University of Glasgow | Institute of Neuroscience & Psychology

CENTRE FOR COGNITIVE NEUROIMAGING

Interpreting interactions

Guillaume Rousselet
@robustgar
https://garstats.wordpress.com

"The strategy of run-a-crappy-study, get p less than .05, come up with a cute story based on evolutionary psychology, and PROFIT . . . well, it does not work anymore. OK, maybe it still can work if your goal is to get published in PPNAS, get tenure, give Ted talks, and make boatloads of money in speaking fees. But it will not work in the real sense, the important sense of learning about the world."

Andrew Gelman, 2018, The Failure of Null Hypothesis Significance Testing When Studying Incremental Changes, and What to Do About It. *Personality and Social Psychology Bulletin*

References

- Loftus, Geffrey R. 1978. On interpretation of interactions. Memory & Cognition 6(3). 312-319.

- Wagenmakers, Eric-Jan, Angelos-Miltiadis Krypotos, Amy H. Criss and Geoff Iverson. 2012. On the interpretation of removable interactions: A survey of the field 33 years after Loftus. Memory & Cognition 40(2). 145-160.

- Wagenmakers, E.-J. (2015) A quartet of interactions. Cortex, 73, 334–335.

- Vanhove, J. (2019) Interactions in logistic regression models. Blog post:

https://janhove.github.io/analysis/2019/08/07/interactions-logistic

Common statistical myths & fallacies

SCIENCE FORUM

Ten common statistical mistakes to watch out for when writing or reviewing a manuscript

eLIFE
elifesciences.org

Abstract Inspired by broader efforts to make the conclusions of scientific research more robust, we have compiled a list of some of the most common statistical mistakes that appear in the scientific literature. The mistakes have their origins in ineffective experimental designs, inappropriate analyses and/or flawed reasoning. We provide advice on how authors, reviewers and readers can identify and resolve these mistakes and, we hope, avoid them in the future.

TAMAR R MAKIN* AND JEAN-JACQUES ORBAN DE XIVRY

Common statistical myths & fallacies

datamethods

https://discourse.datamethods.org

Reference Collection to push back against “Common Statistical Myths”

data analysis | journal | teaching

ADAlthousePhD 40 Nov '19

Note: This topic is a wiki, meaning that this main body of the topic can be edited by others. Use the *Reply* button only to post questions or comments about material contained in the body, or to suggest new statistical myths you’d like to see someone write about.

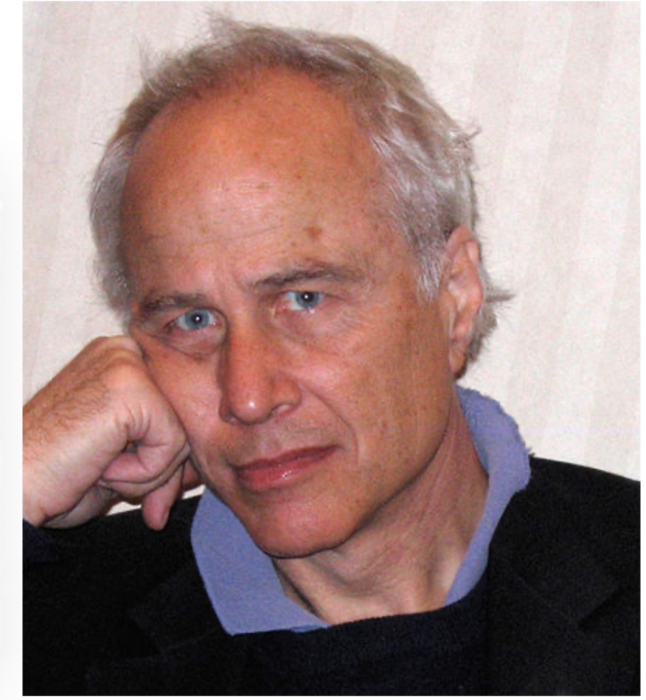
Jun 2019

1 / 34
Jun 2019

Memory & Cognition
1978, Vol. 6 (3), 312-319

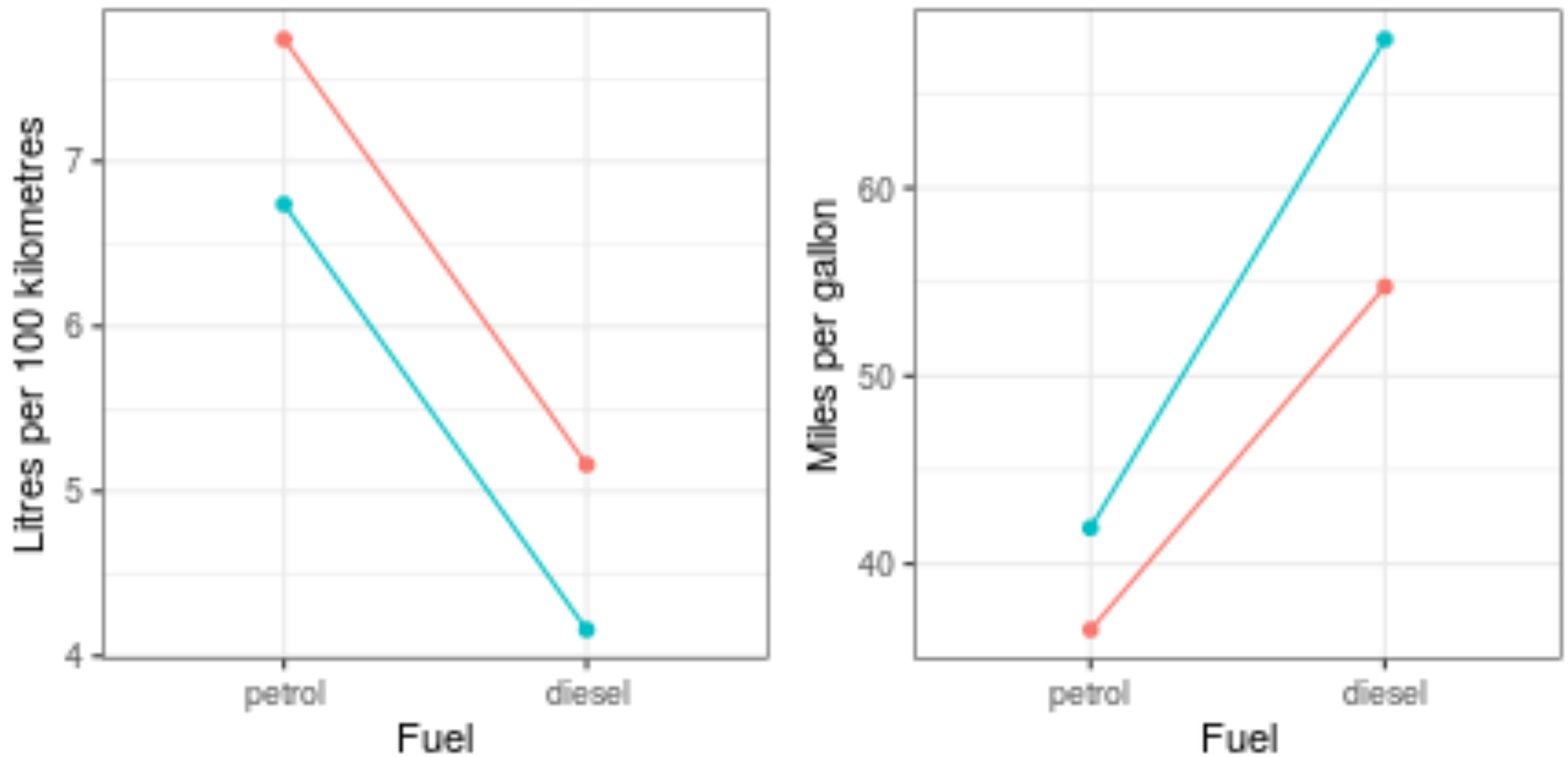
On interpretation of interactions

GEOFFREY R. LOFTUS
University of Washington, Seattle, Washington 98195



"certain types of interactions make sense only if a particular scale is assumed"

[1] A mapping problem

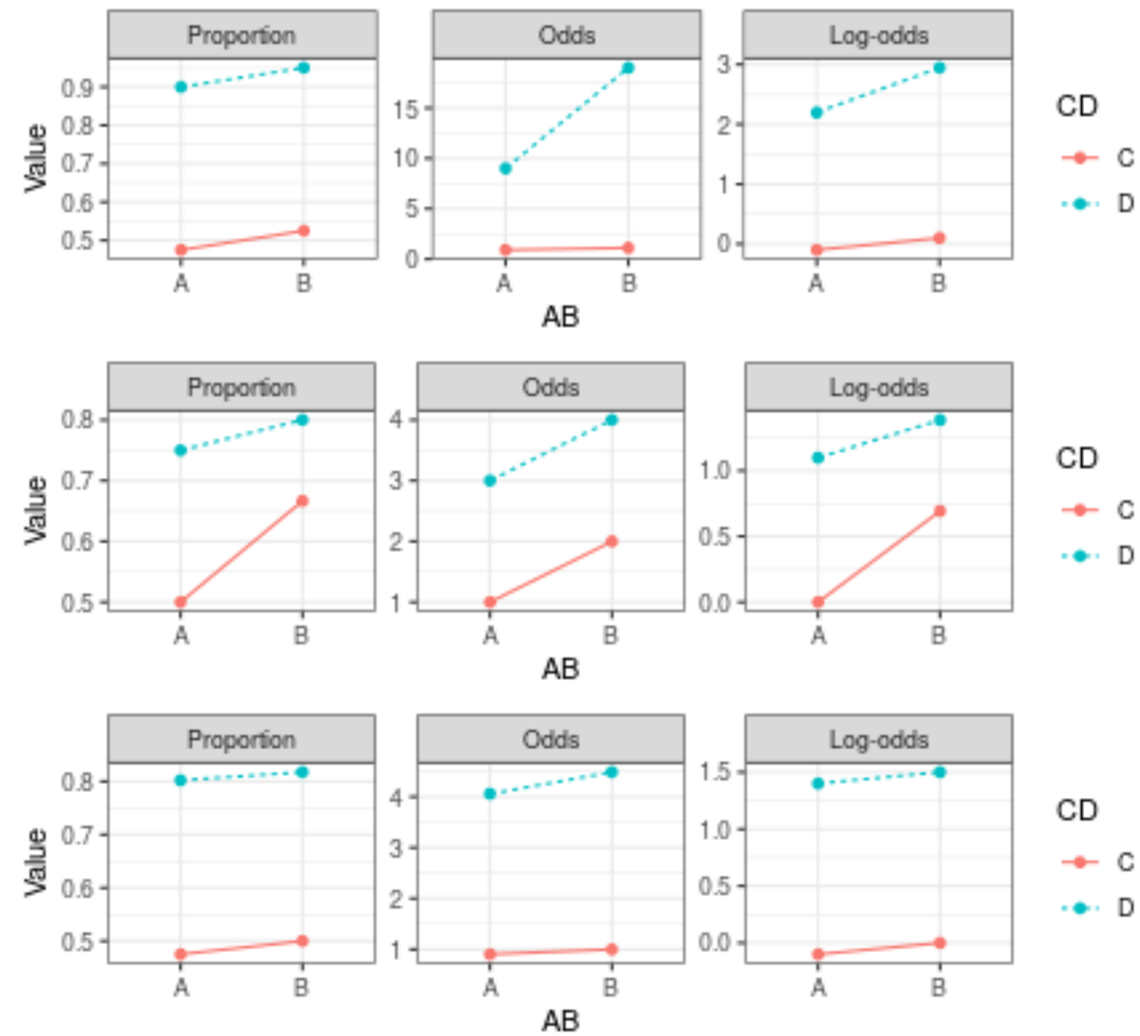


Fuel efficiency of blue cars and red cars: interaction?

https://janhove.github.io/analysis/2019/08/07/interactions-logistic

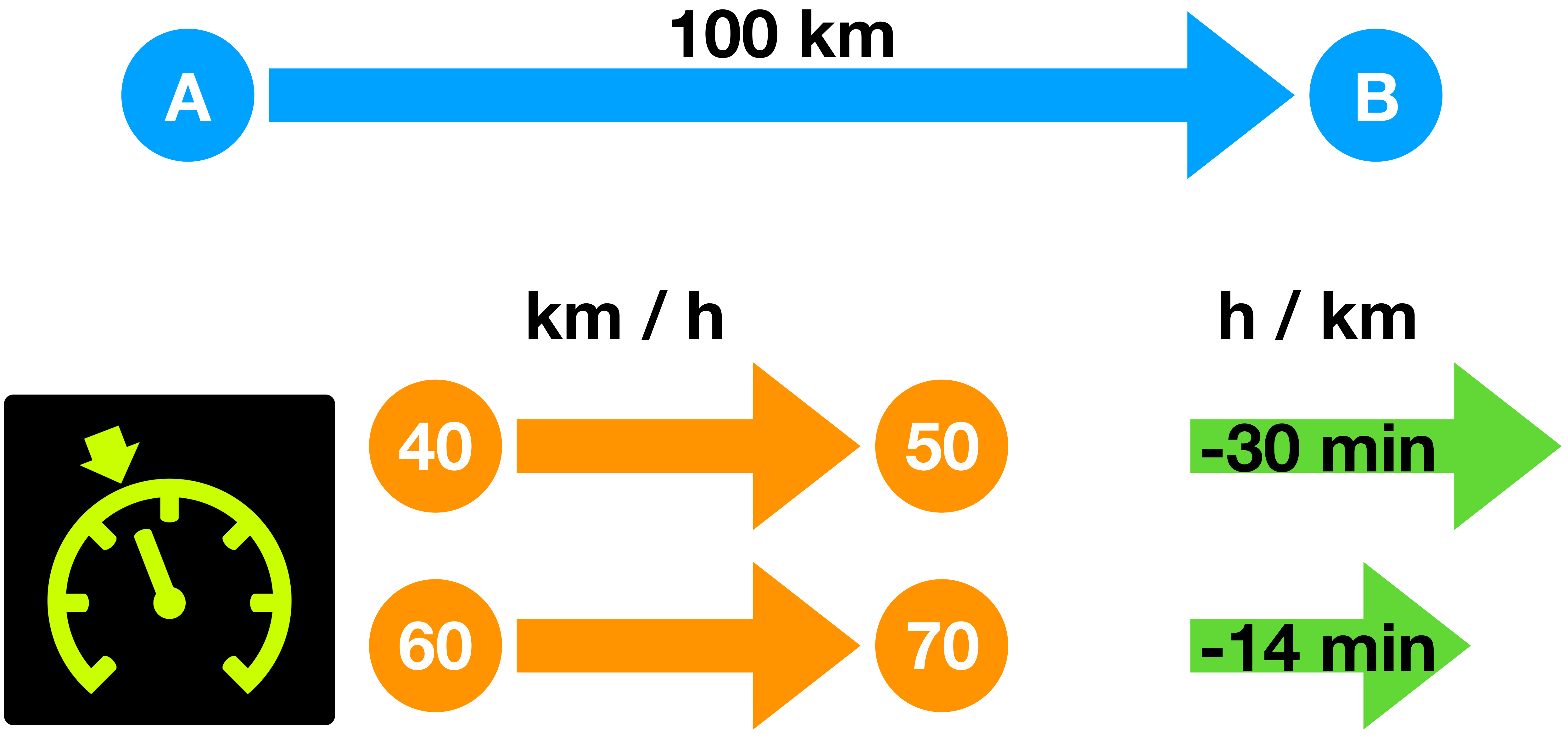
[1] A mapping problem

“The coefficients in logistic models are estimated on the log-odds scale, but such models are more easily interpreted when the coefficients or its predictions are converted to odds or to proportions. Both the exponential and the logistic function are nonlinear, so that you end up with the same problem as above: Whether or not you observe an interaction may depend on how you express the outcome variable.”



