

# Can Students Reproduce?

Ian E. Holliday, Aston University, UK  
&  
Patrick Wessa, University of Leuven, B

# Can Students Reproduce?

- Are they able to reproduce?



# Can Students Reproduce?

- Are they able to reproduce?
- How much effort does it take them to reproduce?





# Can Students Reproduce?

- Are they able to reproduce?
- How much effort does it take them to reproduce?
- Or do they simply trust that reproduction is not necessary?



Btw, we are talking about reproducing **statistical results** ....

# The *Science* Paper Experiment

- Assignment: reproduce and review the correlation matrix and both regression models in

Allison, T. and Cicchetti, DV (1976), Sleep in mammals: ecological and constitutional correlates, ***Science***, 194 (4266): 732-734

- N = 263
- Assignment was graded by instructors (accounts for 15% of final grade)
- Article and dataset (raw data) was made available
- Deadline: 2 weeks

# A&C Sleep Study

Variables	Correlation coefficients								
	SWS	PS	L	$w_b$	$w_{br}$	$t_g$	P	S	D
SWS	1.000	.582	-.377	-.712	-.679	-.589	-.369	-.580	-.542
PS		1.000	-.342	-.370	-.435	-.651	-.536	-.591	-.686
L			1.000	.685	.777	.682	.018	.518	.226
$w_b$				1.000	.945	.692	.253	.662	.432
$w_{br}$					1.000	.781	.192	.624	.377
$t_g$						1.000	.158	.588	.363
P							1.000	.680	.930
S								1.000	.819
D									1.000

Correlations between Sleep Variables (SWS, PS), Constitutional variables (e.g. body weight,  $w_b$ ) and ecological variables (e.g. Danger of predation, D)

Factor analysis revealed 'Danger' and 'Size' dominated – the constitutional and ecological groups of variables.

Main predictors of SWS were body size ( $w_b$ ) and predation risk (D).

$$\text{SWS} = 11.7 - (1.82 \pm 0.37) \log w_b - (0.799 \pm 0.33) D \quad (1)$$

$$\text{Log PS} = 1.07 - (0.109 \pm 0.022) D - (0.300 \pm 0.068) \log t_g \quad (2)$$

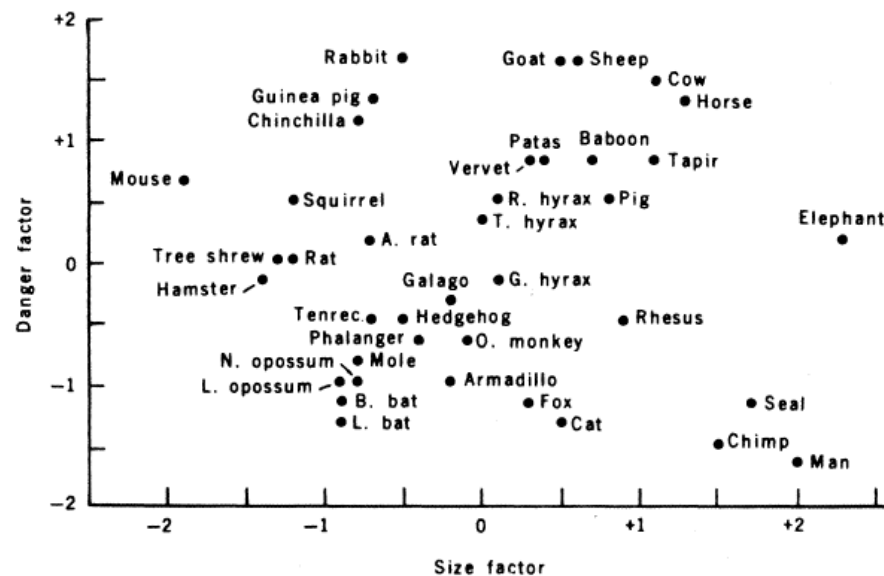


Fig. 1. Factor scores ( $\bar{X} = 0$ , S.D. = 1) for 39 species used in the factor analysis of Table 2. Smaller species appear to the left, larger species to the right. Species heavily preyed upon appear at the top, those less subject to predation at the bottom. Some clusters are perceptible, for example domesticated farm animals (goat, sheep, cow, horse), small surface-dwelling herbivores (rabbit, guinea pig, chinchilla), and African savanna primates (vervet, patas, baboon). The

