P-value project

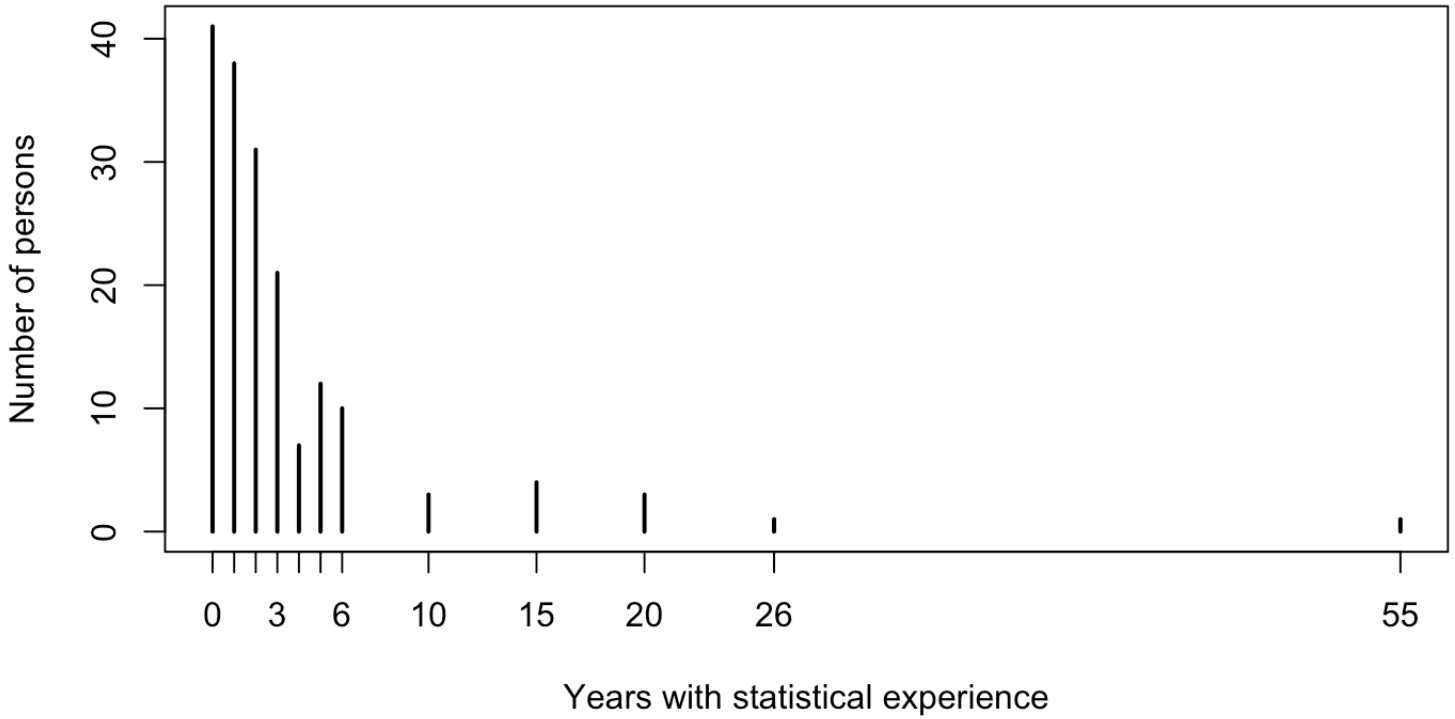
Blockkurs 2018

28.03.2018

Data

- Number of persons interviewed: 172
- Proportion of females: 0.56
- Average proportion of correct answers: 0.67

Statistical experience



3/13

Statistical methods

- Logistic regression (outcome 0 or 1)
- Person ID as random effect to account for repeated measurements (each person gave 10 answers)
- Experimental treatment: figures with or without p-value
- Covariate: statistical experience of person (in log(years+1))
- Interaction treatment x experience

Main result

	Estimate	Std. Error	P-Value
Intercept	0.118	0.274	0.668
Statistical experience [log(year+1)]	1.244	0.237	0.000
Treatment	0.341	0.382	0.372
Treatment x Experience	-0.955	0.314	0.002