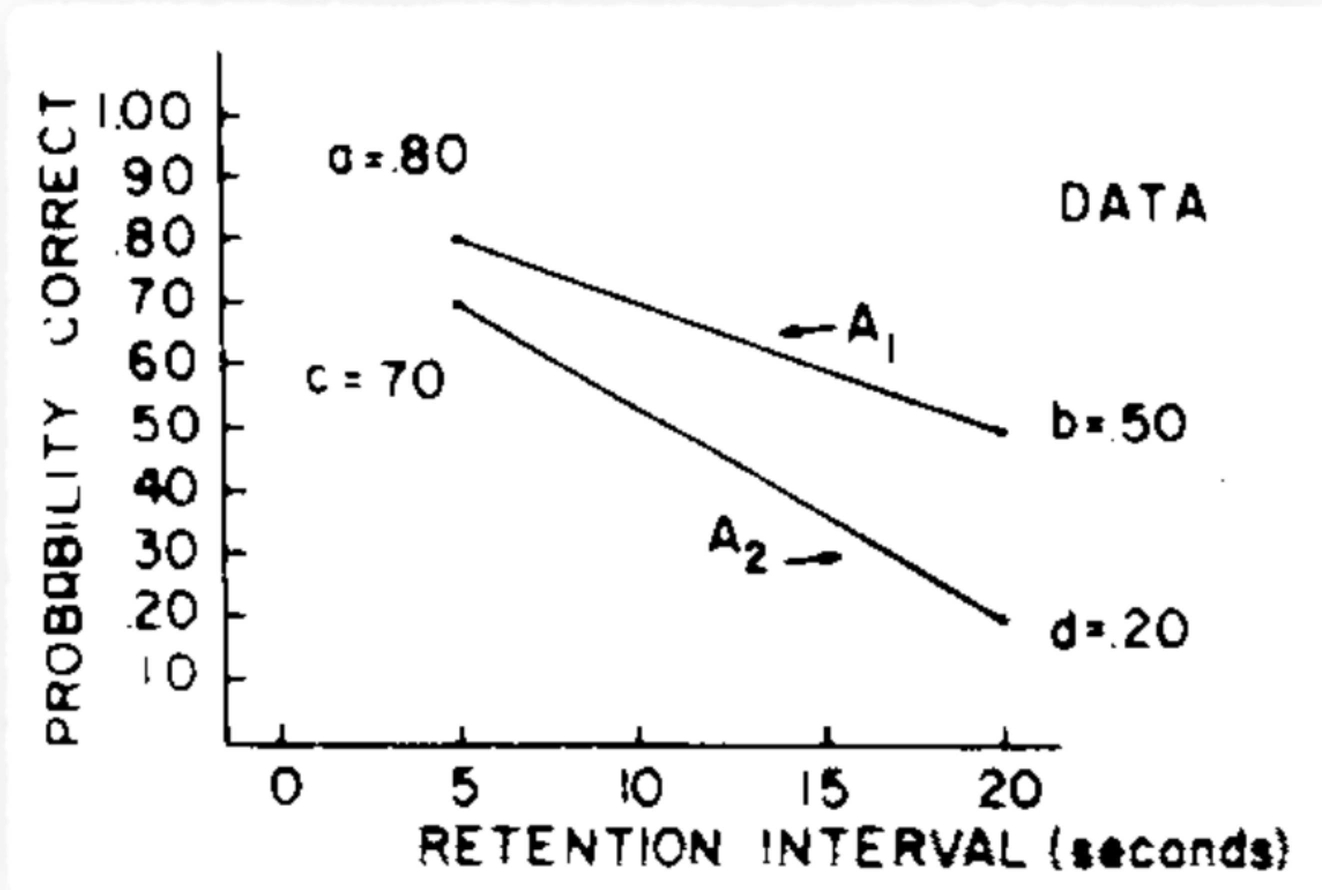
https://janhove.github.io/analysis/2019/08/07/interactions-logistic

How much faster?



Wagenmakers et al. 2012

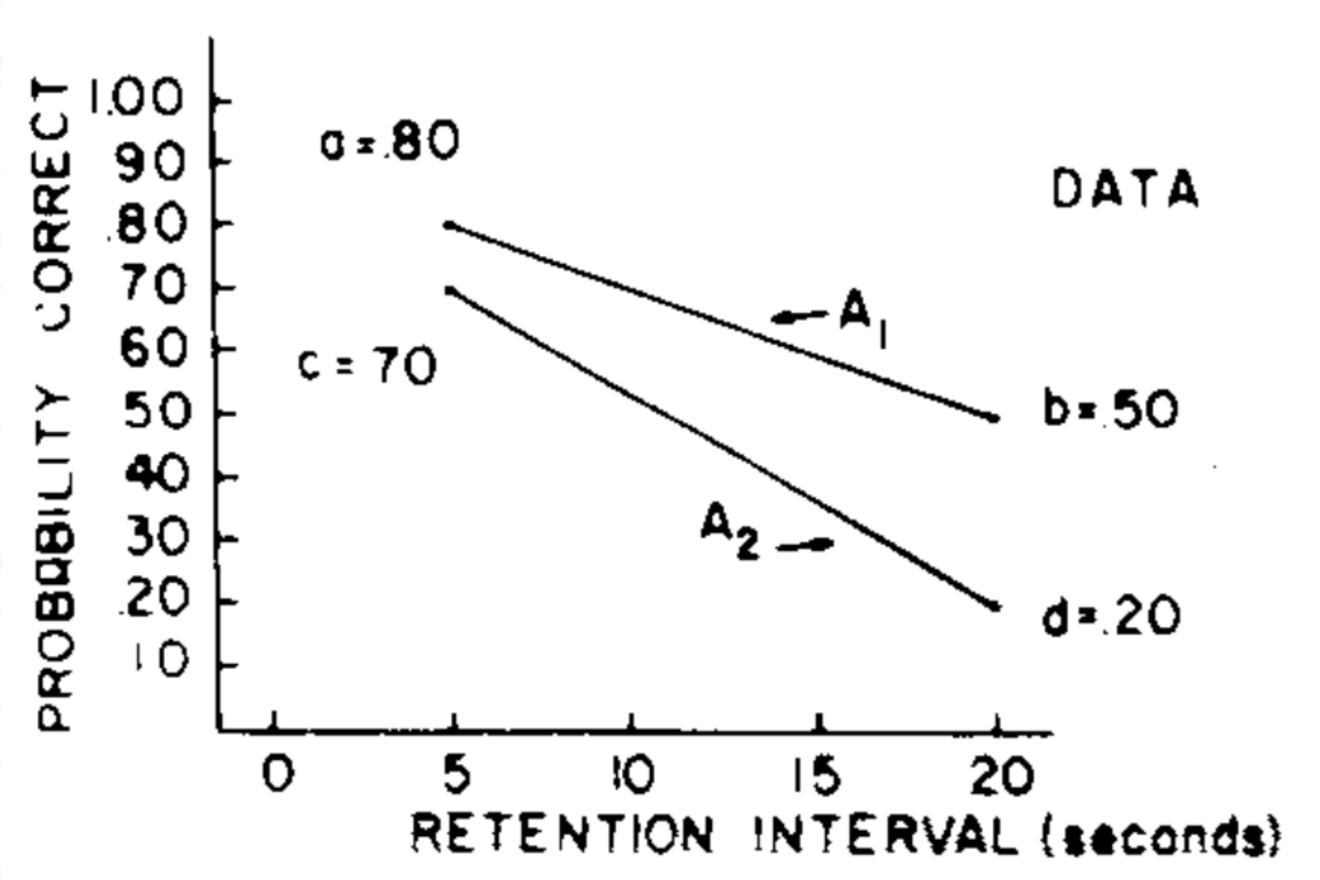
[2] Loftus 1978



Simple main effects + interaction are significant:

"Condition A1 leads to better overall memory performance than does condition A2 and overall memory performance decreases over retention interval."

"Forgetting is faster in A2 than in A1."

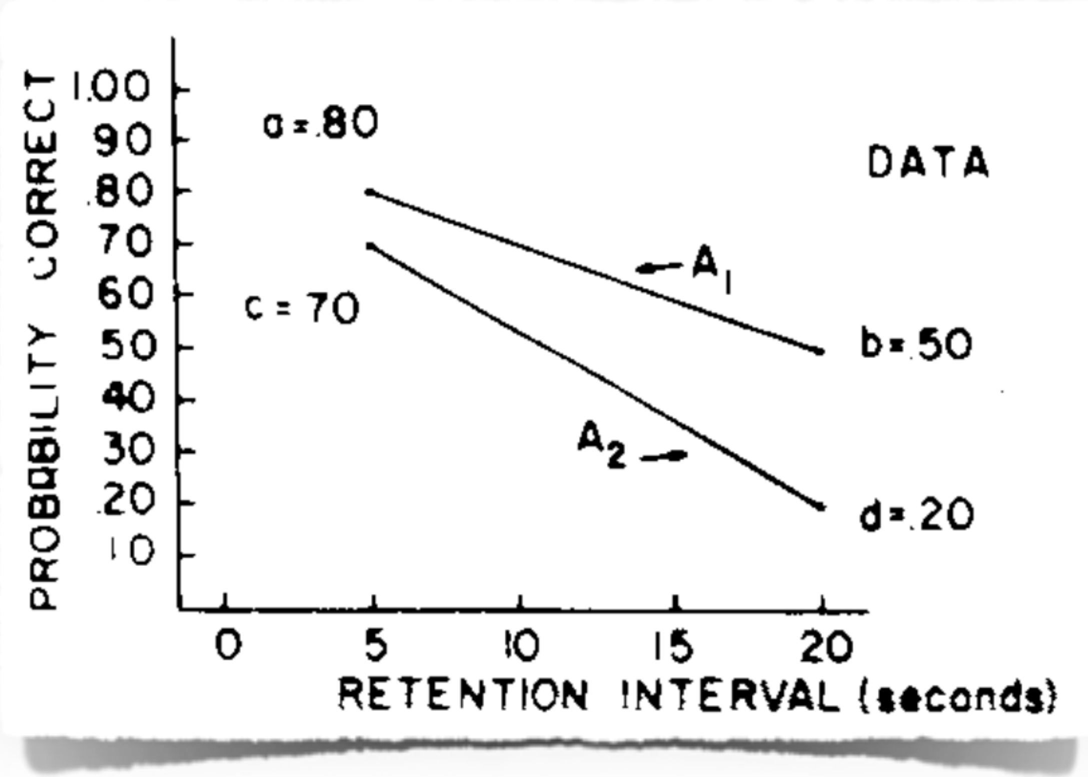


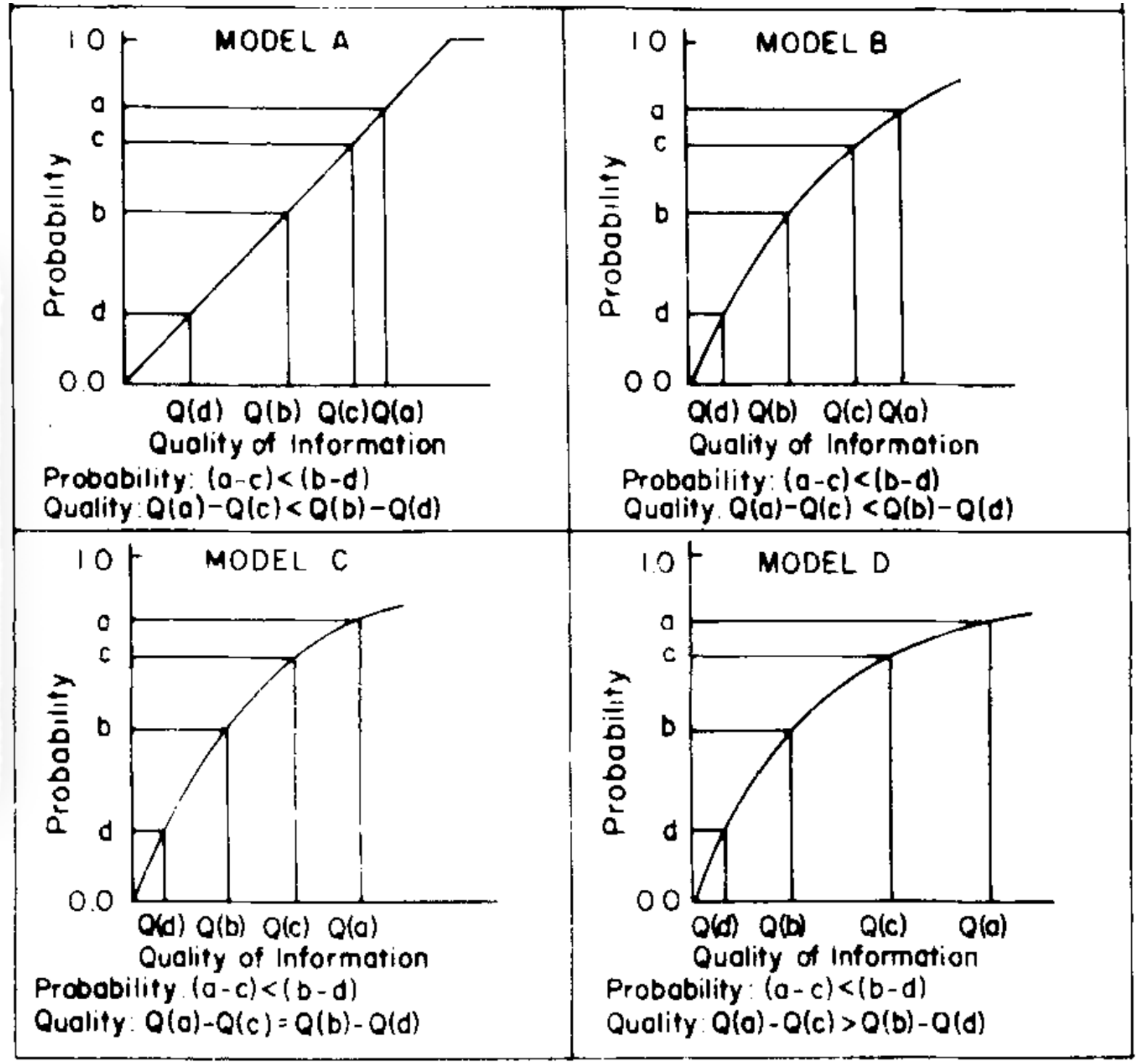
Assumptions underlying the conclusions:

- (1) correct response based on *stored* information *about* the stimulus
- (2) greater information *quality* translates into higher proportion correct.

