	Lie	Lie
3	2	Training	Truth	Truth	Truth	Lie
4	3	Control	Lie	Truth	Lie	Lie
5	4	Training				

Match Veracity

	A	B	C	D	E	F
1	**Participant**	**Condition**	**AnswerV1**	**AnswerV2**	**AnswerV3**	**AnswerV4**
2	1	Control	Correct	Incorrect	Incorrect	Incorrect
3	2	Training	Incorrect	Correct	Correct	Incorrect
4	3	Control	Correct	Correct	Incorrect	Incorrect
5	4	Training				

Tally Correct

	A	B	C	D
1	**Participant**	**Condition**	**Truth.#**	**Lie.#**
2	1	Control	5	3
3	2	Training	3	1
4	3	Control	3	0
5	4	Training		

Percentage

	A	B	C	D
1	**Participant**	**Condition**	**Truth.%**	**Lie.%**
2	1	Control	50%	30%
3	2	Training	30%	10%
4	3	Control	30%	0%
5	4	Training	40%	20%

Typical Data Processing: NOT DONE YET!!!

Response Bias

Participant	Condition	AnswerV1	AnswerV2	AnswerV3	AnswerV4
1	Control	Lie	Lie	Lie	Lie
2	Training	Truth	Truth	Truth	Lie
3	Control	Lie	Truth	Lie	Lie
4	Training	Lie	Lie	Truth	Lie

Participant	Condition	AnswerV1	AnswerV2	AnswerV3	AnswerV4
1	Control	-1	-1	-1	-1
2	Training	1	1	1	-1
3	Control	-1	1	-1	-1
4	T			1	-1

Participant	Condition	Bias
1	Control	-4
2	Training	2
3	Control	-2
4	Training	-2

Signal Detection Theory (SDT)

Participant	Condition	CR	FA	HIT	MISS
1	Control	4	2	3	2
2	Training	3	1	1	3
3	Control	4	3	0	1
4	Training	1	3	2	0

Do some shady 0 and 1 correction

Participant	Condition	dprime	c	bias
1	Control	0.39	0.21	0.11
2	Training	0.87	0.63	0.19
3	Control	0	0.26	0.05
4	Training	0.21	1.01	-0.17

Typical Analyses

- Then, we report up to 4 (unconnectable) analyses (ANOVAs)

misleading & wrong

Accuracy %

Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p
Veracity	156.641	1	156.641	74.413	< .001
Veracity ✻ Condition	2.150	2	1.075	0.511	0.602
Residuals	216.817	103	2.105		

Note. Type III Sum of Squares

Between Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p
Condition	10.459	2	5.230	1.700	0.188
Residuals	316.809	103	3.076		

Note. Type III Sum of Squares

Bias

ANOVA - BDS_Bias

Cases	Sum of Squares	df	Mean Square	F	p
Condition	17.199	2	8.600	0.511	0.602
Residuals	1734.537	103	16.840		

Note. Type III Sum of Squares

OK-ish (but could be better)

SDT – Accuracy (d')

ANOVA - BDS_Aprime

Cases	Sum of Squares	df	Mean Square	F	p
Condition	0.102	2	0.051	1.433	0.243
Residuals	3.655	103	0.035		

Note. Type III Sum of Squares

SDT – Bias (c)

ANOVA - BDS_b2p ▼

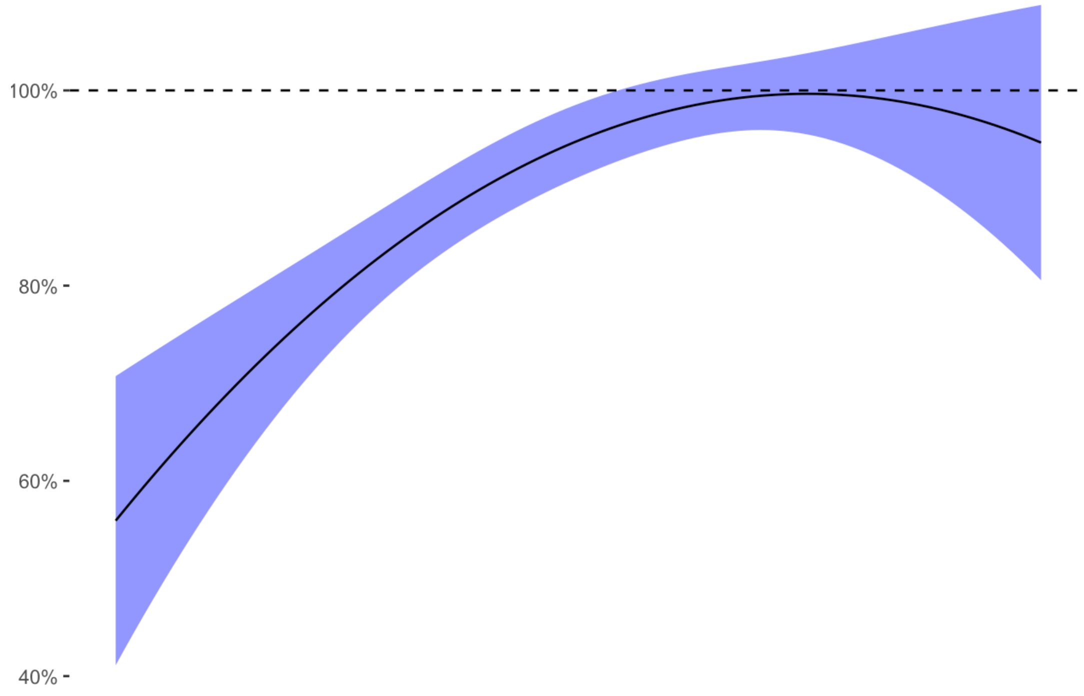
Cases	Sum of Squares	df	Mean Square	F	p
Condition	0.167	2	0.083	0.617	0.541
Residuals	13.906	103	0.135		