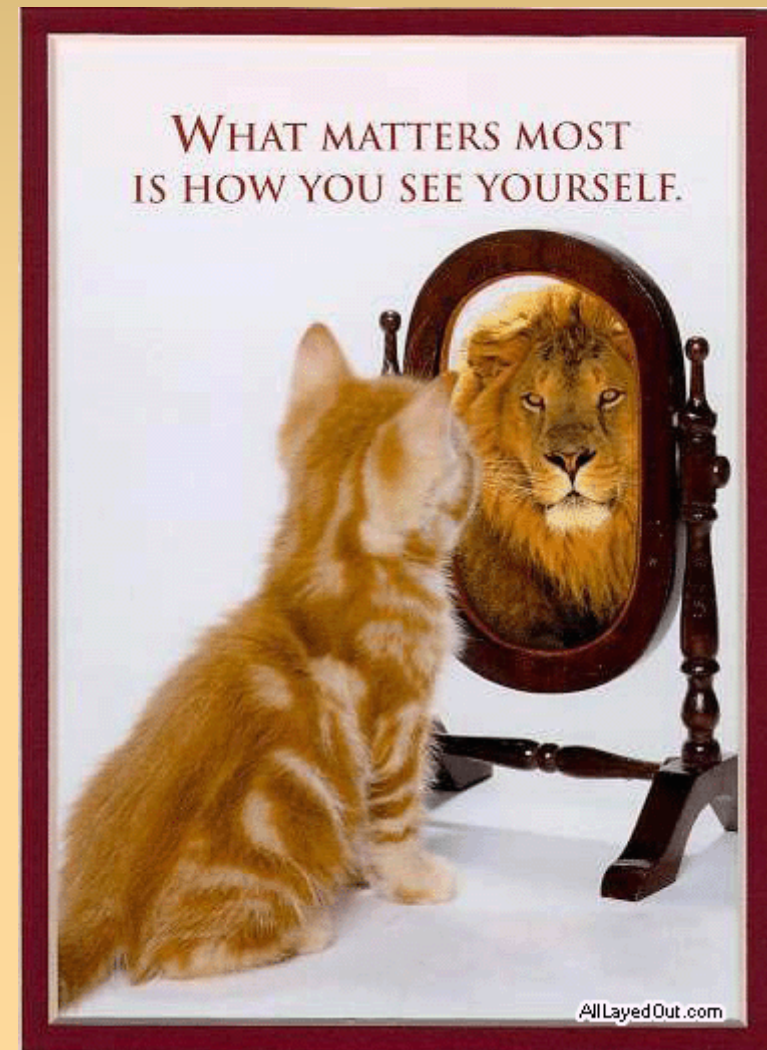
# The *Science* Paper Experiment

- Article has 145 citations (Google Scholar)
- Both authors from Yale Univ.
- Easy to read, even for economists :-)
- Simple statistical methods
- Permission to use the raw data
- Assignment at the end of the semester (chapters on correlation and regression were completed)
- Students “blogged” computations (we were able to verify the students' findings)
- **We assert that there were NO errors in the statistical analysis of the paper**



# At the start of the Experiment...

- ...most students were confident that they would be easily able to reproduce the statistical analyses and boost their final grade!



# Illustration of differences

	SWS	PS	L	Wb	Wbr	tg	P	S	D
SWS	1	0.518	-0.372	-0.394	-0.387	-0.606	-0.353	-0.58	-0.535
PS	0.518	1	-0.268	-0.075	-0.074	-0.409	-0.398	-0.504	-0.572
L	-0.372	-0.268	1	0.47	0.629	0.646	-0.17	0.316	0.015
Wb	-0.394	-0.075	0.47	1	0.956	0.714	0.096	0.406	0.259
Wbr	-0.387	-0.074	0.629	0.956	1	0.734	-0.015	0.323	0.151
tg	-0.606	-0.409	0.646	0.714	0.734	1	0.091	0.573	0.306
P	-0.353	-0.398	-0.17	0.096	-0.015	0.091	1	0.626	0.927
S	-0.58	-0.504	0.316	0.406	0.323	0.573	0.626	1	0.79
D	-0.535	-0.572	0.015	0.259	0.151	0.306	0.927	0.79	1