- Conclusion 1: information quality is greater A1 than in A2
- Conclusion 2: information quality declines with longer retention
- Conclusion 3: quality decline is faster in A2 than A1

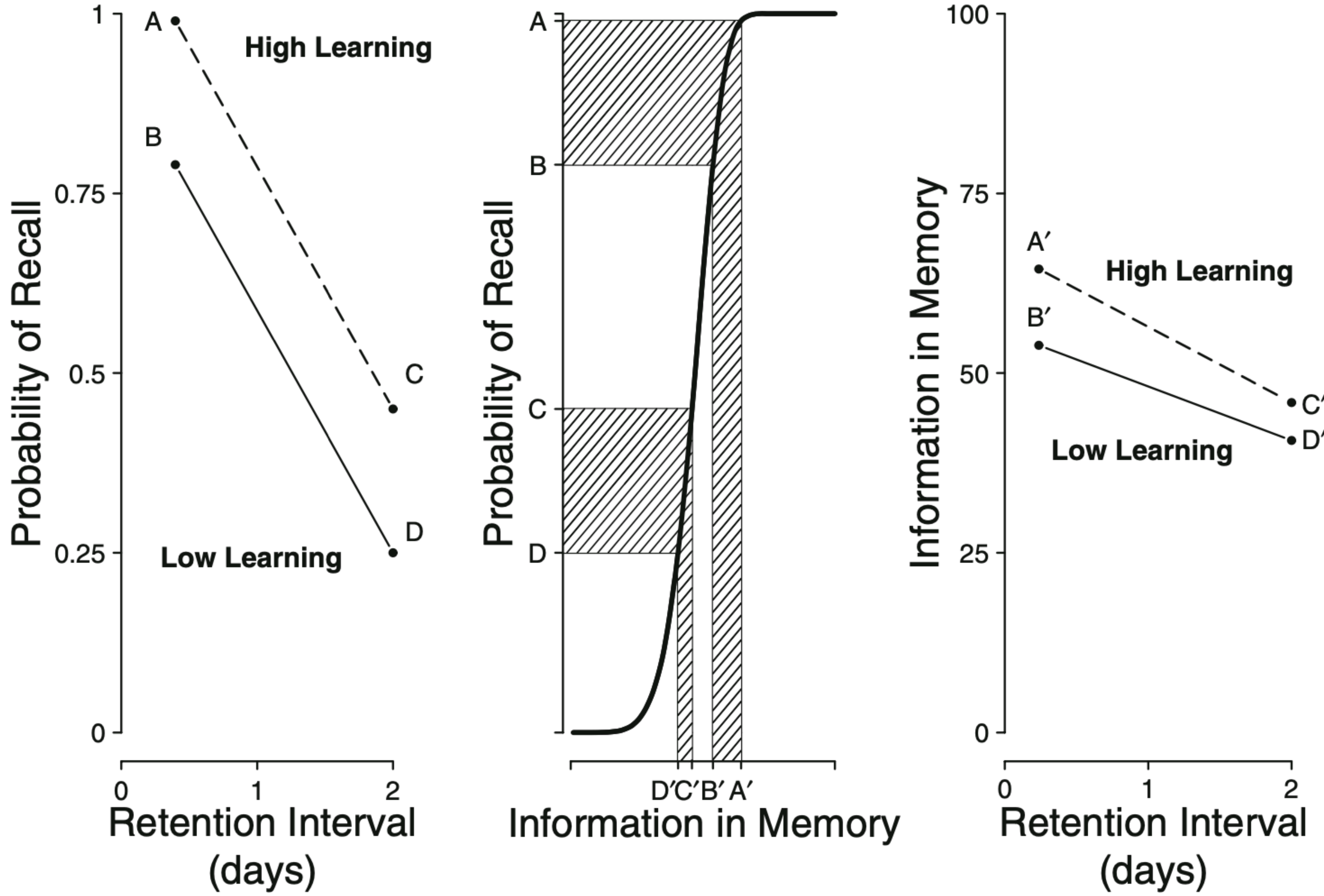
PROBLEM:
"Conclusion 3 can be made only within the context of a more specific model than the one described above."





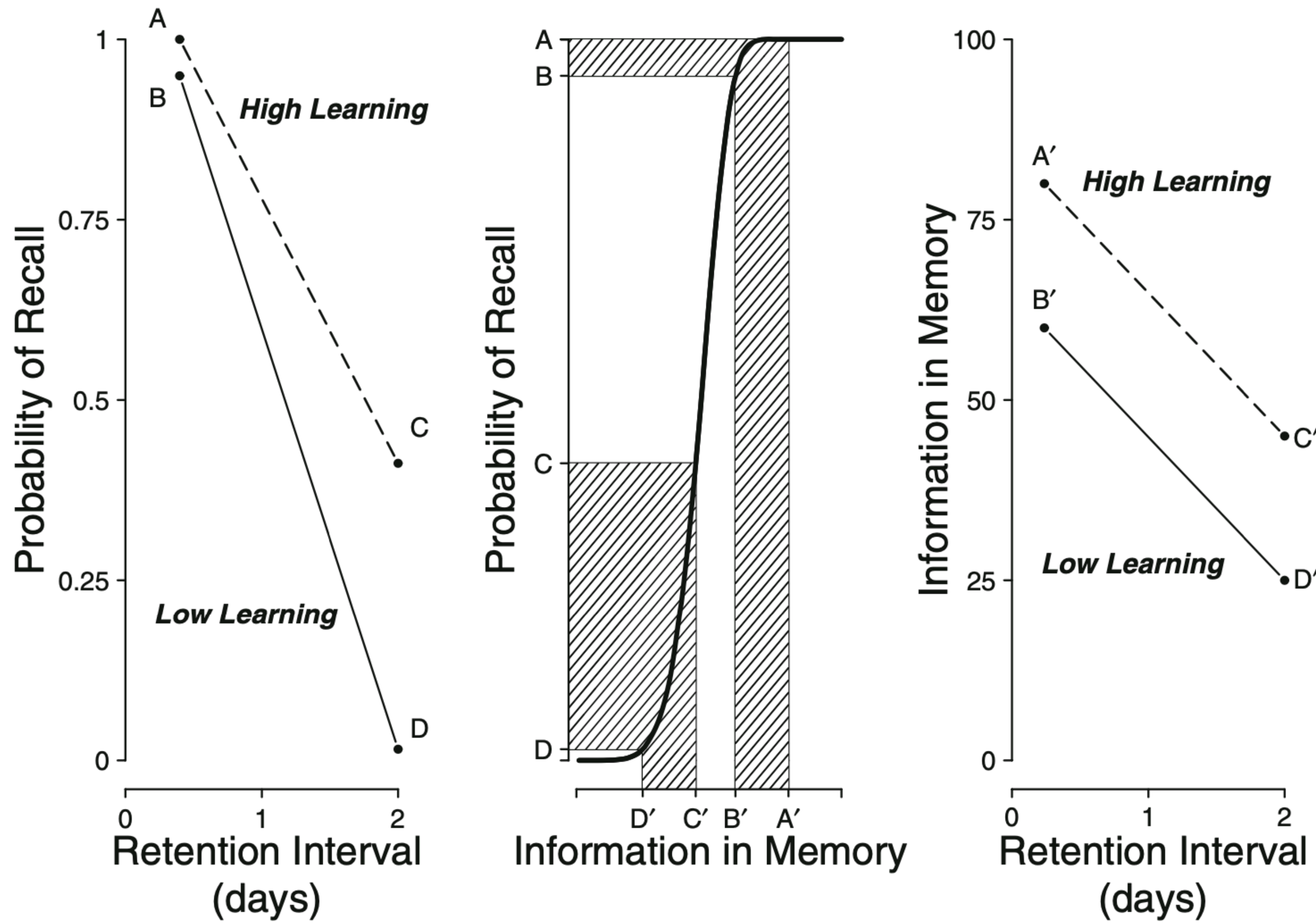
“The point of this example is to illustrate that, when a negatively accelerated function maps some theoretical component-in this case, quality-onto response probability, the sort of interaction depicted in the top panel of Figure 2 is uninterpretable. That is to say, one cannot tell whether the interaction will be the same, will be transformed away, or will reverse itself in terms of the theoretical component. Which of these three outcomes will obtain depends entirely on the exact quantitative form of the mapping function.”

Additive effects on the probability of recall correspond to interaction effects on information in memory



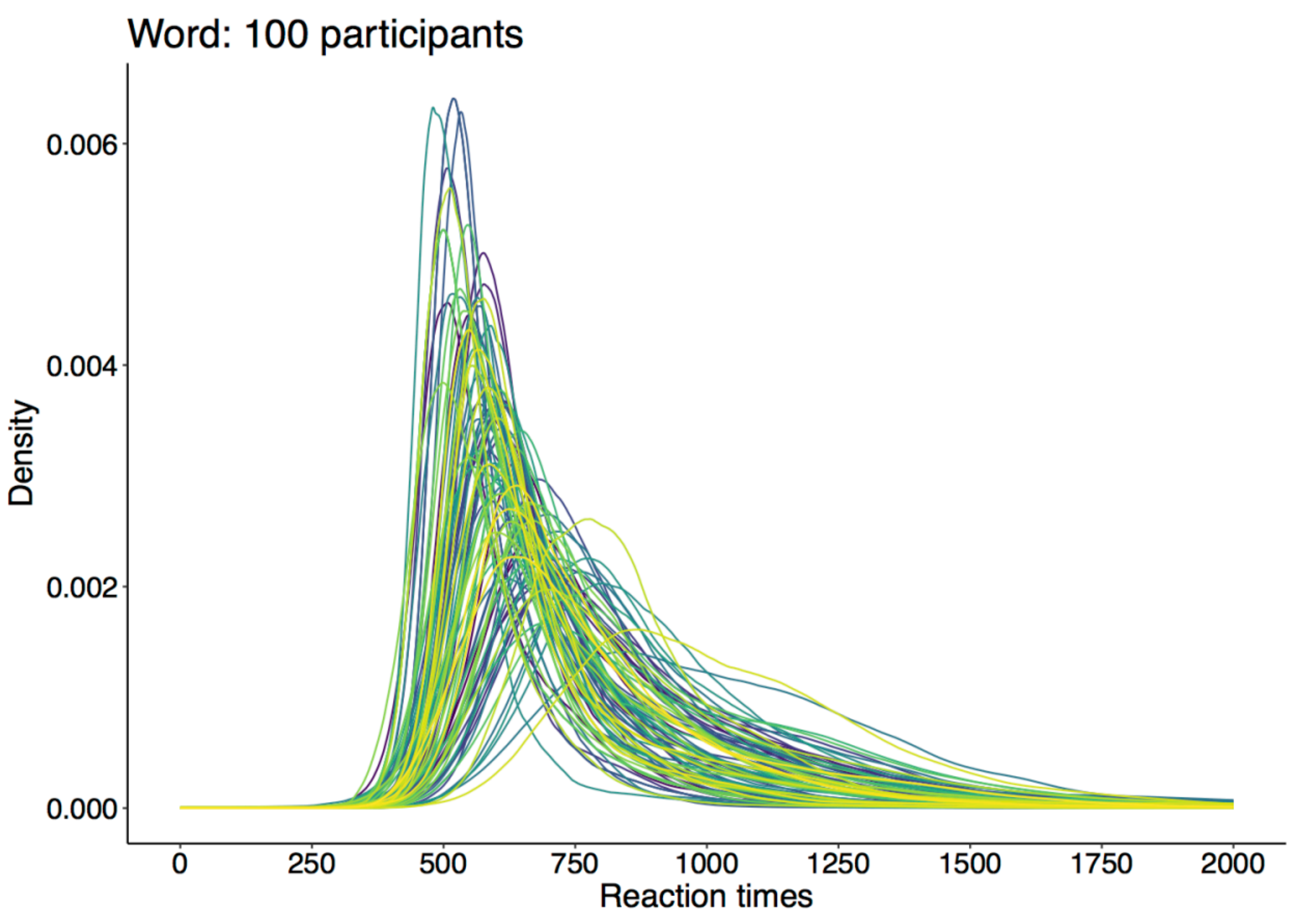
Wagenmakers et al. 2012

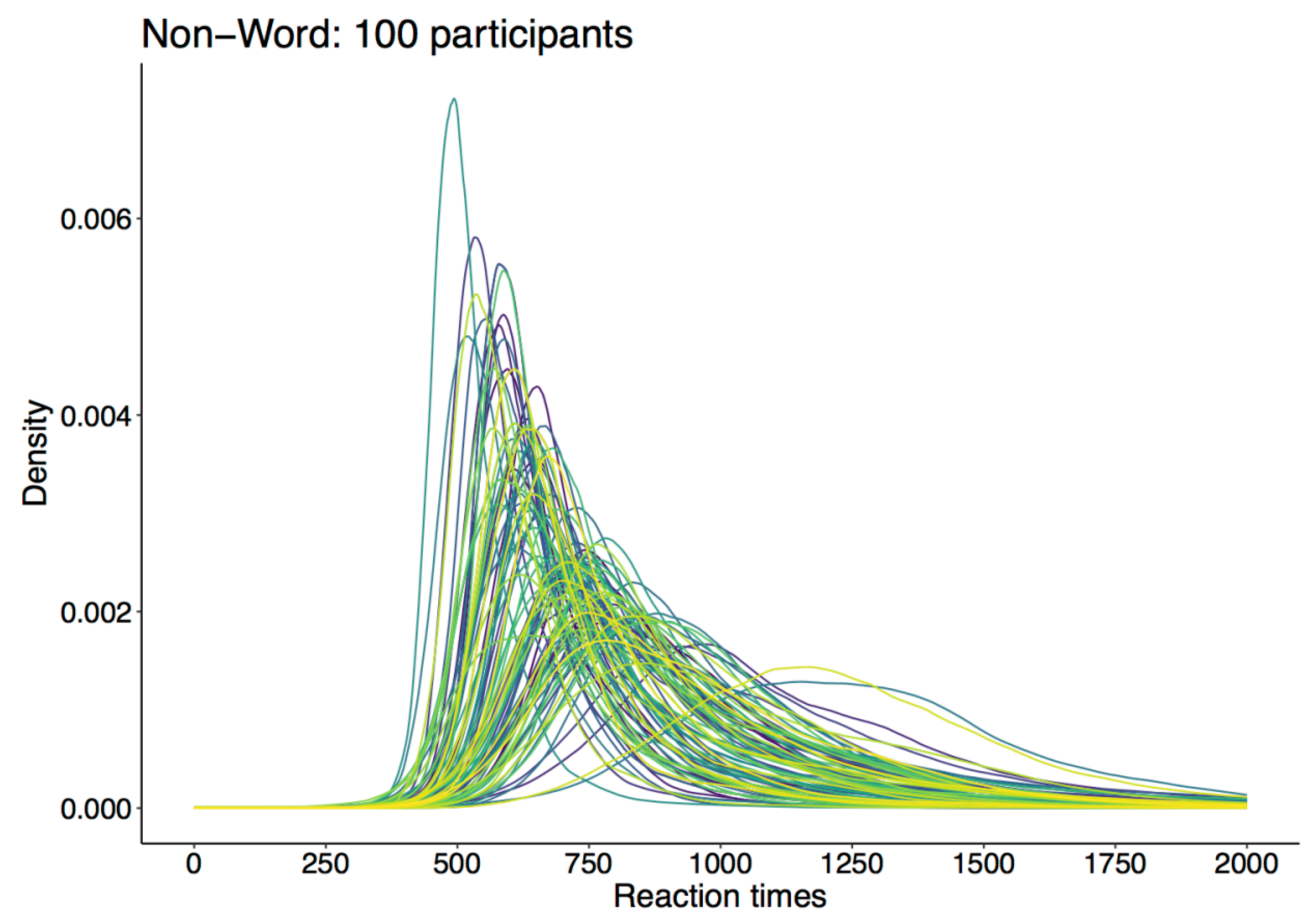
Interaction effects on the probability of recall correspond to additive effects on information in memory