Note. Type III Sum of Squares

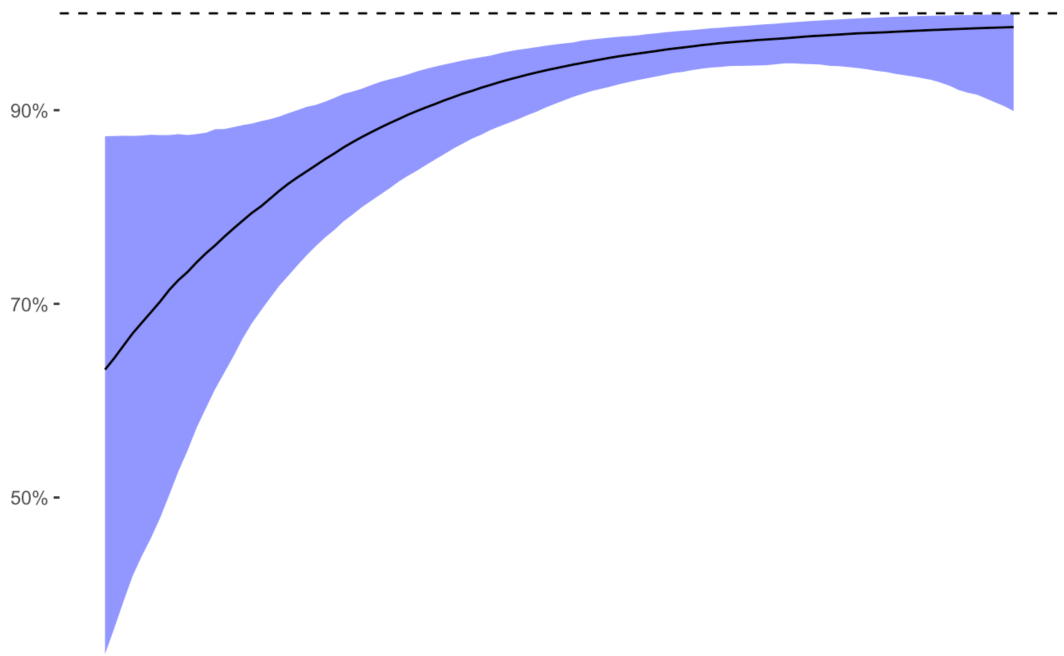
The tools we are using are not fit for purpose

- Models that assume **unbounded** and **normal** data will produce inappropriate estimates
  - ANOVAs will estimate 120% or -10% values (which are impossible), and ‘create’ differences that don’t exist

Ordinary Least Squares (OLS)



Ordered Beta Regression



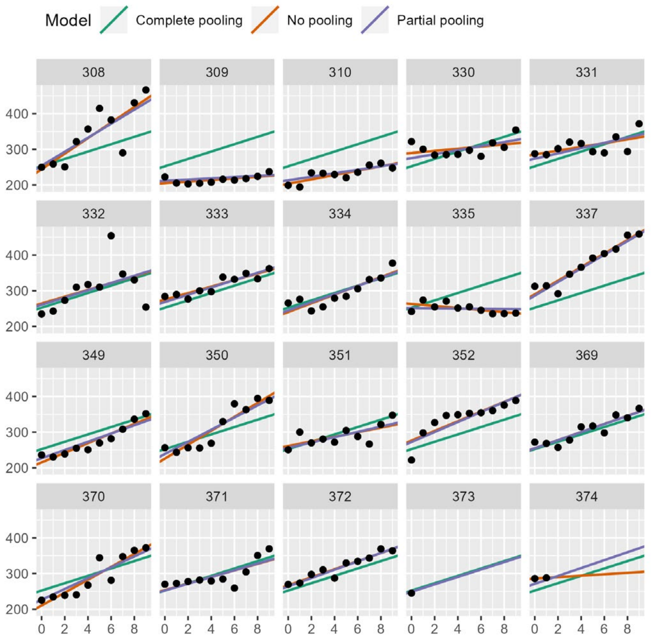
We can do better: Bayesian Mixed Effects Models

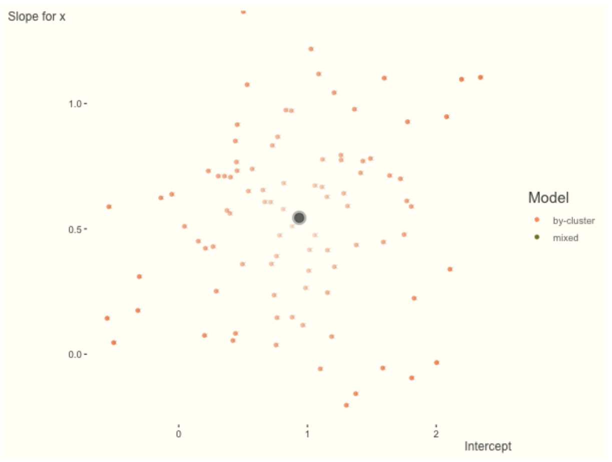
Advantages of LMMs over RM-ANOVA

- Same results as simple analyses
- All data types permitted
- Complex designs
- Missing data
- Differing number of repeats (unbalanced data)

Bayesian vs. Frequentist

- Better chance of model fit
- No issues with dfs and computing “significance”
- Provides the information one wants (Prob. of hypothesis given data)
- Richer data (everything has a posterior distribution to interpret)
- No dichotomous thinking (~~p-values~~)
- More flexibility in exploring competing models/explanations





We can do better: Bayesian Mixed Effects Models

- ANOVAs treat all Trials/Videos as equivalent, but:

Video 1

Video 2

Video 3

Video 4

Not all the same, regardless of labels

- ANOVAs treat all Participants as homogenous, but:

P 1

P 2

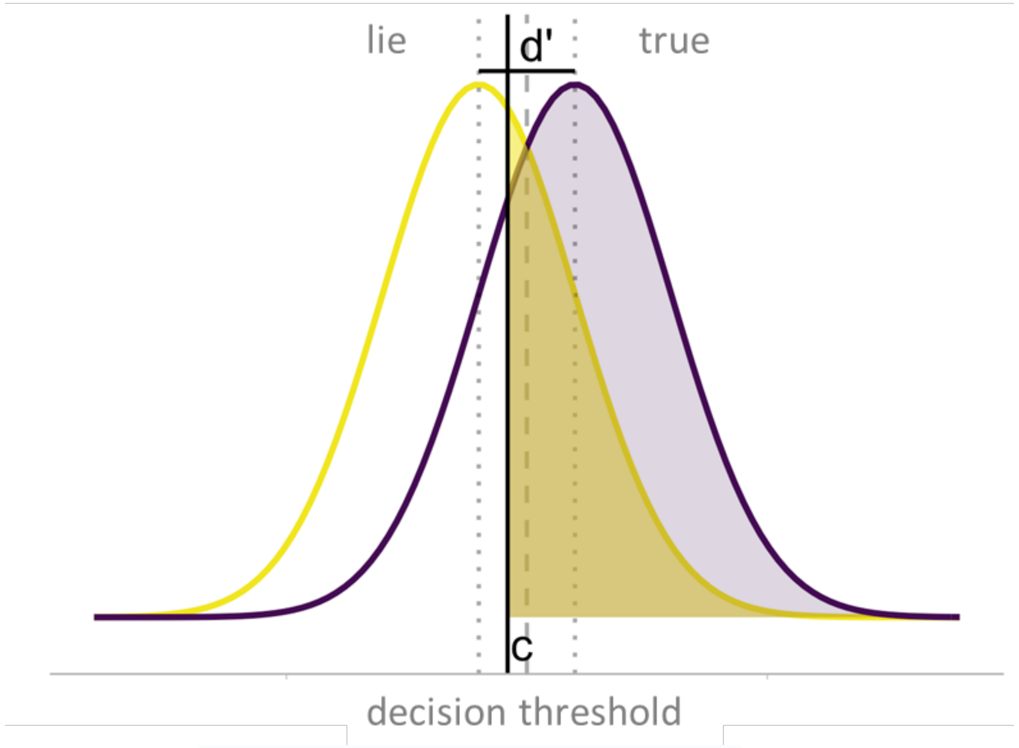
P 3

P 4

All response differences are ignored

Signal Detection Theory (briefly)

- Classify a stimulus into one of two categories, e.g., Y/N, Old/New, Lie/Truth
- Measuring a *latent psychological dimension*
- Uses the binary classification: “lie” and “truth”
- Participants judge if a stimulus belongs to A or B, based on some internal process (decision rule)
  - d’: sensitivity; distance between distributions
  - c: location of decision threshold (relative to no-bias)

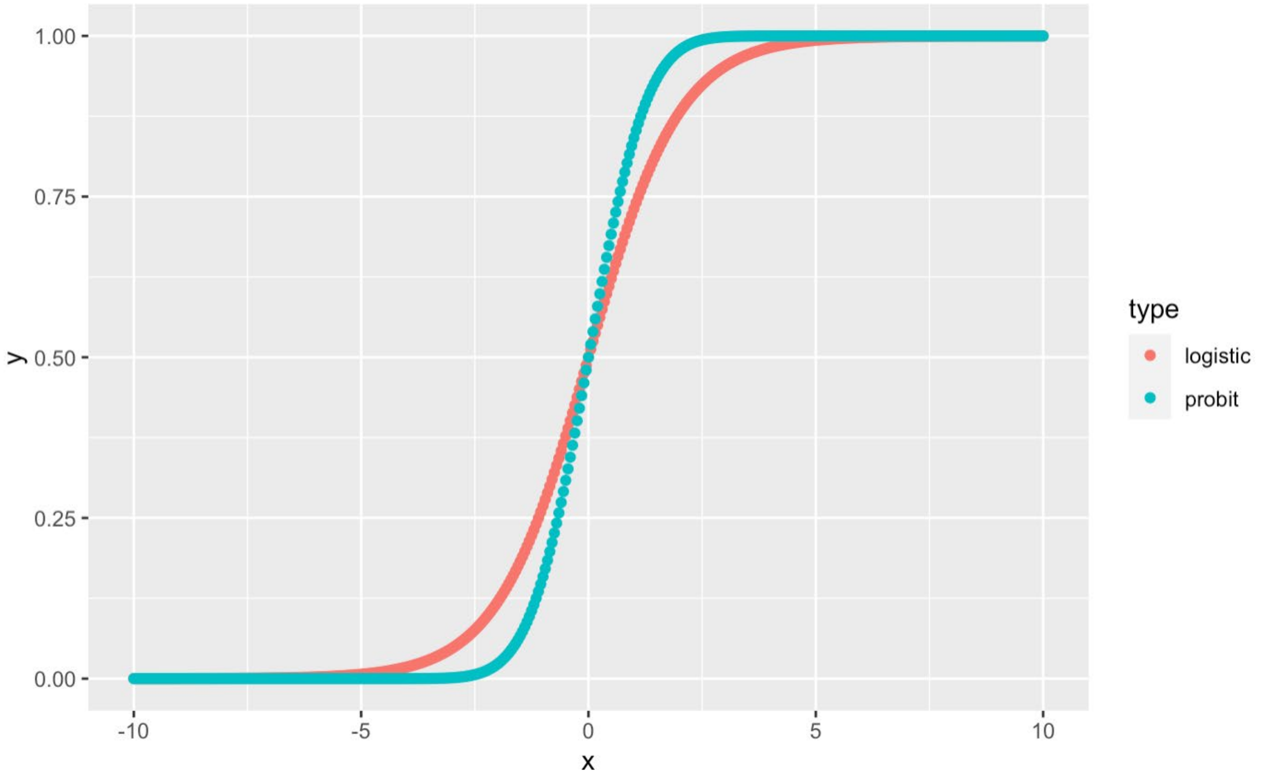


		Perceiver’s response	
		“Yes”	“No”
Stimulus	Present	Hit	Miss
	Absent	False alarm	Correct rejection