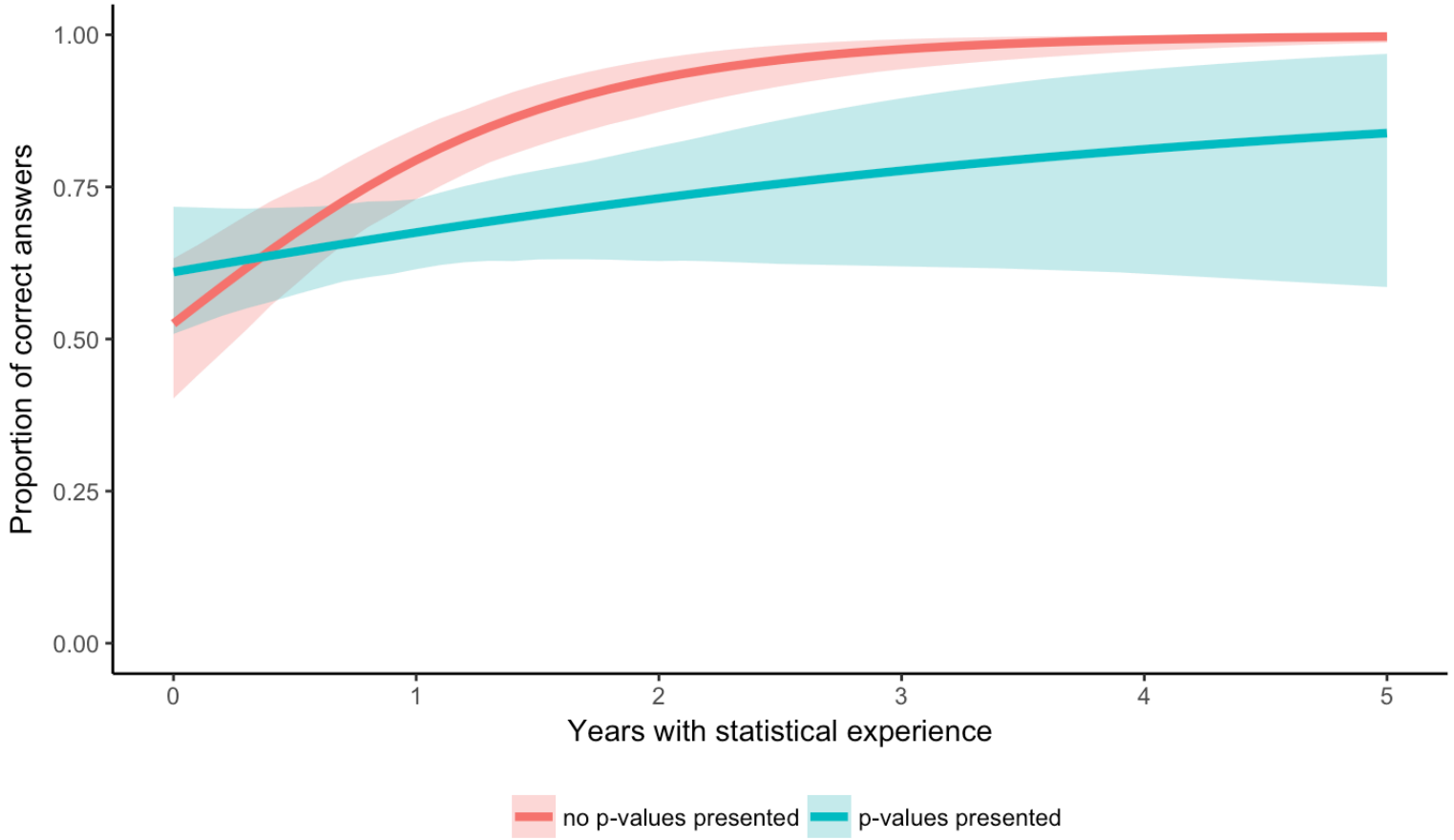
Experienced persons excluded

	Estimate	Std. Error	P-Value
Intercept	0.018	0.294	0.952
Statistical experience [log(year+1)]	1.437	0.300	0.000
Treatment	0.291	0.418	0.485
Treatment x Experience	-0.971	0.404	0.016

- Persons with more more than 9 years of statistical experience excluded.

6/13

Deceived by p-values



Correct 7/13

What affects the answers?

- Only data from figures that shows confidence intervals AND p-values
- Logistic regression
- Outcome variable: 1 if person say left figure is correct, 0 otherwise.
- First predictor: difference in confidence interval length
- Second predictor: difference in p-values
- Person ID as random effect to account for repeated measurements (each person gave 5 answers)

What affects the answers?

If figures with confidence intervals and p-values are presented, the answers are more strongly guided by the difference in p-value than the differences in confidence interval length.

	Estimate	Std. Error	P-Value
Intercept	-0.154	0.114	0.177
Delta conf. intervals	0.907	0.143	0.000
Delta p-values	1.564	0.392	0.000

9/13

Boxplots vs. confidence intervals?

- Only data from figures that do NOT show p-values
- Logistic regression
- Outcome variable: 1 answer is correct, 0 answer is not correct
- Predictor: 1 Figure shows a box-plot, 0 otherwise.
- Person ID as random effect to account for repeated measurements (each person gave 5 answers)

Boxplots vs. confidence intervals?

	Estimate	Std. Error	P-Value
Intercept	1.711	0.316	0.000
Boxplot	-0.174	0.197	0.377

The proportion of correct answers was slightly higher for figures with confidence intervals (0.85) than for figures with box-plots (0.82). But the difference was not significant.

11/13

Effect of observer and gender

	Estimate	Std. Error	P-Value
Intercept	0.268	0.321	0.403
Statistical experience [log(year+1)]	1.247	0.238	0.000
Treatment	0.341	0.381	0.372
Observer	-0.113	0.243	0.643
Gender	-0.203	0.246	0.409
Treatment x Experience	-0.971	0.314	0.002

12/13

Learning effect

	Estimate	Std. Error	P-Value
Intercept	0.015	0.298	0.959
Statistical experience [log(year+1)]	1.244	0.237	0.000
Treatment	0.341	0.382	0.372
Learn-effect	0.019	0.021	0.376
Treatment x Experience	-0.955	0.314	0.002

13/13