Wagenmakers et al. 2012

[3] Reaction times

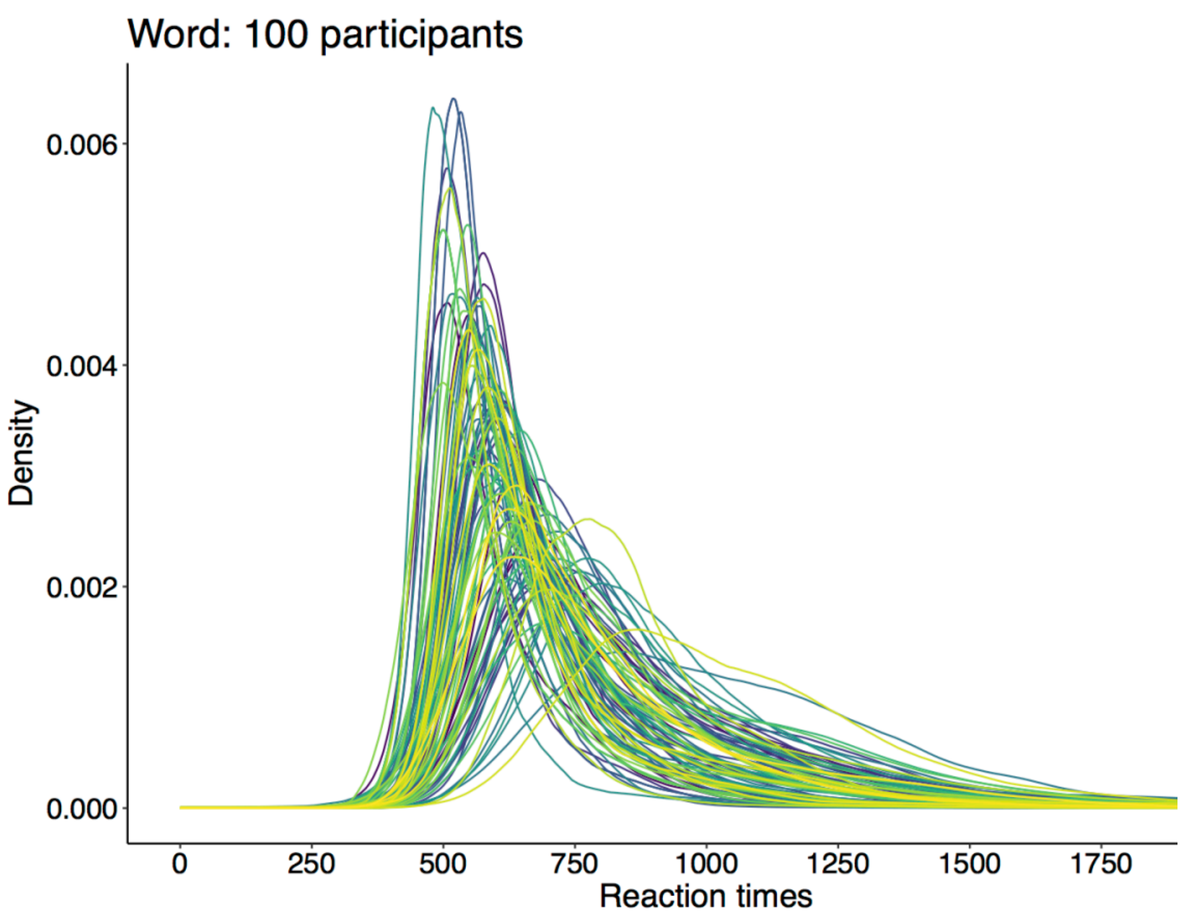


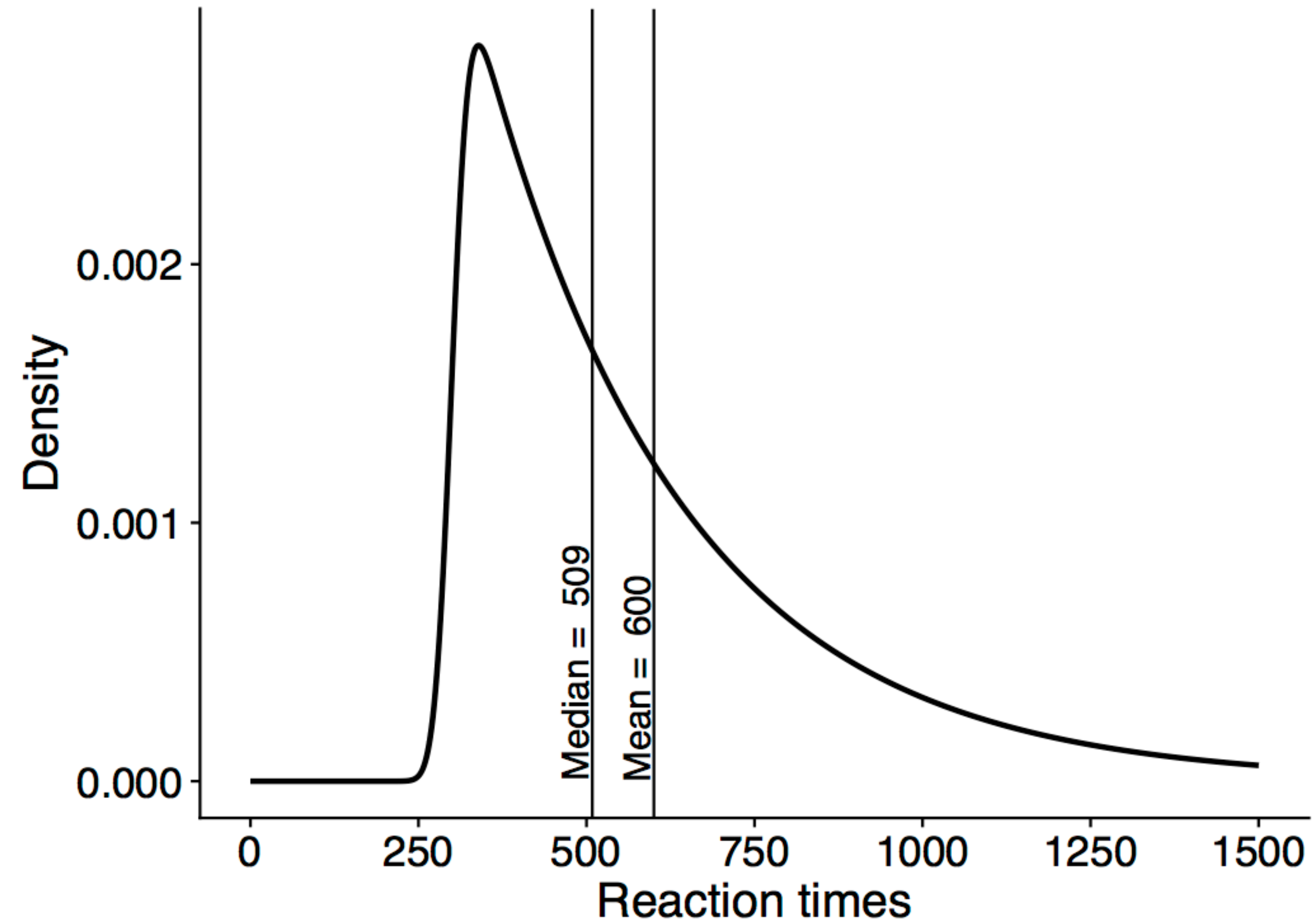


- **processing speed interpretation**
- **distribution transformations**
- **multiple methods / scales: eye tracking, manual responses, EEG, LFP, single units…**
- **same mapping for all parts?**

https://garstats.wordpress.com/2018/04/25/rtbias4/

[3] Reaction times

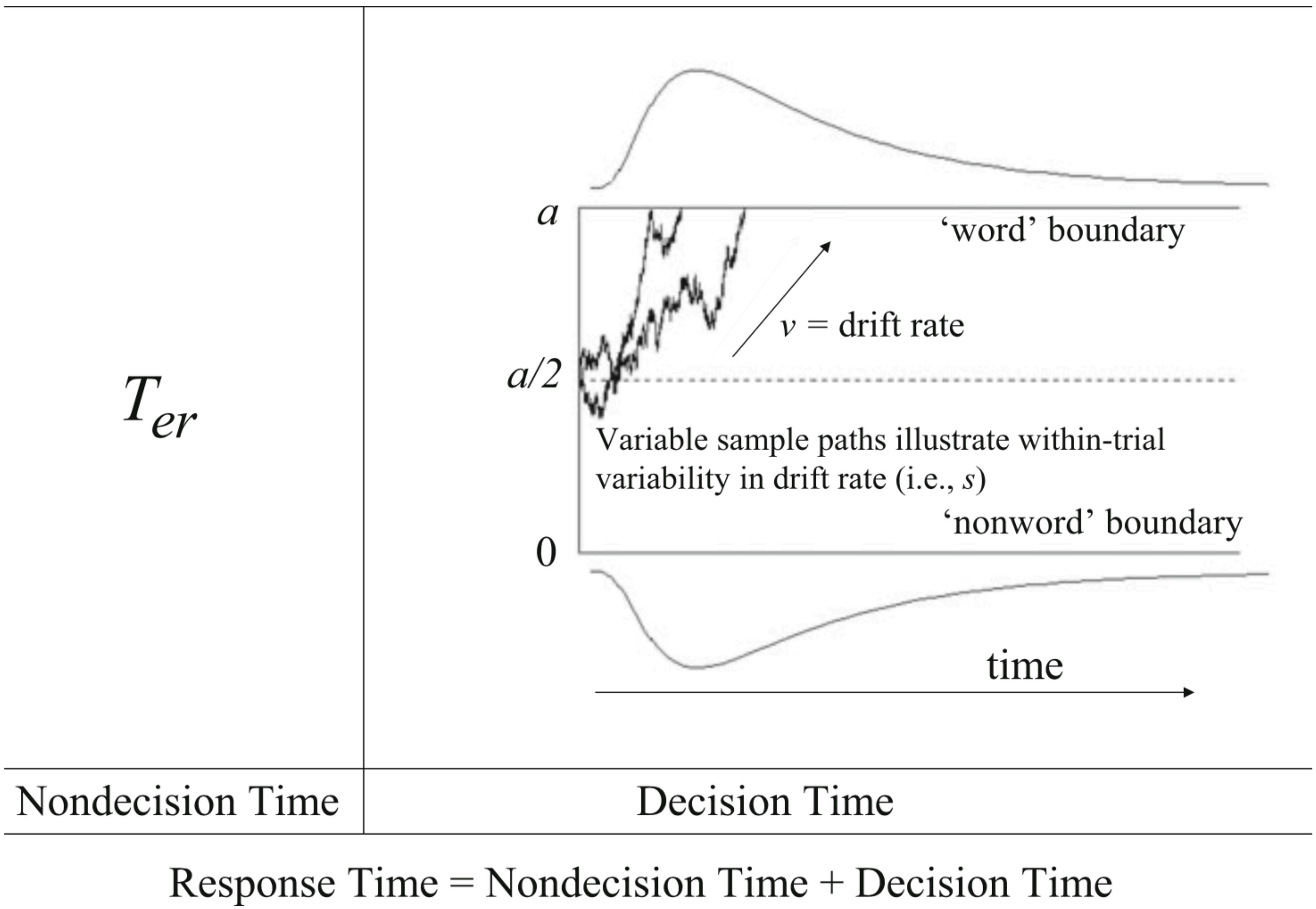




- **processing speed interpretation**
- **distribution transformations**
- **multiple methods / scales: eye tracking, manual responses, EEG, LFP, single units…**
- **same mapping for all parts?**

