

Meta-analysis of response-time tests of the sequential two-systems model of moral judgment

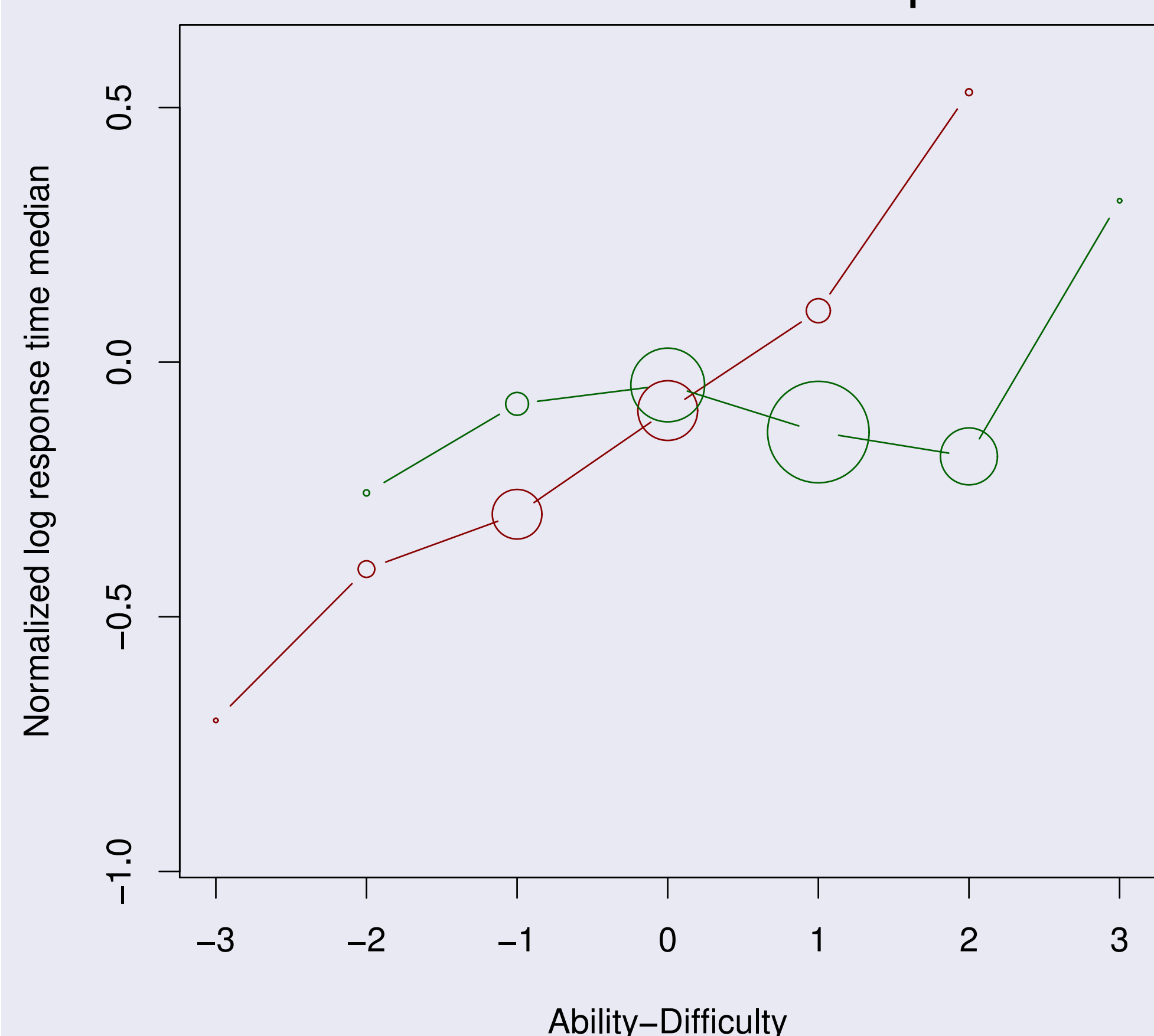
Jonathan Baron and Burcu Gürçay, University of Pennsylvania. Full paper at <http://finzi.psych.upenn.edu/~baron/ms/bg/meta.pdf>.

Introduction

- The sequential two-system (“default interventionist”) model of utilitarian moral judgment predicts that utilitarian responses (e.g., push a man off a bridge to stop a trolley headed for 5 others) often arise from a system-two correction of system-one deontological intuitions (don’t push).
- Utilitarian responses should thus take longer.
- An alternative model, the **conflict model**, says that utilitarian and deontological responses vie with each other during the entire decision process.
- To compare models, we need to equate probability of a utilitarian response pU , which itself affects response time (RT).
- We don’t know when a response has $pU = .5$, so we must estimate what RT would be if $pU = .5$
- Baron, Gürçay, Moore & Starcke (2012) used a Rasch model to predict pU from each subject’s tendency to make utilitarian responses (“Ability”, A) and each dilemma’s tendency to elicit deontological responses (“Difficulty”, D). We fit the model to the data from each experiment with many subjects and about 10 moral dilemmas.
- The Rasch model says that pU is a logistic function of $A - D$. It fit well enough.
- The sequential model says that RT for utilitarian responses should be greater than for deontological responses when $A - D = 0$. The conflict model says these RTs should be the same.