We can do better: Bayesian Mixed Effects Models

Solution:

- Probit Models !
- Directly model the binary (L/T) responses
- Translate into a SDT model (DeCarlo, 1998)
  - intercept = -c (**bias)**
  - $\beta 1$ = d' **(sensitivity)**
- Incorporates all sources of variance
- Bayesian inferences
  - Magnitude & uncertainty
  - Existence
  - Importance
  - (prediction)



Probit Models

Data Preparation

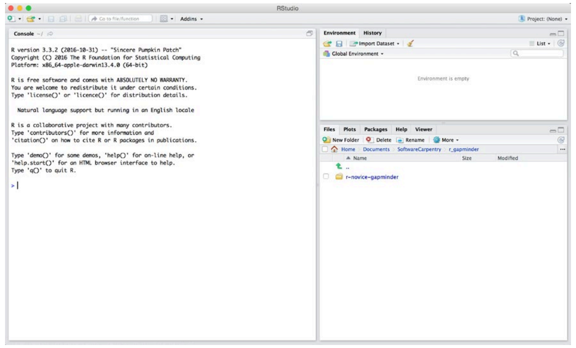
Long format

	A	B	C	D	E
1	Participan	Stimuli	Veracity	Answer	Condition
2	1	video1L	1	1	ERT
3	1	video2L	0	0	ERT
4	1	video3T	0	1	ERT
5	1	video4T	1	1	ERT
6	2	video1L	0	0	BT
7	2	video2L	0	1	BT
8	2	video3T	1	1	BT
9	2	video4T	0	1	BT
10	3	video1L	0	1	CT
11	3	video2L	1	0	CT
12	3	video3T	0	1	CT
13	3	video4T	1	1	CT

Truth = 0
Lie = 1

R package

```
library(brms)
library(bayestestR)
library(emmeans)
```



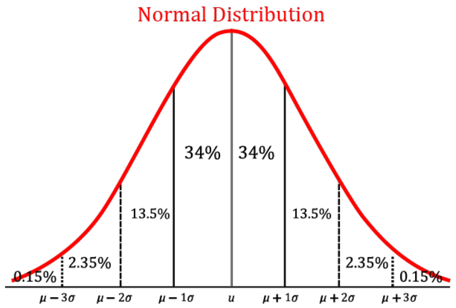
Bayesian Priors (optional)

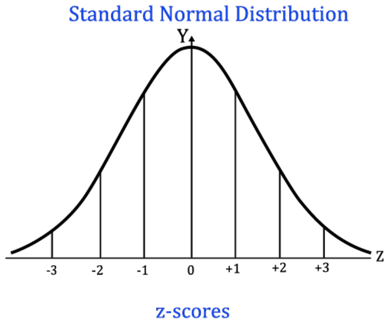
As the probit CDF back-transforms to a Z-curve, we can see the parameter space the same as a Normal distribution with mean = 0, and SD = 1.

So, think in terms of Cohen's d!

Some acceptable priors for the coefficients is N ~ (0,1)

But, avoid being too specific for interactions.



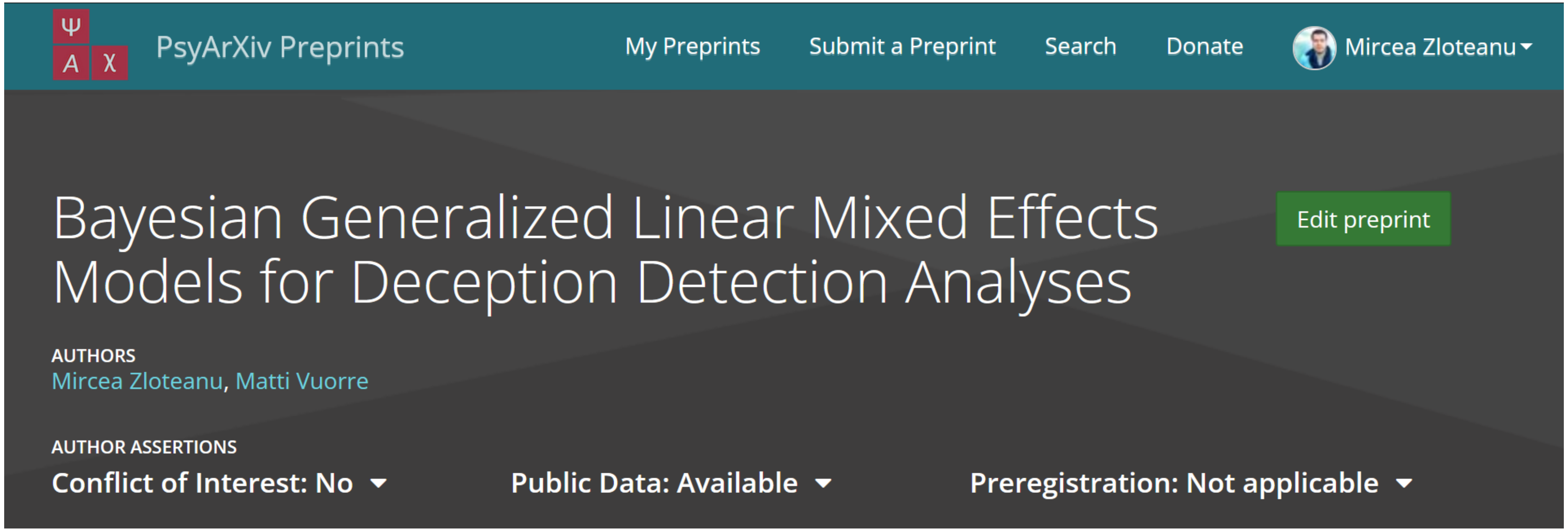


Full details in paper:

Preprint PsyArxiv

(submitted to JONB)

- https://psyarxiv.com/fdh5b/



Probit Models

- Easy to do

-c

d’

Participant variance

Video variance

Syntax

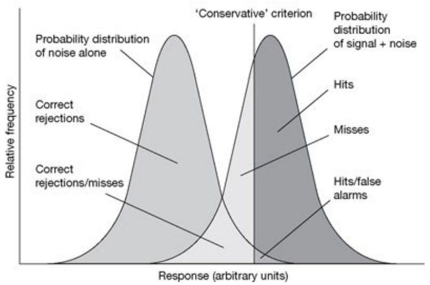
```
m0 <- brm(sayLie ~ 1 + isLie + (1 + isLie | Participant) + (1 | Stimulus),
          data = sdt.data,
          family = bernoulli(link = probit))
```

Output

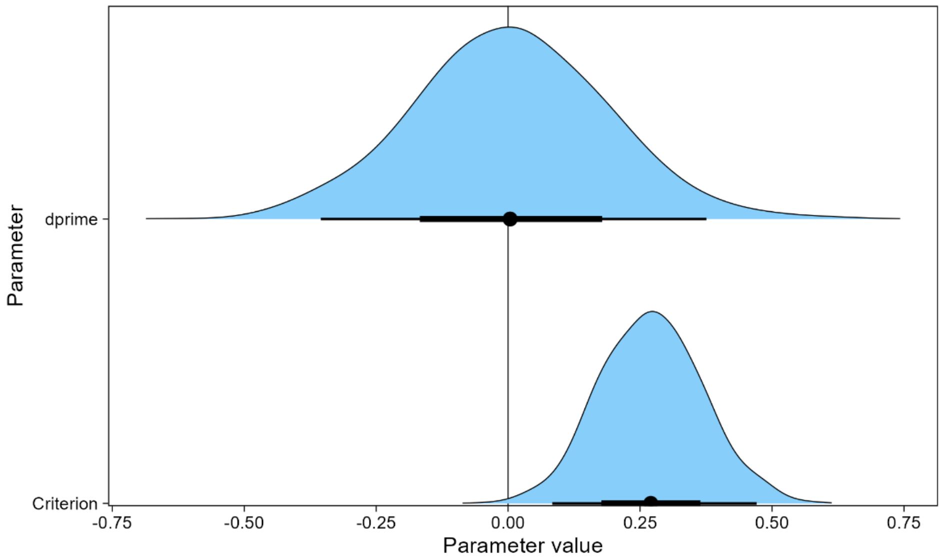
```
# Fixed Effects
Parameter    |    Mean |         95% CI |     pd |  Rhat |    ESS
-----------------------------------------------------------------
(Intercept)  |   -0.27 | [-0.47, -0.08] | 99.83% | 1.009 | 574.00
isLie1       |   0.004 | [-0.35,  0.38] | 50.78% | 1.003 | 819.00
```

Note: c = -Intercept

“0.27” ergo a “truth bias”