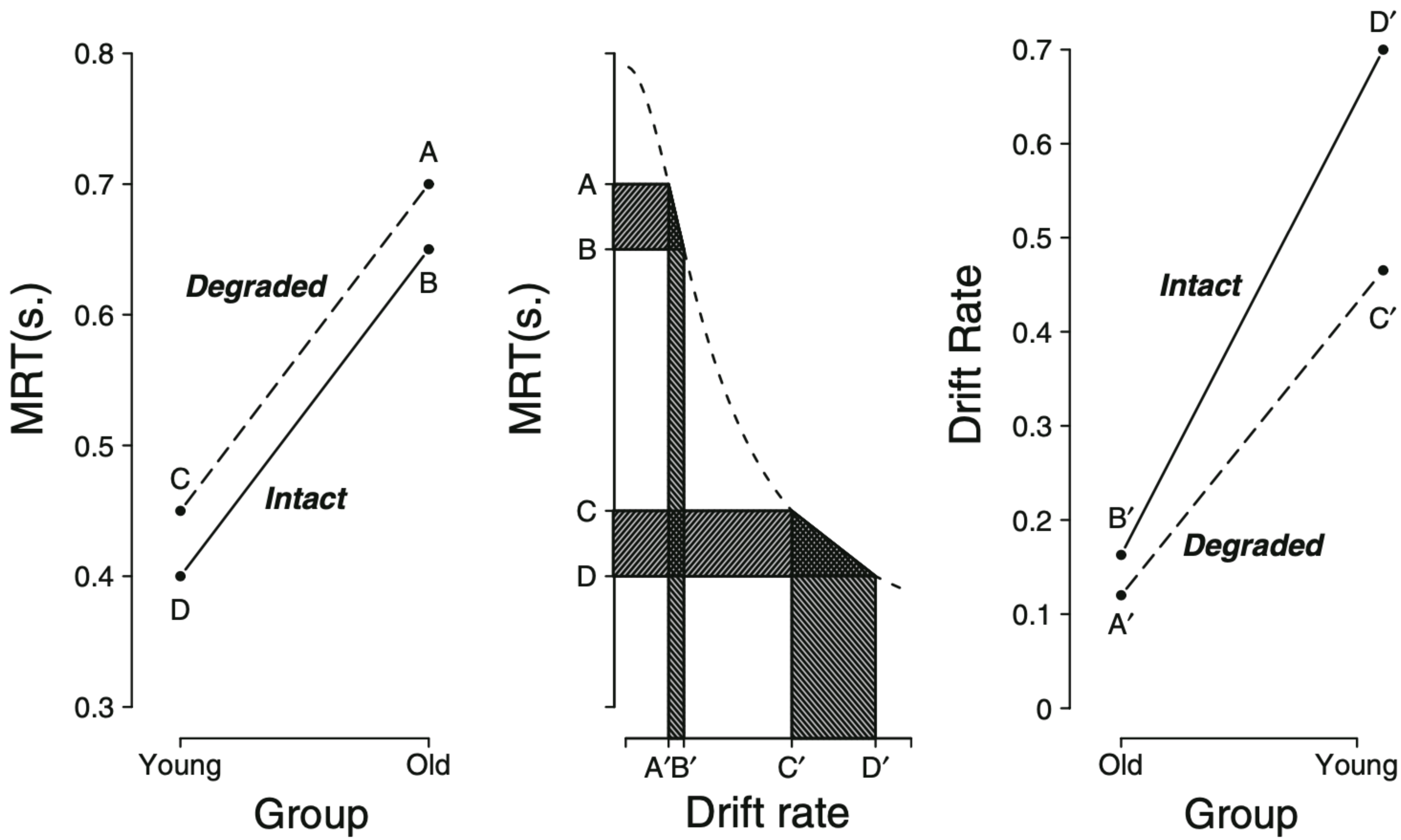
https://garstats.wordpress.com/2018/04/25/rtbias4/

Diffusion model analysis



Wagenmakers et al. 2012

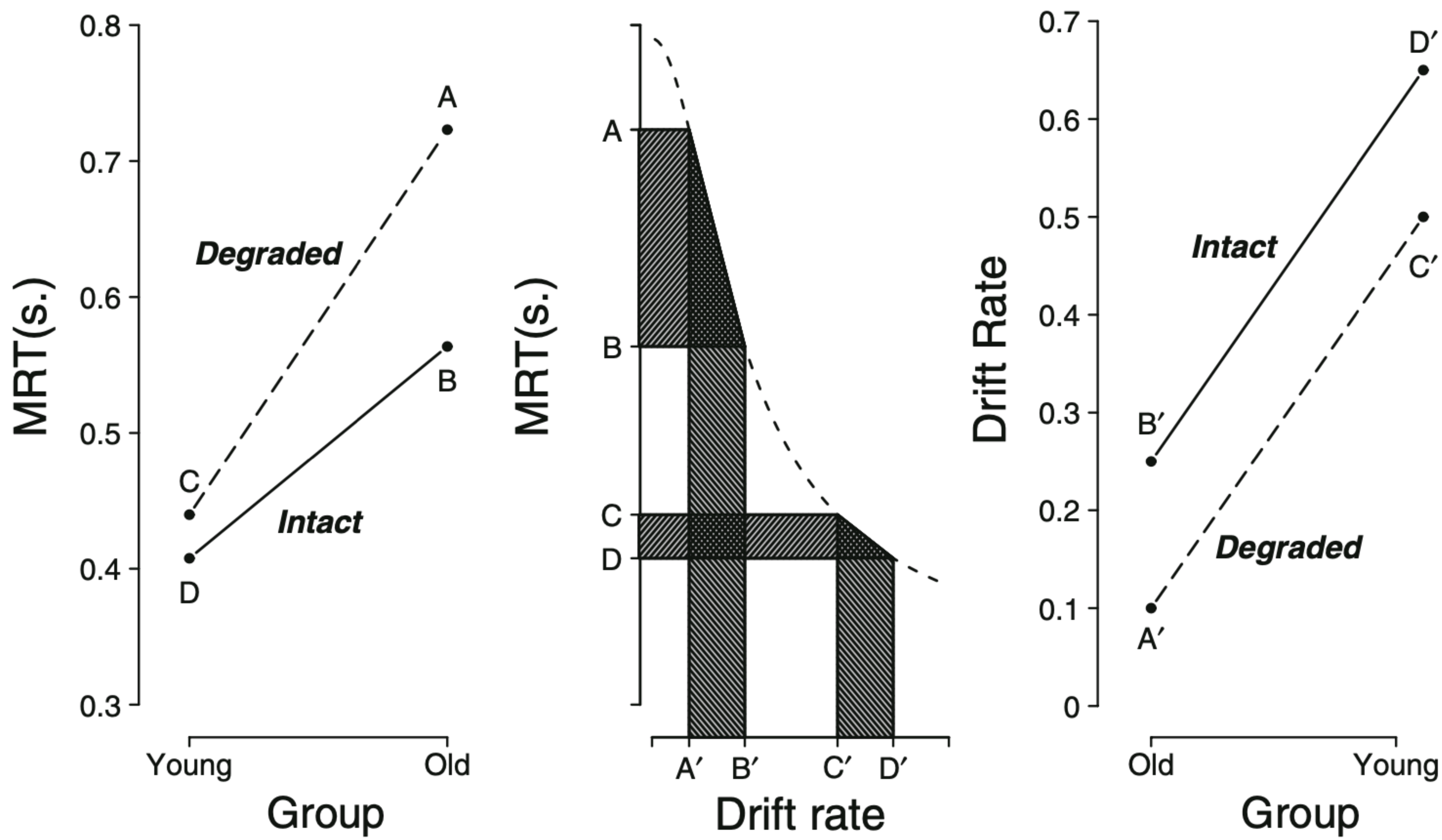
Diffusion model analysis



Additive effects on MRT correspond to interaction effects on drift rate.

Wagenmakers et al. 2012

Diffusion model analysis



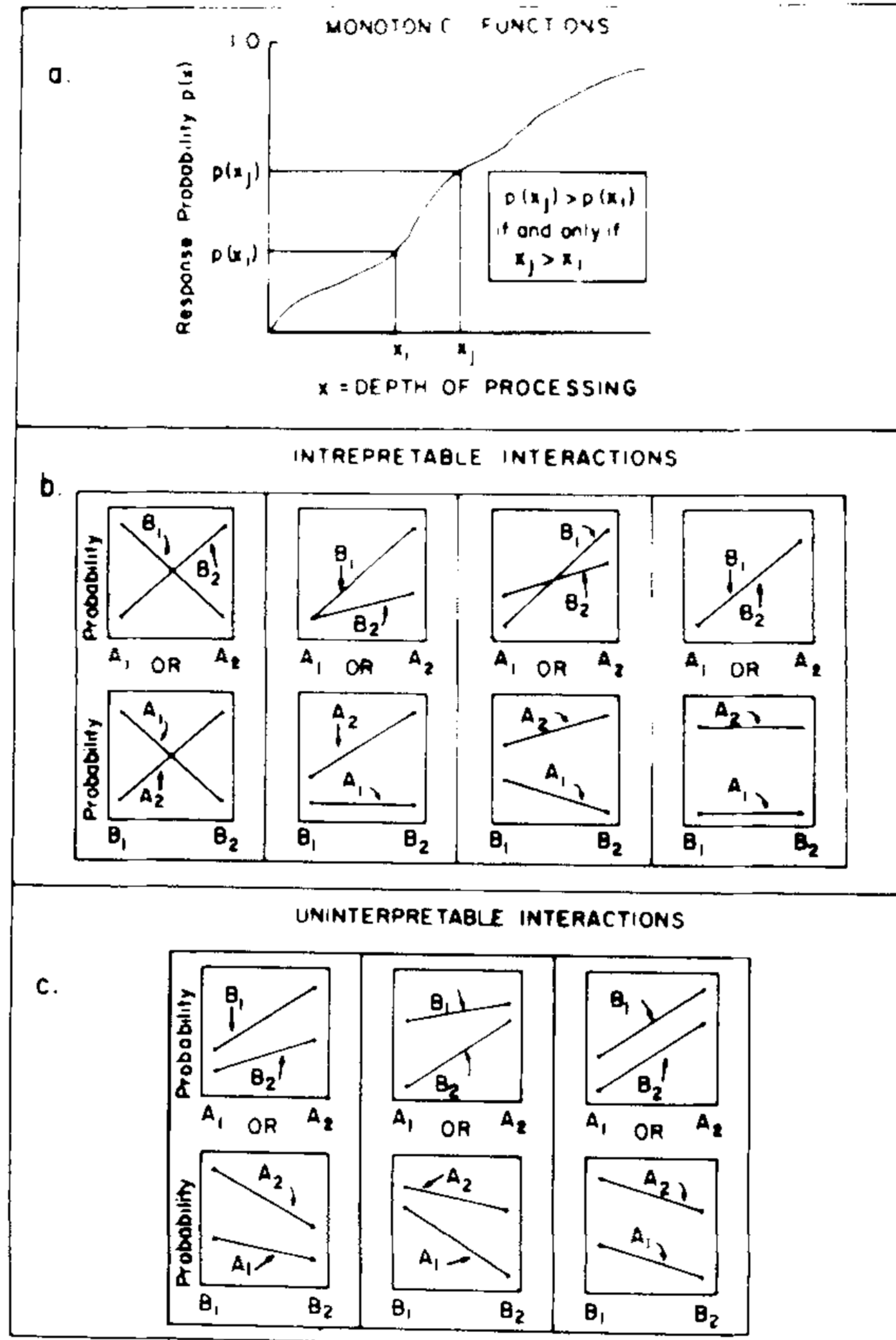
Interaction effects on MRT may correspond to additive effects on drift rate.

Wagenmakers et al. 2012

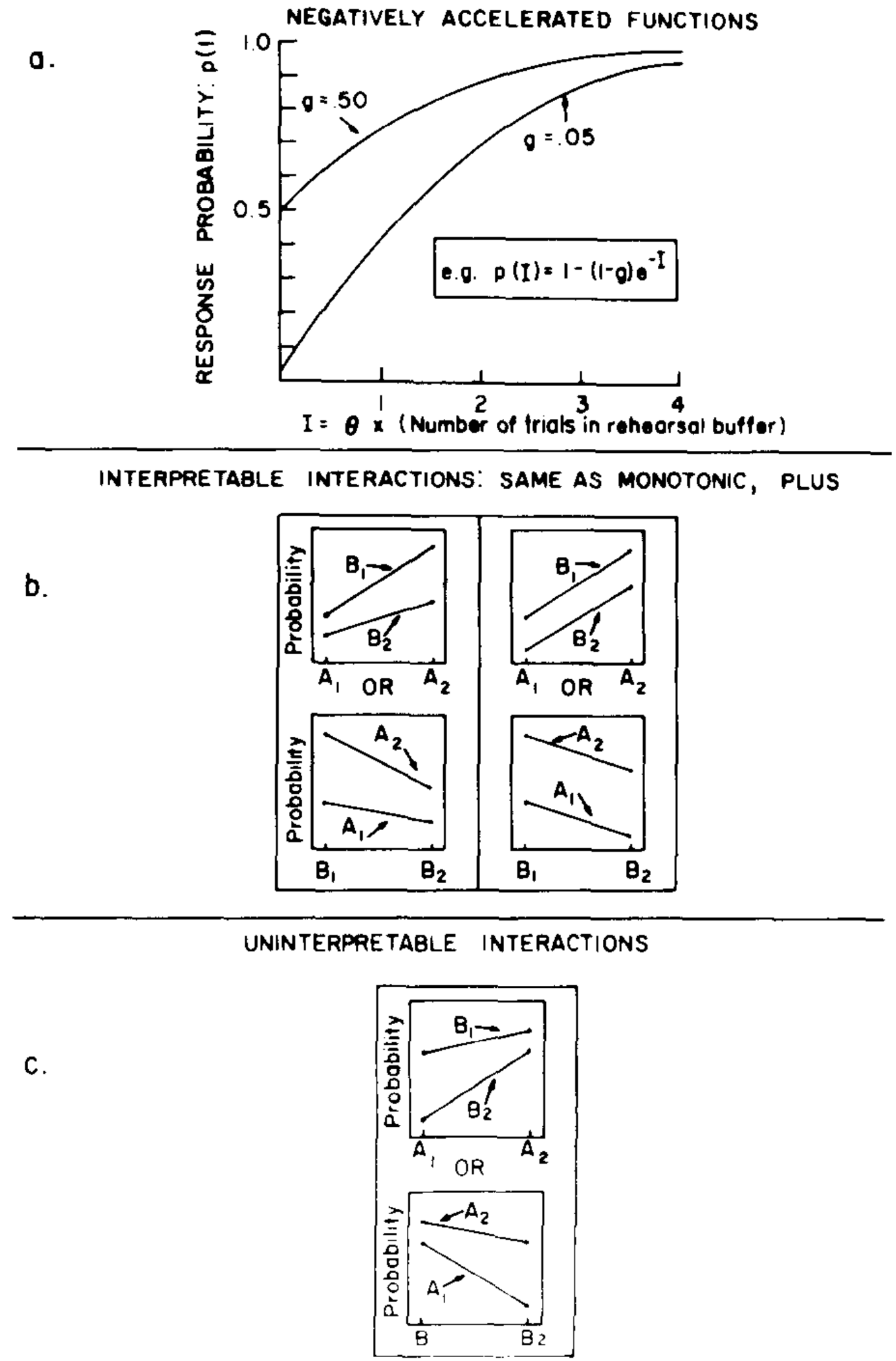
[4]
Loftus 1978's classification

Interpretable and uninterpretable interactions when monotonicity is the only assumption made about the function mapping.

"Any interaction that is not a crossover interaction is not interpretable"



Interpretable and uninterpretable interactions when a negatively accelerated function is assumed for the mapping.