	SWS	PS	L	Wb	Wbr	tg	P	S	D
SWS	1	0.582	-0.378	-0.713	-0.679	-0.601	-0.369	-0.58	-0.542
PS	0.582	1	-0.346	-0.372	-0.438	-0.646	-0.534	-0.591	-0.684
L	-0.378	-0.346	1	0.683	0.777	0.679	0.019	0.519	0.227
Wb	-0.713	-0.372	0.683	1	0.945	0.695	0.252	0.662	0.431
Wbr	-0.679	-0.438	0.777	0.945	1	0.777	0.192	0.624	0.377
tg	-0.601	-0.646	0.679	0.695	0.777	1	0.148	0.58	0.353
P	-0.369	-0.534	0.019	0.252	0.192	0.148	1	0.68	0.93
S	-0.58	-0.591	0.519	0.662	0.624	0.58	0.68	1	0.819
D	-0.542	-0.684	0.227	0.431	0.377	0.353	0.93	0.819	1

ecologic  
ow wav  
ation ti

Variables	Correlation coefficients								
	SWS	PS	L	w <sub>b</sub>	w <sub>br</sub>	t <sub>g</sub>	P	S	D
SWS	1.000	.582	-.377	-.712	-.679	-.589	-.369	-.580	-.542
PS		1.000	-.342	-.370	-.435	-.651	-.536	-.591	-.686
L			1.000	.685	.777	.682	.018	.518	.226
w <sub>b</sub>				1.000	.945	.692	.253	.662	.432
w <sub>br</sub>					1.000	.781	.192	.624	.377
t <sub>g</sub>						1.000	.158	.588	.363
P							1.000	.680	.930
S								1.000	.819
D									1.000

SCIENCE, VOL. 194

# Results of the *Science Paper* Experiment

Able to reproduce?	Correlation matrix		Regression 1		Regression 2	
Yes	61	23.19%	55	20.91%	59	22.43%
No	109	41.44%	117	44.49%	109	41.44%
No answer	93	35.36%	91	34.60%	95	36.12%
Did authors do the analysis correctly?	Correlation matrix		Regression 1		Regression 2	
Yes	25	9.51%	44	16.73%	40	15.21%
No	28	10.65%	42	15.97%	43	16.35%
No answer	210	79.85%	177	67.30%	180	68.44%

# Results of the *Science Paper* Experiment

Authors <i>bona fide</i> ?		
Yes	67	25.48%
No	83	31.56%
Not sure	113	42.97%

Should article be published?		
Yes, w/o changes.	50	19.01%
Minor changes req.	62	23.57%
Major revision req.	59	22.43%
No, reject the paper	8	3.04%
Not sure	84	31.94%

Did you collaborate to review the paper?			Did you ask for expert advice?		
Yes	112	42.59%	Yes	22	8.37%
No	151	57.41%	No	241	91.63%



# Results of the *Science Paper* Experiment

Did you consult any sources?		
Internet (Search Engine)	51	19.39%
Internet (other)	23	8.75%
Handbook(s)	37	14.07%
Articles & other	30	11.41%
None	122	46.39%

How much time did you need to reproduce and review the article?		
< 1 hr	128	48.67%
1-3 hrs	55	20.91%
3-5 hrs	56	21.29%
5-8 hrs	28	10.65%
> 8 hrs	14	5.32%

# A few written remarks from students after the experiment...

- Always make sure that a trained statistician reviews the article and checks whether there was no fraud
  - Authors should not exclude observations to manipulate the statistical results
  - I would require all authors to use Reproducible Computing
  - The data/analysis should not be manipulated. One should always present the results without distortions
  - (many similar remarks)...
  - Articles should always be peer reviewed before publication!
- Etc.

# Conclusions

- Students who successfully reproduced the statistical results needed:
  - > 5 hrs of time
  - Additional, external sources (or expert advice)
- Most students did *not* trust the results
- 1 out of 4 students believed that the authors were *bona fide* - almost 32% thought the authors committed fraud
- 19% of students recommended publication (w/o modifications)