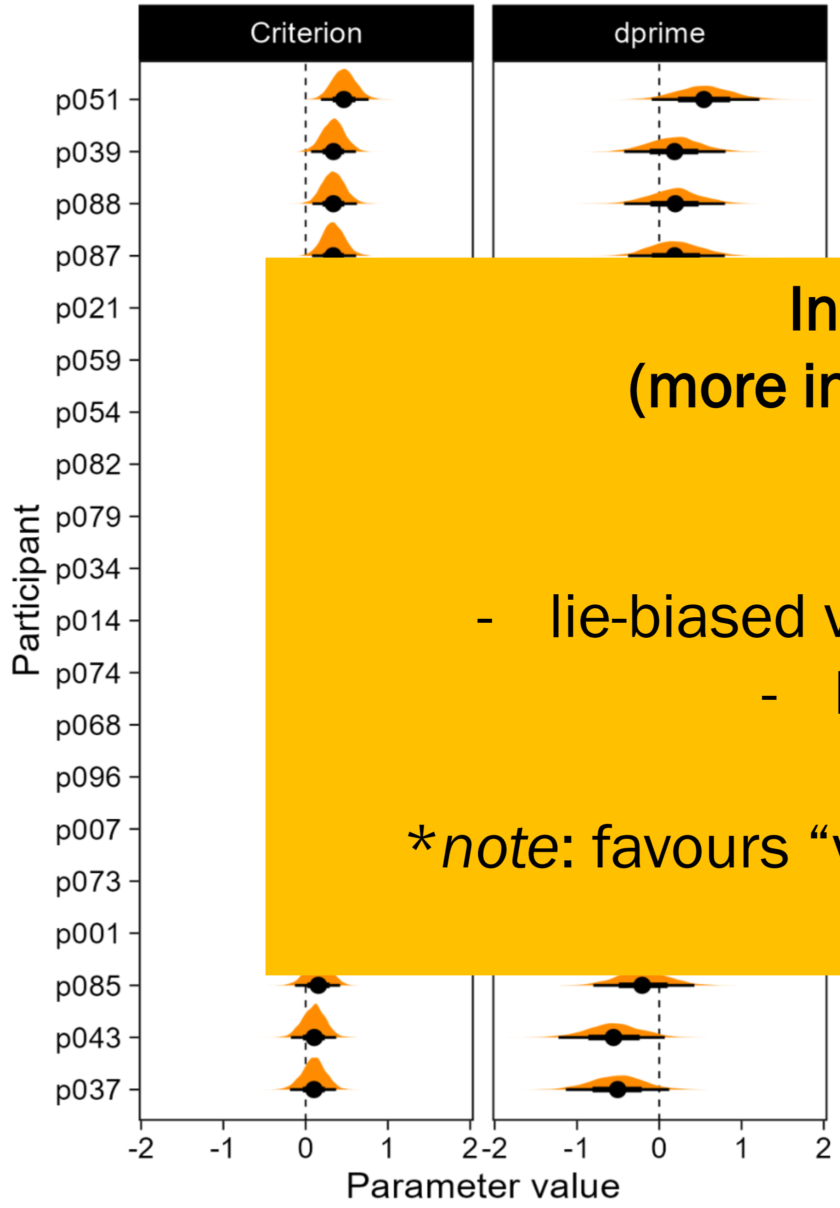


Plots

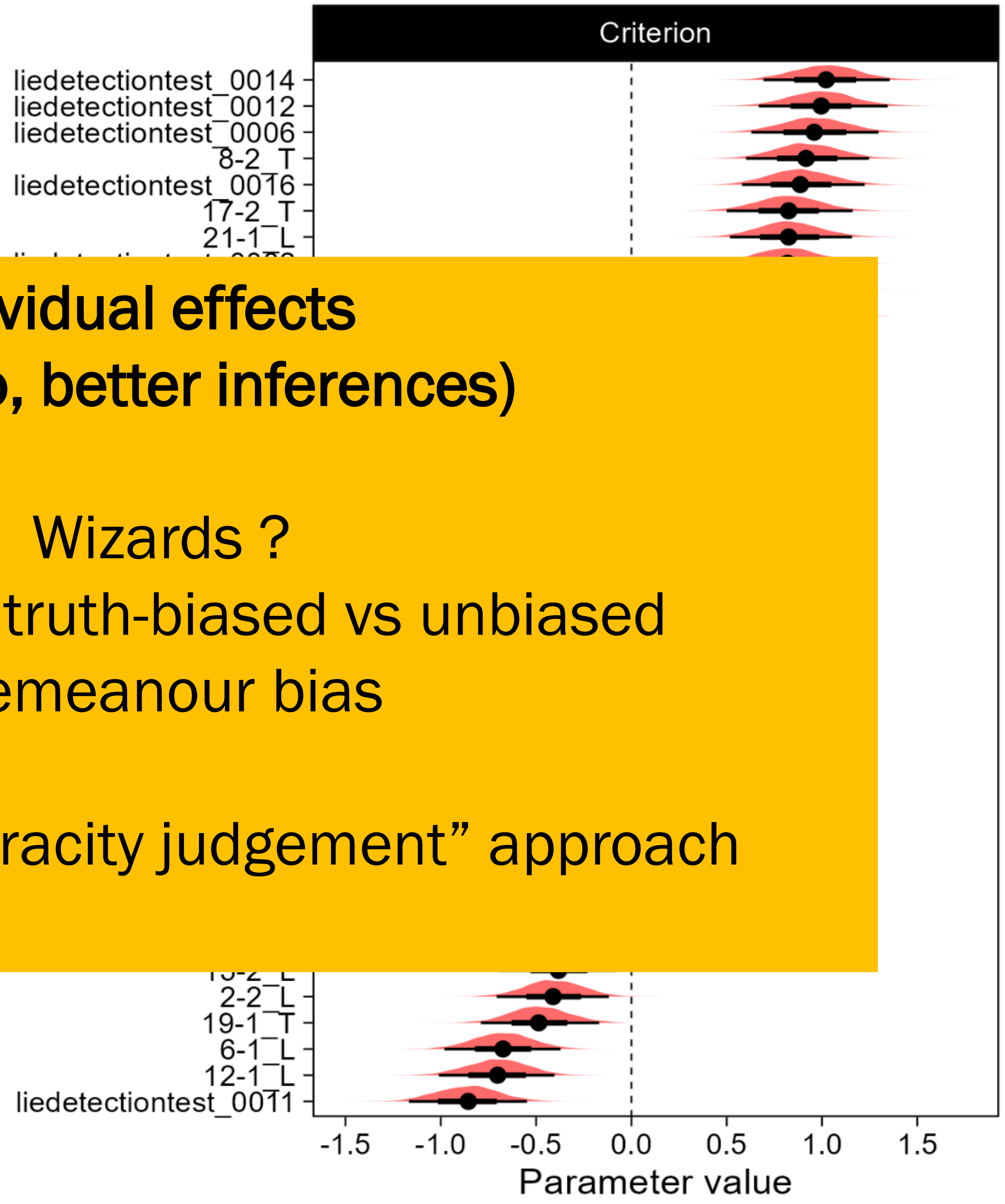


Sensitivity?
... 0

`(1 + isLie | Participant)`



`(1 | Stimulus)`



Individual effects
(more info, better inferences)

- Wizards ?
- lie-biased vs truth-biased vs unbiased
- Demeanour bias

**note*: favours "veracity judgement" approach

Complex designs model, easy:

- Within-participant factors:

Syntax

Within-subjects (repeated)

WS factors require random slopes

```
m2 <- brm(sayLie ~ 1 + LieType + isLie + LieType:isLie + (1 + LieType*isLie | Participant) + (1 | Stimuli),
          data = sdt.data,
          family = bernoulli(link = probit))
```

- Both within and between factors

Syntax

Shorthand for main effects + interactions

```
m3 <- brm(sayLie ~ 1 + Condition*LieType*isLie + (1 + LieType*isLie | Participant) + (1 | Stimuli),
          data = sdt.data,
          family = bernoulli(link = probit))
```

Compare Models

```
> loo_models
   elpd_diff se_diff
m2  0.0       0.0
m0 -2.5       2.7
m3 -2.9       2.7
m1 -3.6       3.3
```

Easy way to select the “best” model (considers model fit and complexity)

All data-types welcome!