[5] Wagenmakers' new classification

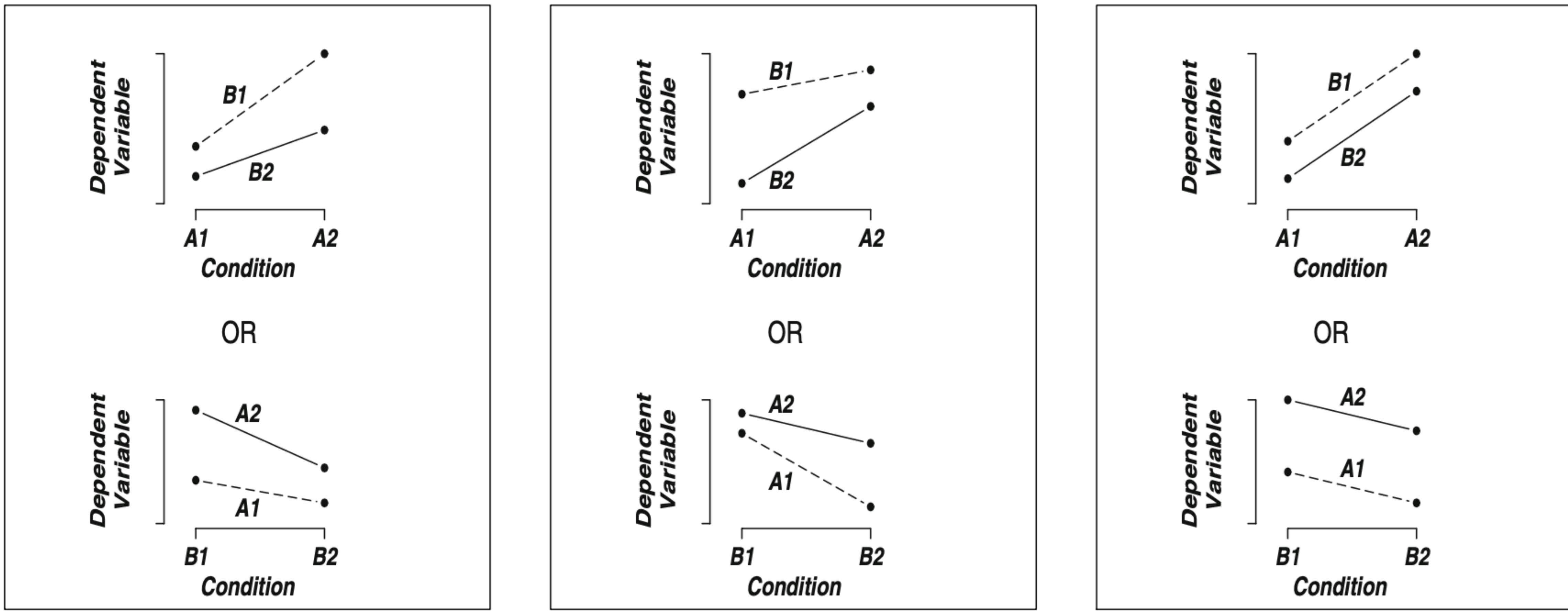
"A <u>nonremovable</u> interaction can never be undone by a monotonic transformation of the measurement scale, and it is therefore also known as qualitative, cross-over, disordinal, nontransformable, order- based, model-independent, or interpretable."

"a <u>removable interaction</u> can always be undone by a monotonic transformation of the measurement scale; such an interaction is also known as quantitative, ordinal, transformable, model-dependent, or uninterpretable."

"<u>borderline nonremovable</u>"
non-removable according to Loftus 1978
actually depends on statistical evidence for equivalence between conditions

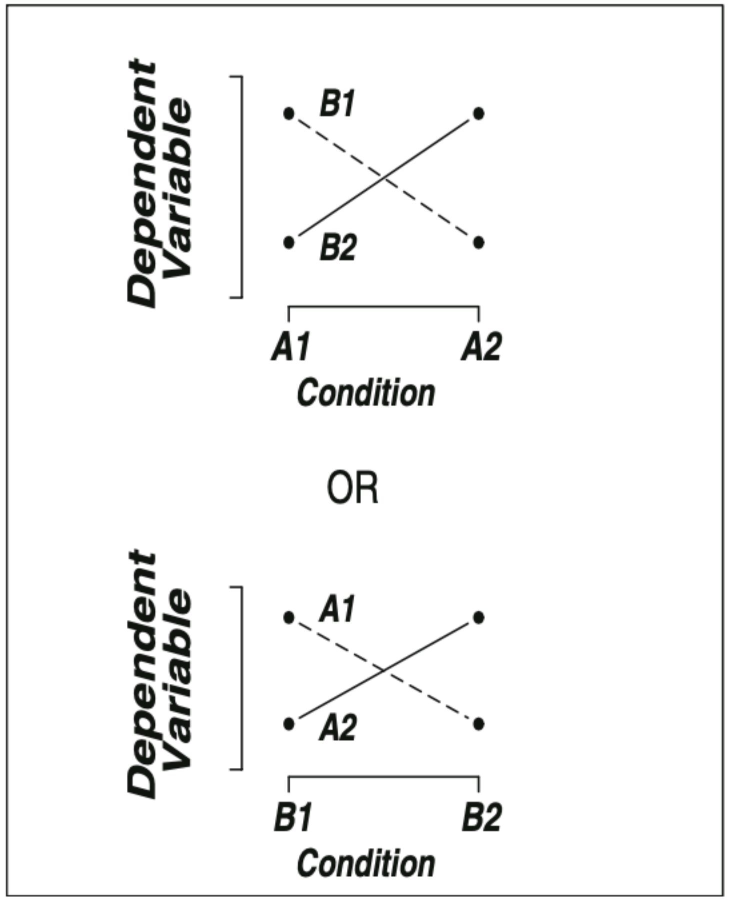
Removable interactions.

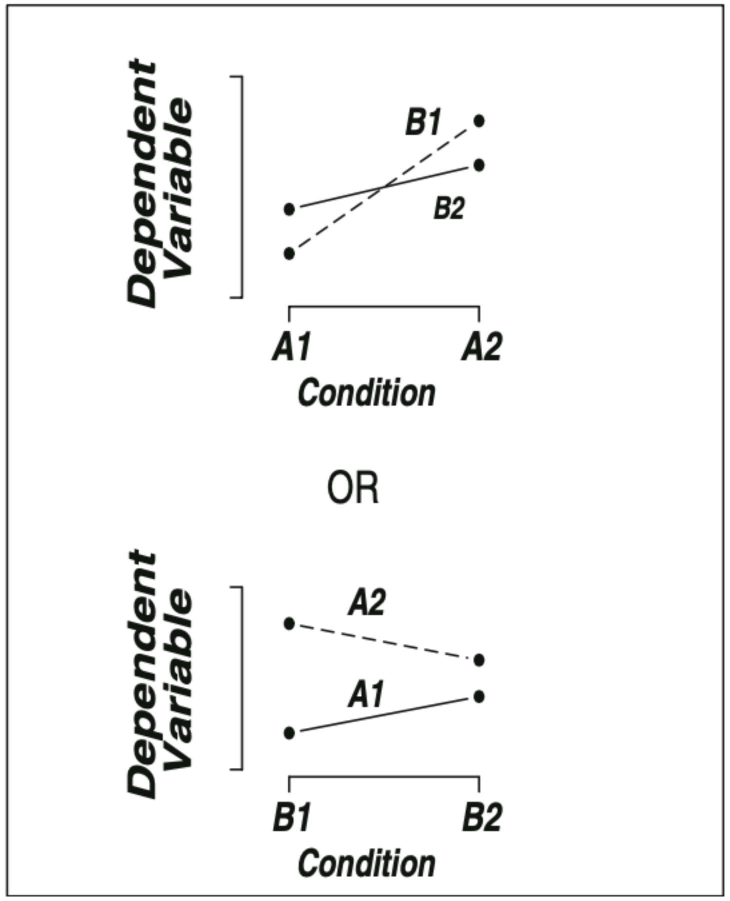
These interactions can be transformed to additivity (or vice versa) by a monotonic change of the measurement scale.



Nonremovable interactions

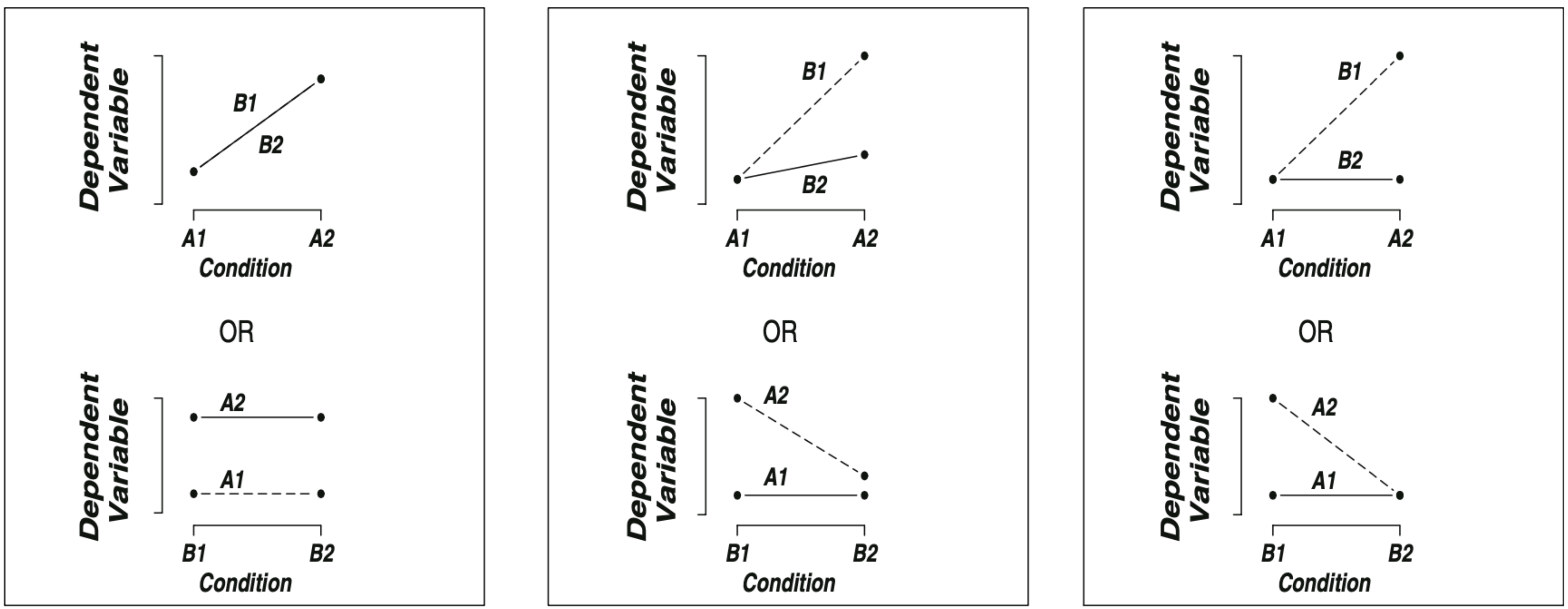
These interactions cannot be transformed to additivity by a monotonic change of the measurement scale.



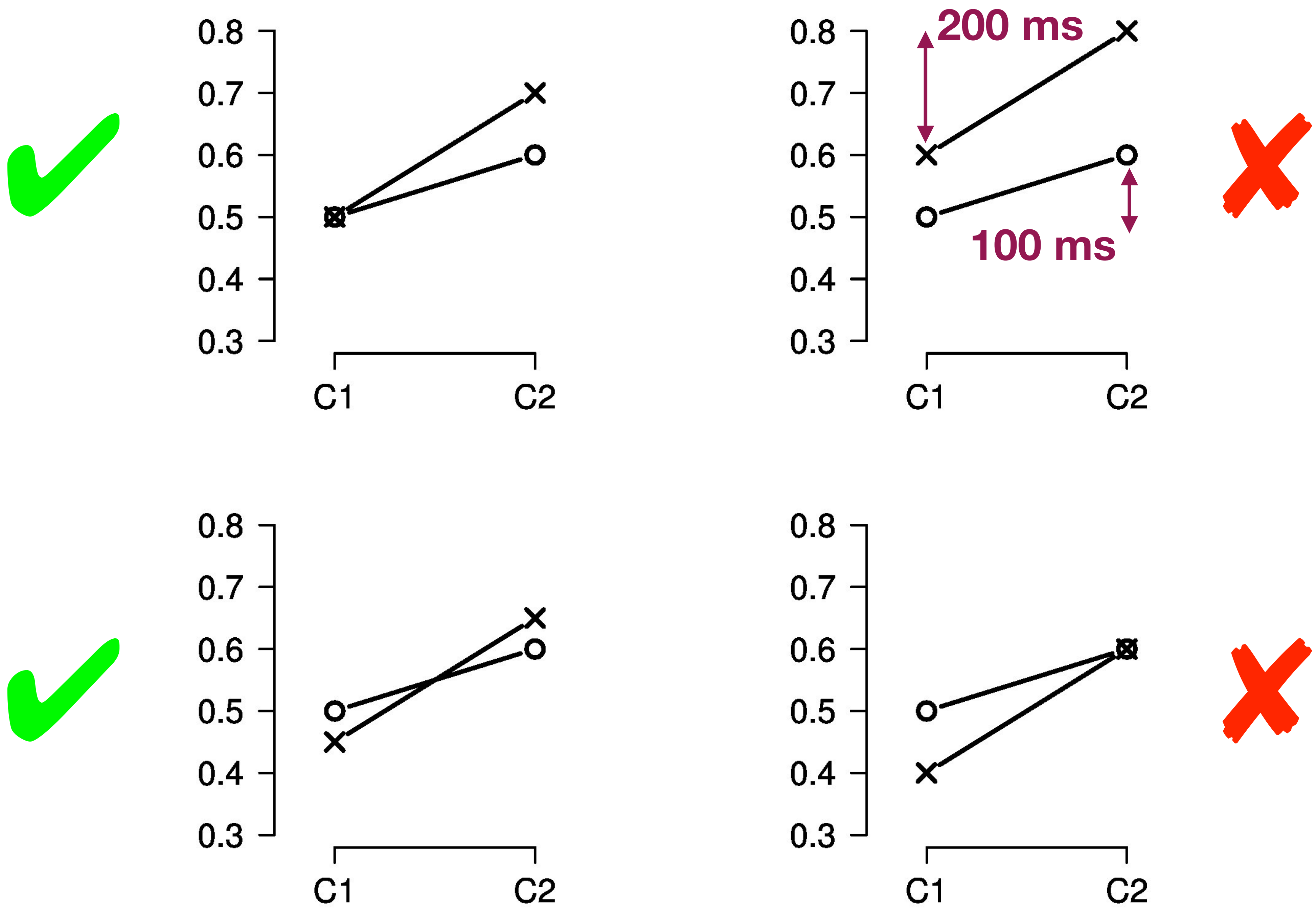


Borderline nonremovable interactions

These interactions are on the cusp between removable and nonremovable. Theoretically, these interactions are nonremovable, but in practice their classification hinges on the statistical evidence in favor of a point-null hypothesis.