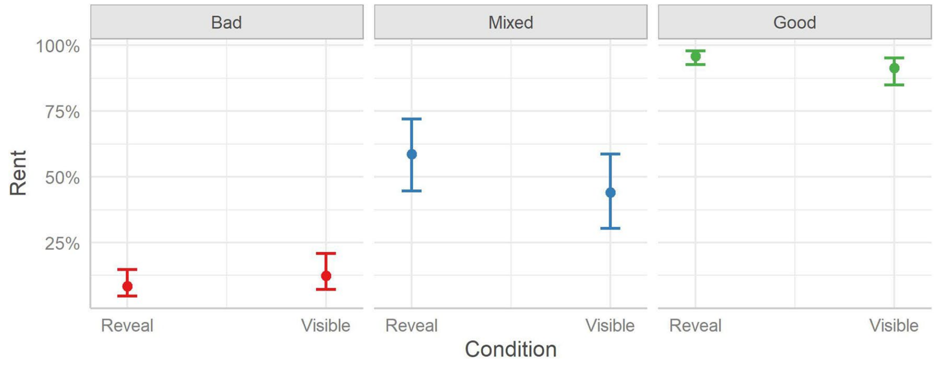
Syntax

```
m2 <- brm(sayLie ~ 1 + isLie + (1 + isLie | Participant) + (1
| Stimuli),
          data = my_data,
          family = categorical())
```

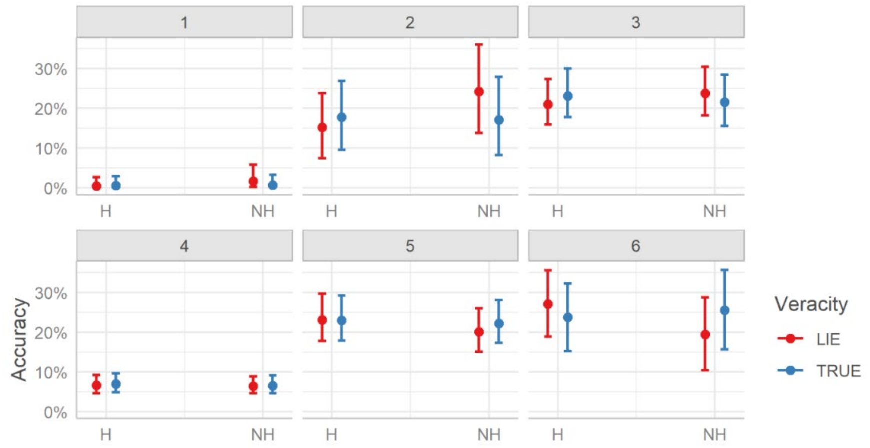
Syntax

```
m3 <- brm(sayLie ~ 1 + isLie + (1 + isLie | Participant) + (1
| Stimuli),
          data = my_data,
          family = cumulative())
```

- Categorical



- Ordinal (Likert, scale)

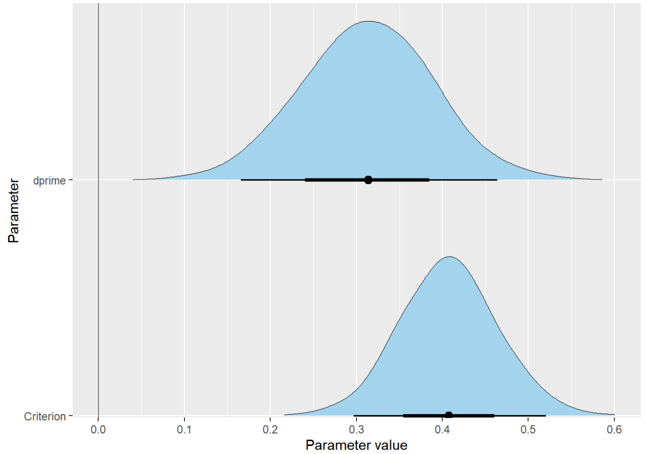


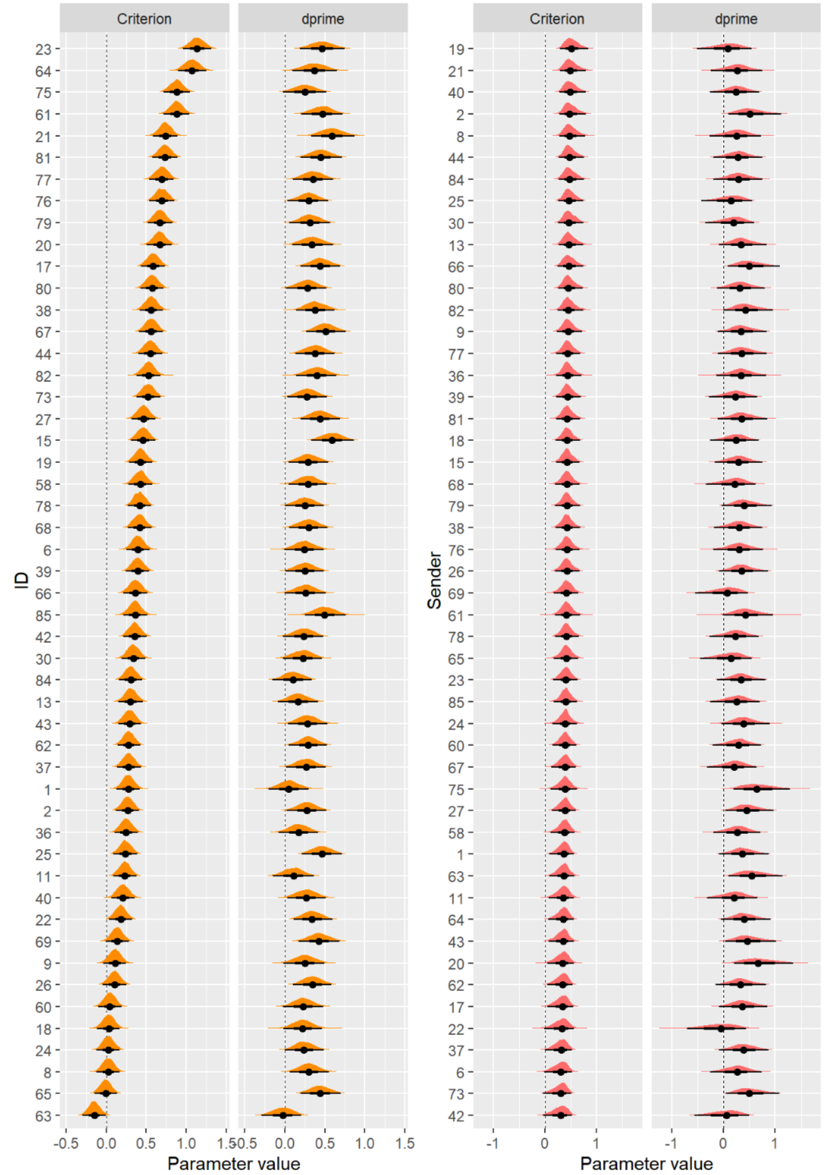
Example of novel insights: Masip, Levine, Somastre, & Herrero (2020)