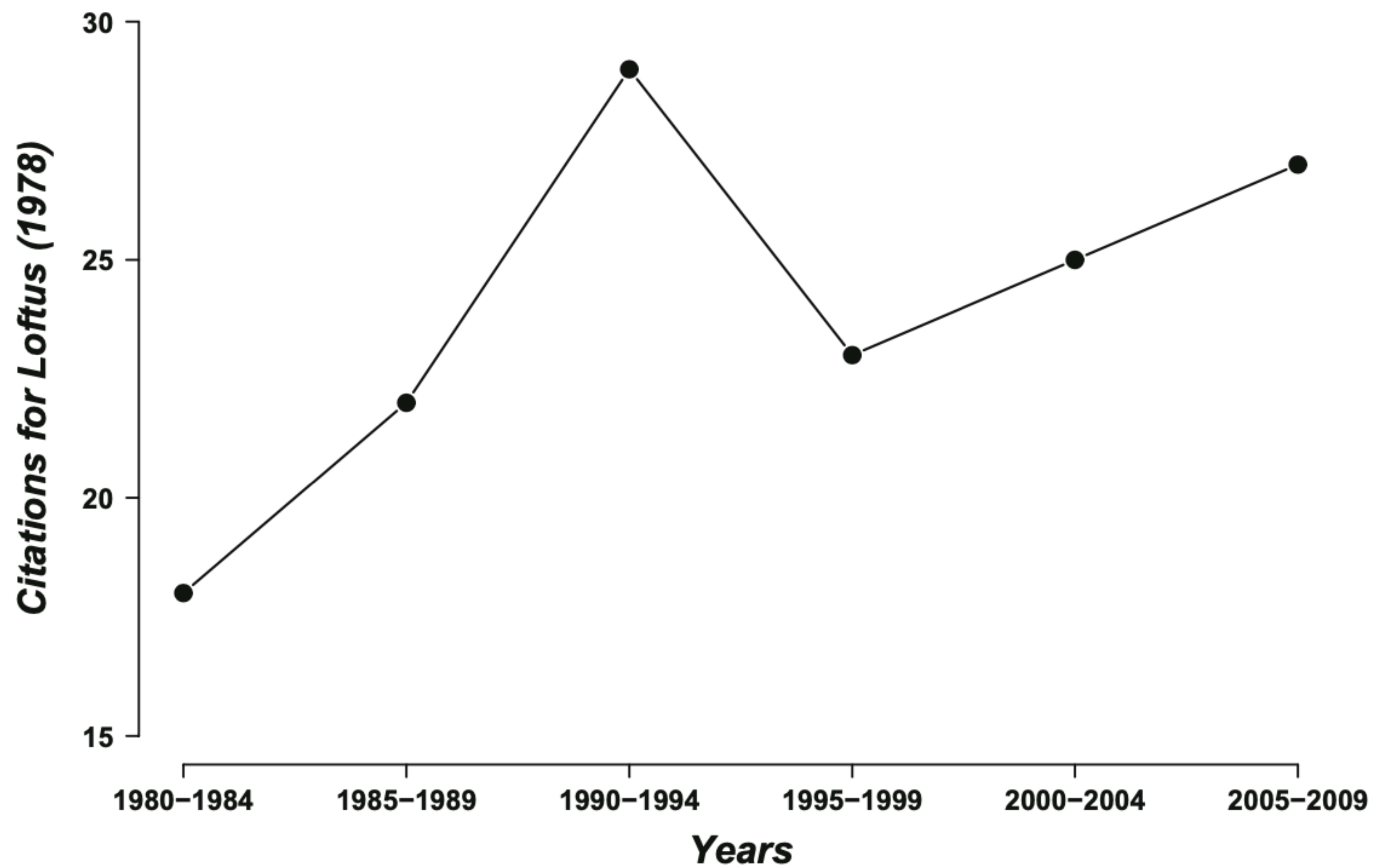


[6] Interaction quartet



Wagenmakers (2015) http://tinyurl.com/p9kl2aa

[7] Citation history



Number of peer-reviewed articles that cite Loftus (1978) in 5-year intervals

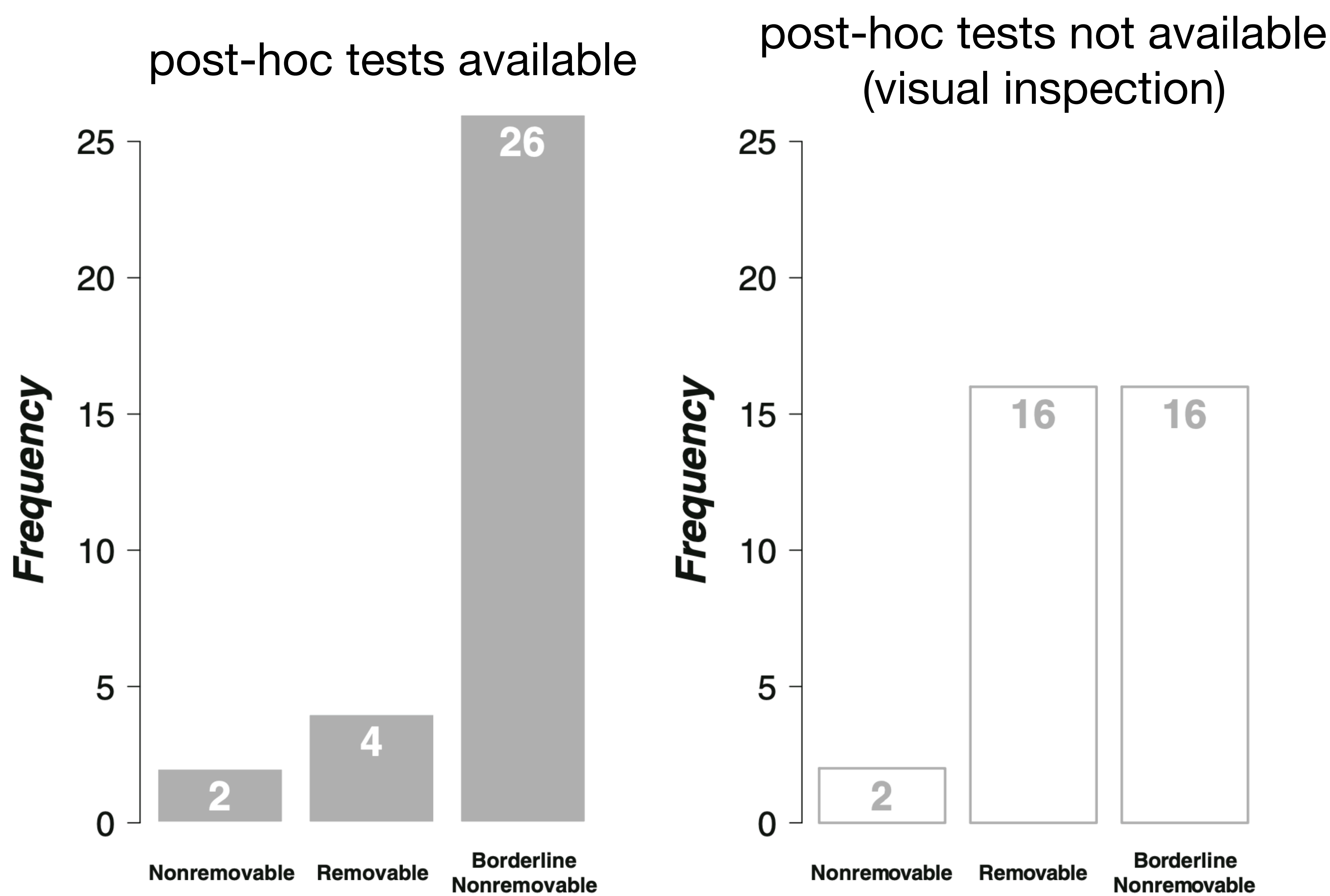
[8] Reference in statistical textbooks

- **14 popular intro textbooks** “Not a single textbook mentioned that certain interactions can be transformed away and should therefore be interpreted with caution.”

- **more advanced textbooks** - 3 books briefly discuss the issue

[9] Literature review

- All 88 articles from *Psychology and Aging* published in 2008.
- 66 significant 2 x 2 interactions
- Loftus (1978) citations?



[10] Questionnaire for students and faculty

3 interactions + cover stories
100 participants:

- **37 master students**
- **36 PhD students**
- **19 professors**

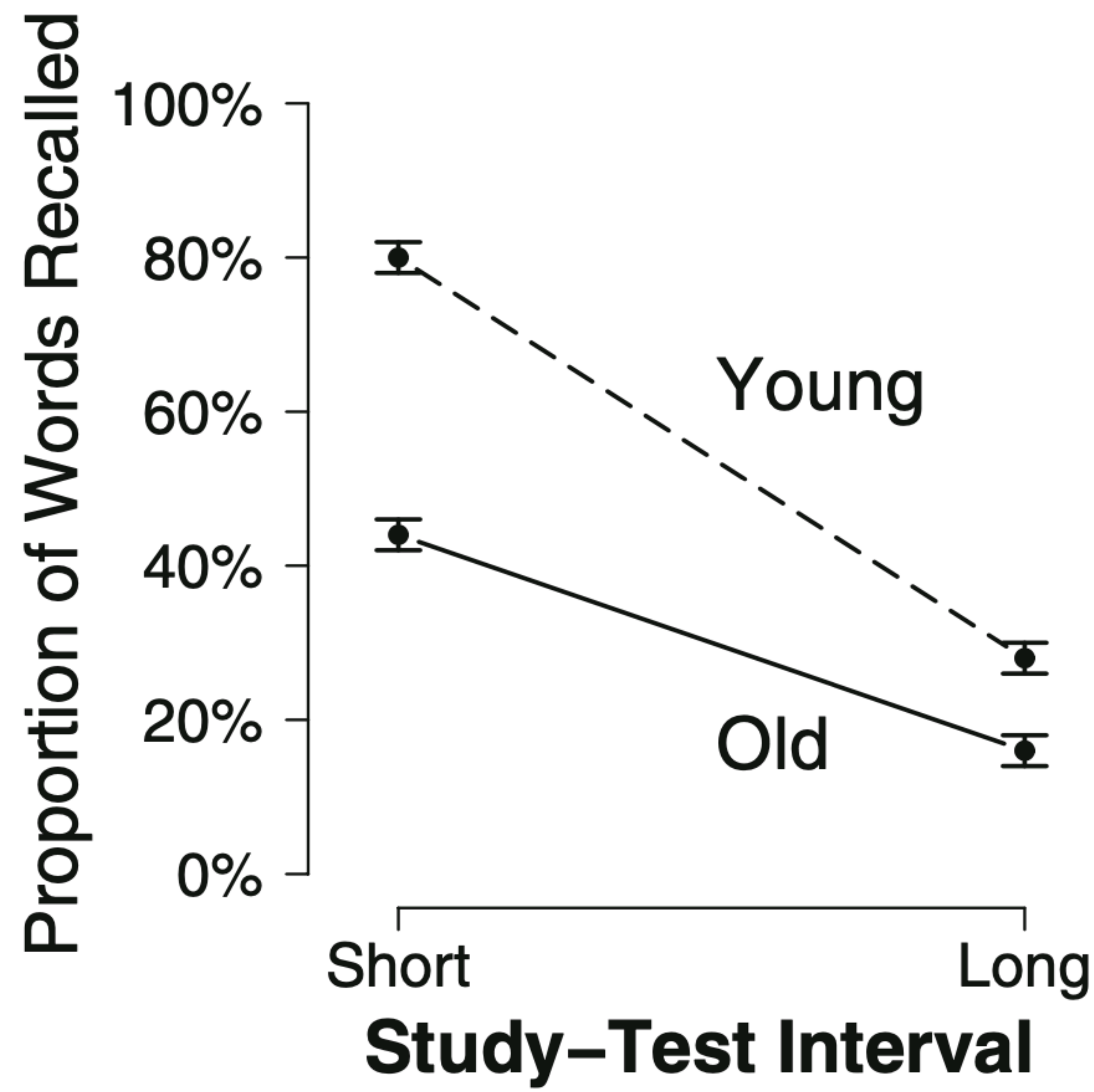
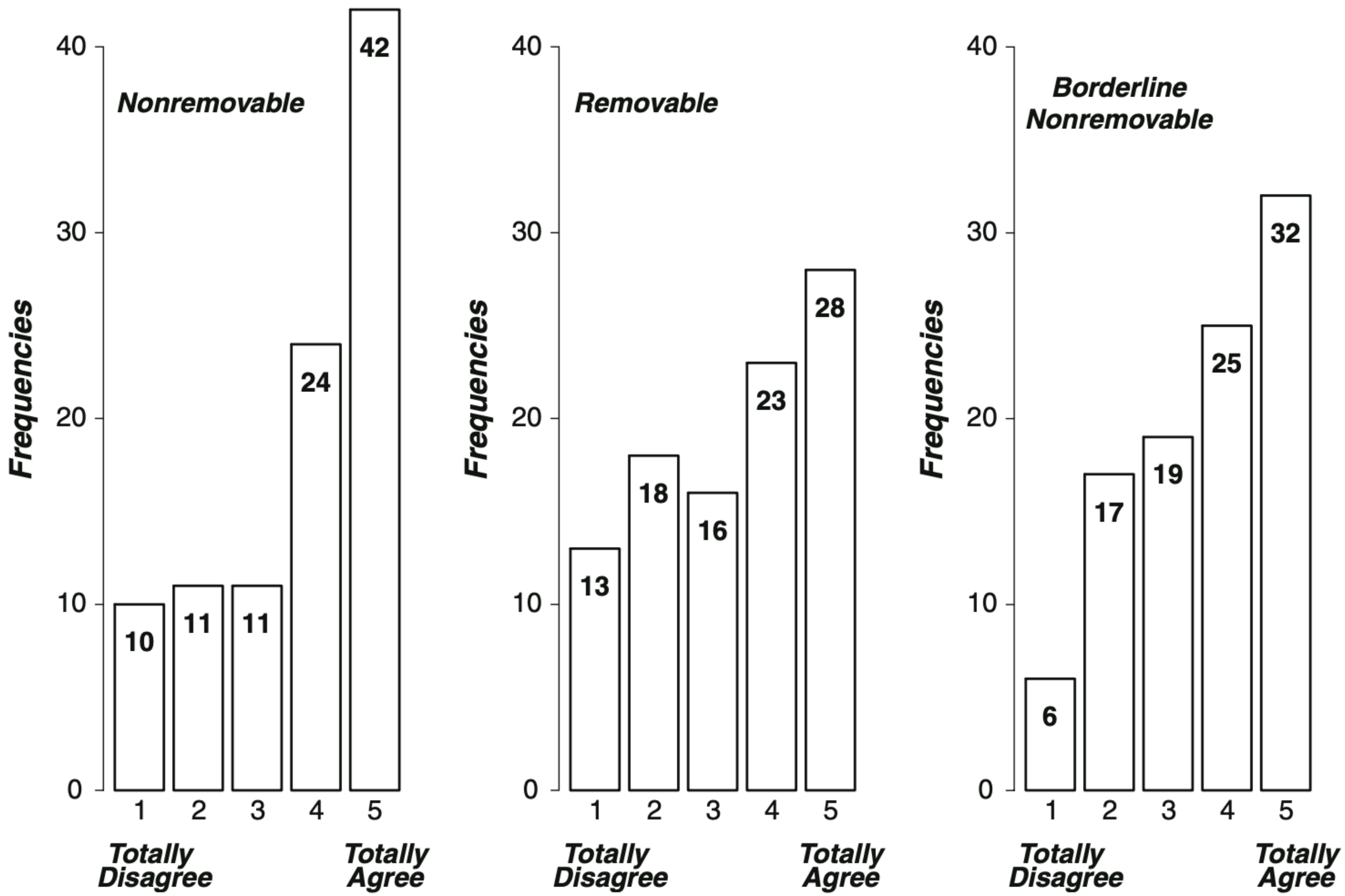


Fig. 10 Example item from a questionnaire that tests knowledge of removable interactions. After reading a cover story, participants were confronted with this figure and had to indicate their level of agreement with the statement "An increase in study–test interval affects long-term memory of young adults more than it affects that of older adults"

Students and faculty members in psychology generally agree with the statement that synthetic data show an interaction, even when this statement is formulated in terms of a latent psychological process.



“In their open-ended responses, only four out of 100 participants correctly identified the removable interaction as such.”

[11] What can we do?

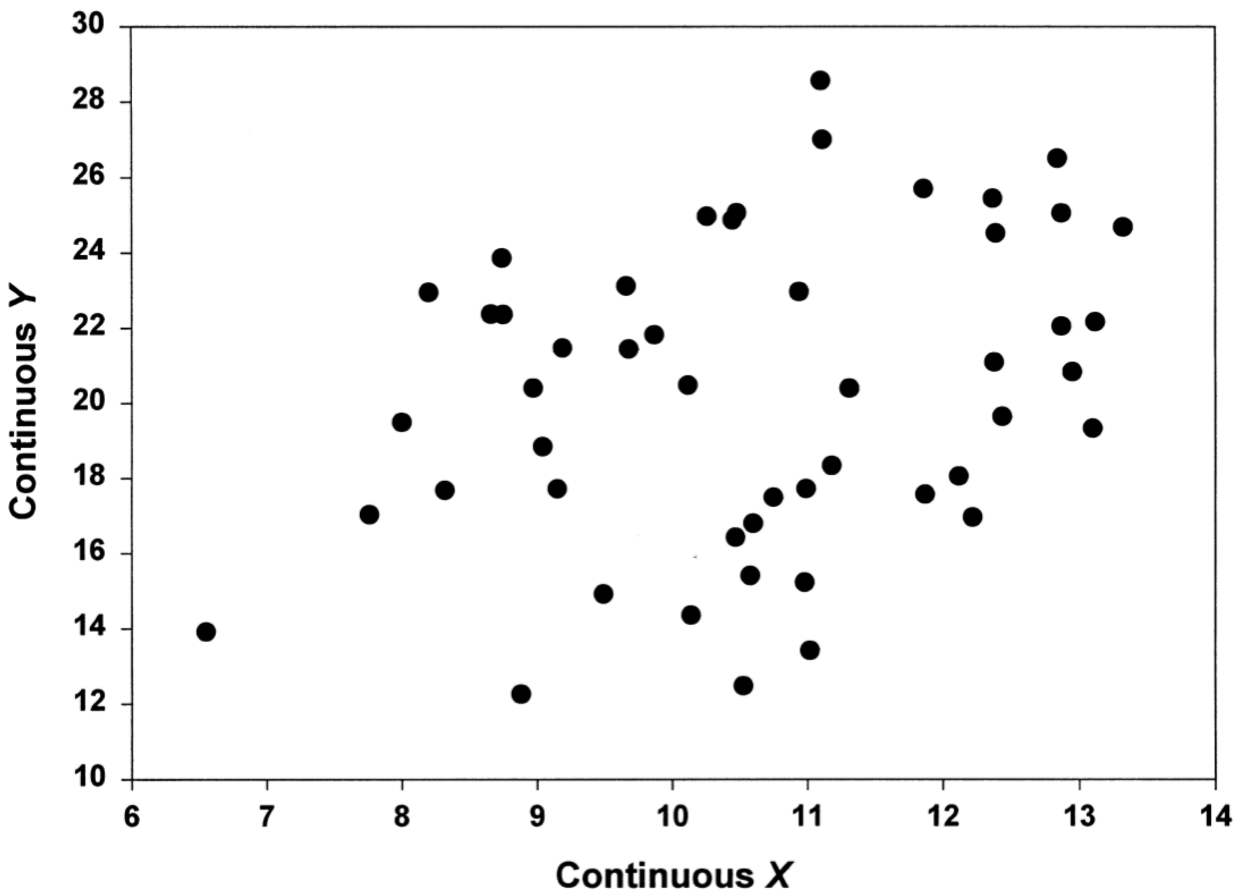
- Teach the problem
- Mention the problem in reviews and editorial decision letters
- Mention the problem in our articles, adding limitations of our interpretations:
- “There is an interaction between A and B at the level of proportion correct measurements; this suggests an interaction at the level of the unobserved variable X, assuming a (highly improbable) linear relationship between measurements and X. A monotonic transformation of the measurement scale could remove the interaction.”

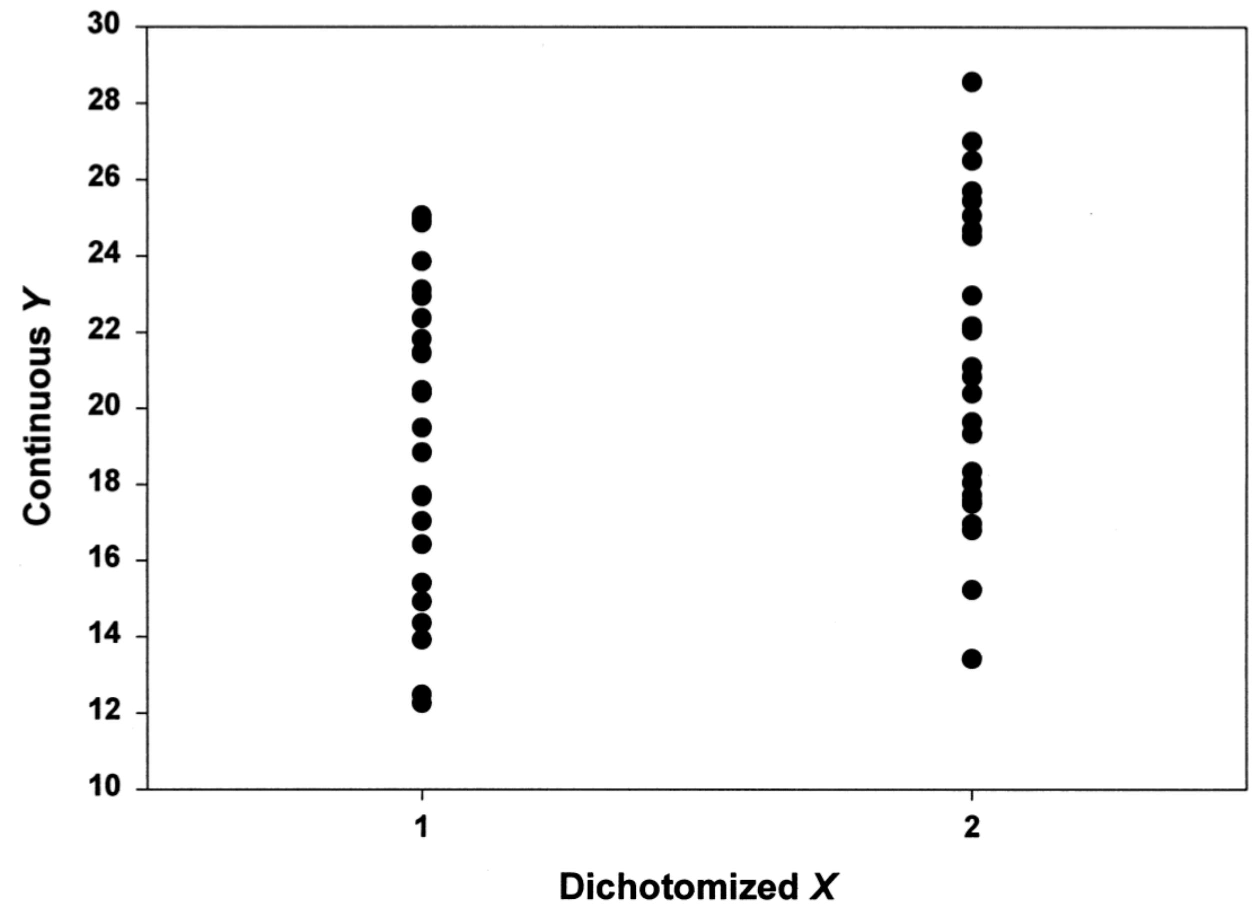
[11] What can we do?

- Use designs in which performance is equated across groups in the easier condition.
- Check the robustness of the interaction to various data transformations.
  - RT: 1/RT, log(RT)
  - PC: logit(p), *d'*

[11] What can we do?

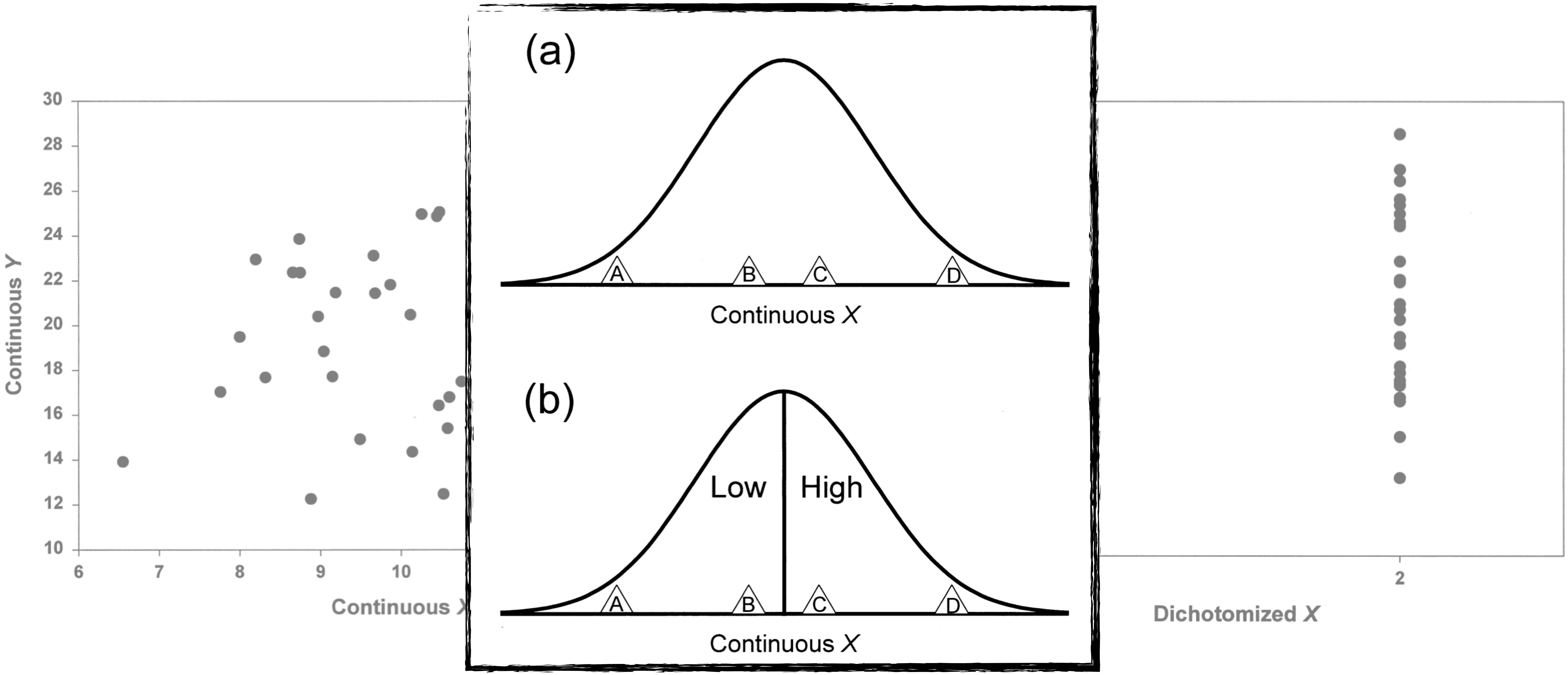
[12] Dichotomisation of continuous variables





Maccallum, R.C., Zhang, S., Preacher, K.J., & Rucker, D.D. (2002)
On the practice of dichotomization of quantitative variables.
Psychological methods, **7**, 19–40.

[12] Dichotomisation of continuous variables



Maccallum, R.C., Zhang, S., Preacher, K.J., & Rucker, D.D. (2002)
On the practice of dichotomization of quantitative variables.
Psychological methods, **7**, 19–40.

Dichotomania

It's all in the title...(1994-2006)

1. **Problems in dichotomizing continuous variables** (Altman 1994)
2. **Dangers of using "optimal" cutpoints** in the evaluation of prognostic factors. (Altman et al 1994)
3. **How bad is categorization?** (Weinberg; 1995)
4. **Seven reasons why you should NOT categorize continuous data** (Dinero; 1996)
5. Breaking Up is Hard to Do: The **Heartbreak of Dichotomizing Continuous Data** (Streiner; 2002)
6. **Negative consequences of dichotomizing continuous predictor variables** (Irwin & McClelland; 2003)
7. **Why carve up your continuous data?** (Owen 2005)
8. **Chopped liver? OK. Chopped data? Not OK. Chopped liver? OK. Chopped data? Not OK** (Butts & Ng 2005)
9. **Categorizing continuous variables resulted in different predictors** in a prognostic model for nonspecific neck pain (Schellingerhout et al 2006)

6

It's all in the title...(2006-2014)

10. **Dichotomizing continuous predictors in multiple regression: a bad idea** (Royston et el 2006)
11. The **cost of dichotomising continuous variables** (Altman & Royston; 2006)
12. **Leave 'em alone - why continuous variables should be analyzed as such** (van Walraven & Hart; 2008)
13. **Dichotomization of continuous data--a pitfall** in prognostic factor studies (Metze; 2008)
14. Analysis by **categorizing or dichotomizing continuous variables is inadvisable**: an example from the natural history of unruptured aneurysms (Naggara et al 2011)
15. **Against quantiles: categorization of continuous variables** in epidemiologic research, and its discontents (Bennette & Vickers; 2012)
16. **Dichotomizing continuous variables** in statistical analysis: **a practice to avoid** (Dawson & Weiss; 2012)
17. The **danger of dichotomizing continuous variables**: A visualization (Kuss 2013)
18. The "**anathema**" of **arbitrary categorization of continuous predictors** (Vintzileos et al; 2014)
19. Ophthalmic statistics note: **the perils of dichotomising continuous variables** (Cumberland et al 2014)

https://twitter.com/GSCollins/status/1026541340748701698

Key problem: can introduce spurious interactions!