Syntax

```
m1 <- sayLie ~ 1 + isLie + (1 + isLie | ID) + (1 + isLie | Sender) + (1 | Sender : Question)
```

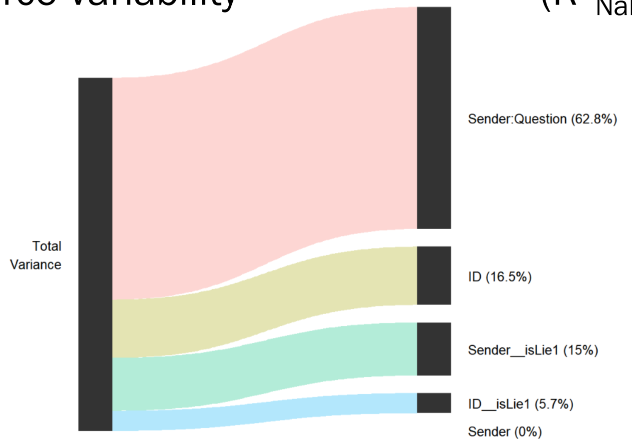
- Demeanour bias and Judge variability:





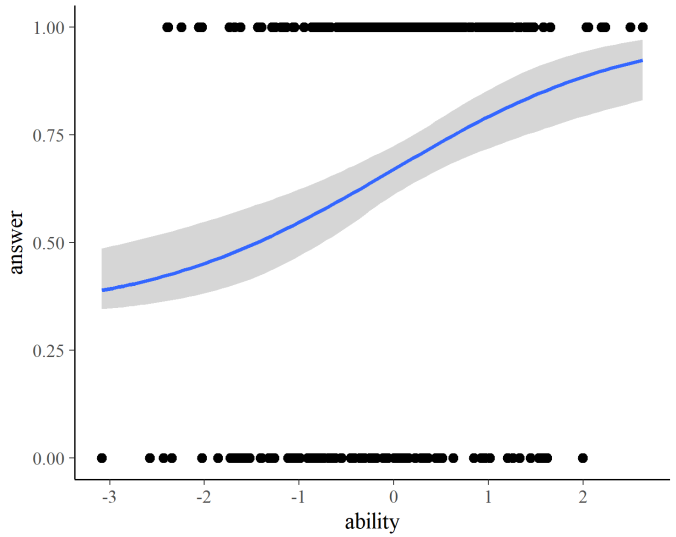
- Source variability

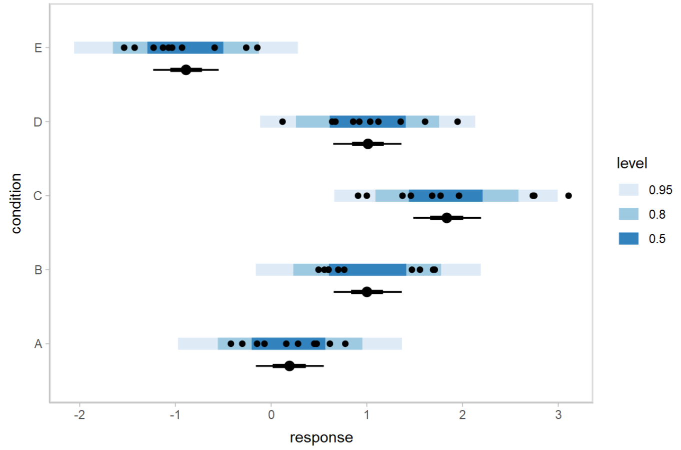
($R^2_{Nakagawa}$: 0.31)



Just scratching the surface! More options:

- Type-II SDT
  - Veracity + Confidence – meta-detection
- Different SDT models
  - Mixture, dual-process, etc.
- Unequal variance (honesty scales)
  - see Zloteanu, et al. 2022; in SI
- Taking guessing into account
  - 50% [chance] correction
- Dealing with outliers (w/o needing to remove)





Open Research

Reproducible

- Share your analysis code

Open

- Make your data public
- Share your videos!!!

Interpretable

- Use estimation language
- Plot your data

Transparent

- Pre-register your study!
- Differentiate Planned from Exploratory analyses

Conclusion

Benefits

- 1 analysis instead of 4
- Easy way to incorporate complex designs
- Take into account variability and uncertainty from ALL sources (in your data)
  - *Demeanour bias* (important and overlooked)
- Can easily be adapted to theoretical decision (e.g., are people guessing? How does suspiciousness move the threshold for judging something as a 'lie'?)
- Works with small samples!
- Clear inferences

Thank you!

- Questions ???

Matti Vuorre
Tilburg University

Contact

: m.zloteanu@kingston.ac.uk
: @mzloteanu
: mz555.github.io

Kingston University London

OSF

osf.io/kuhzj
For R script