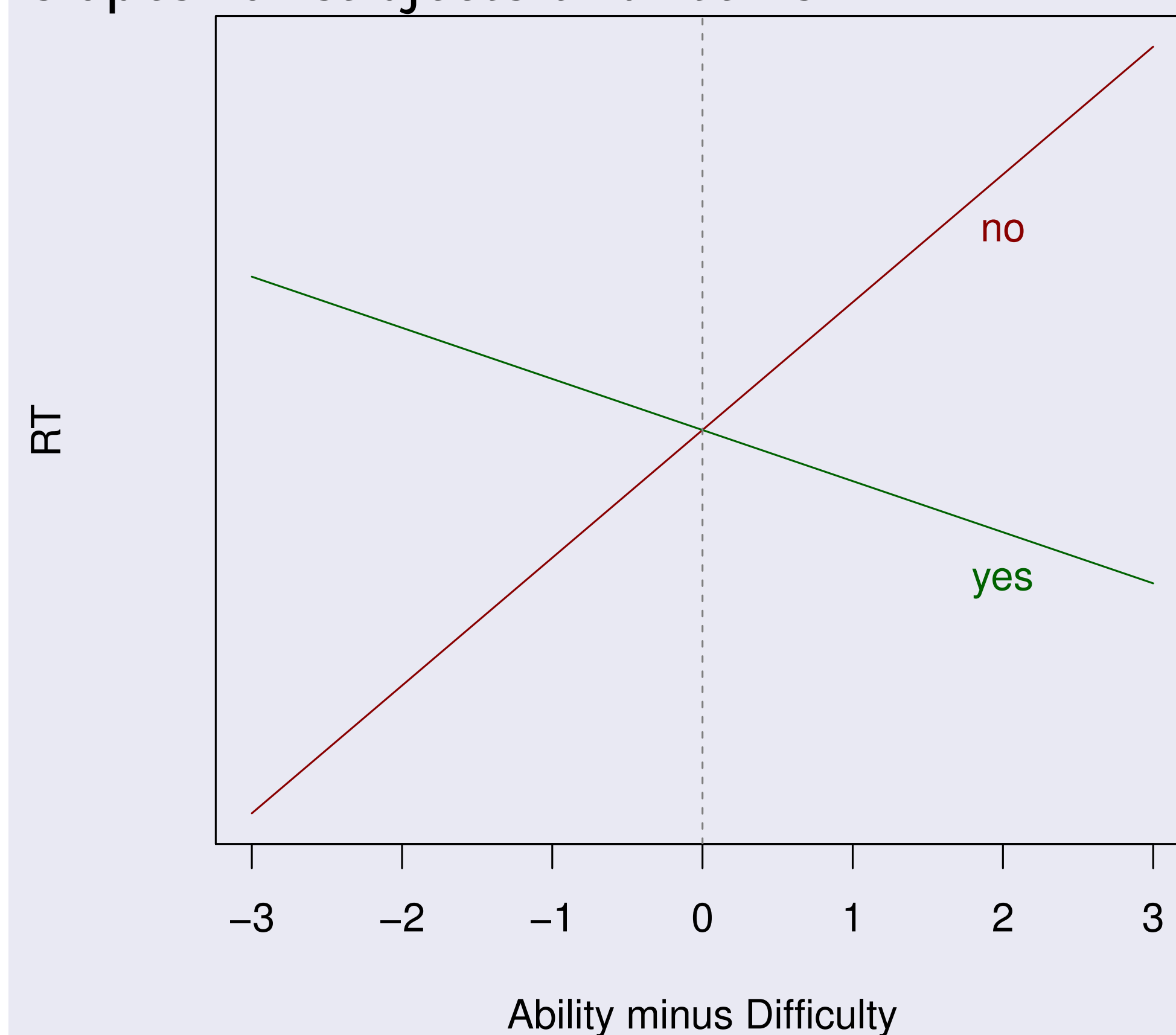
Example of data from one study (BG2)

Red for deontological responses, green for utilitarian. Circle areas are proportional to the number of observations for each point.



Linear model fit to data each study

To compare models, Baron et al. fit the data from each of 5 studies with a model like that in the following figure, here showing the predictions of the conflict model (identical RT when $A - D = 0$). We used the `lmer()` function in the R package `lme4`, with random slopes for subjects and items.



$$RT = b_0 + b_1AD + b_2U + b_3AD \times U$$

RT is $\log(\text{Response time})$.

U is the response, 1 (Yes) for utilitarian, -1 (No) for deontological.

AD is Ability–Difficulty. When AD is 0 RT for yes and no responses should be equal. So the main effect of U , the intercept, at this point, is the critical parameter, the distance between the two lines at this point. AD also tells us about the relative speed of the two conflicting processes.

$AD \times U$, the interaction term, shows how RT to Yes and No responses changes as a function of pU , i.e., as a function of AD .

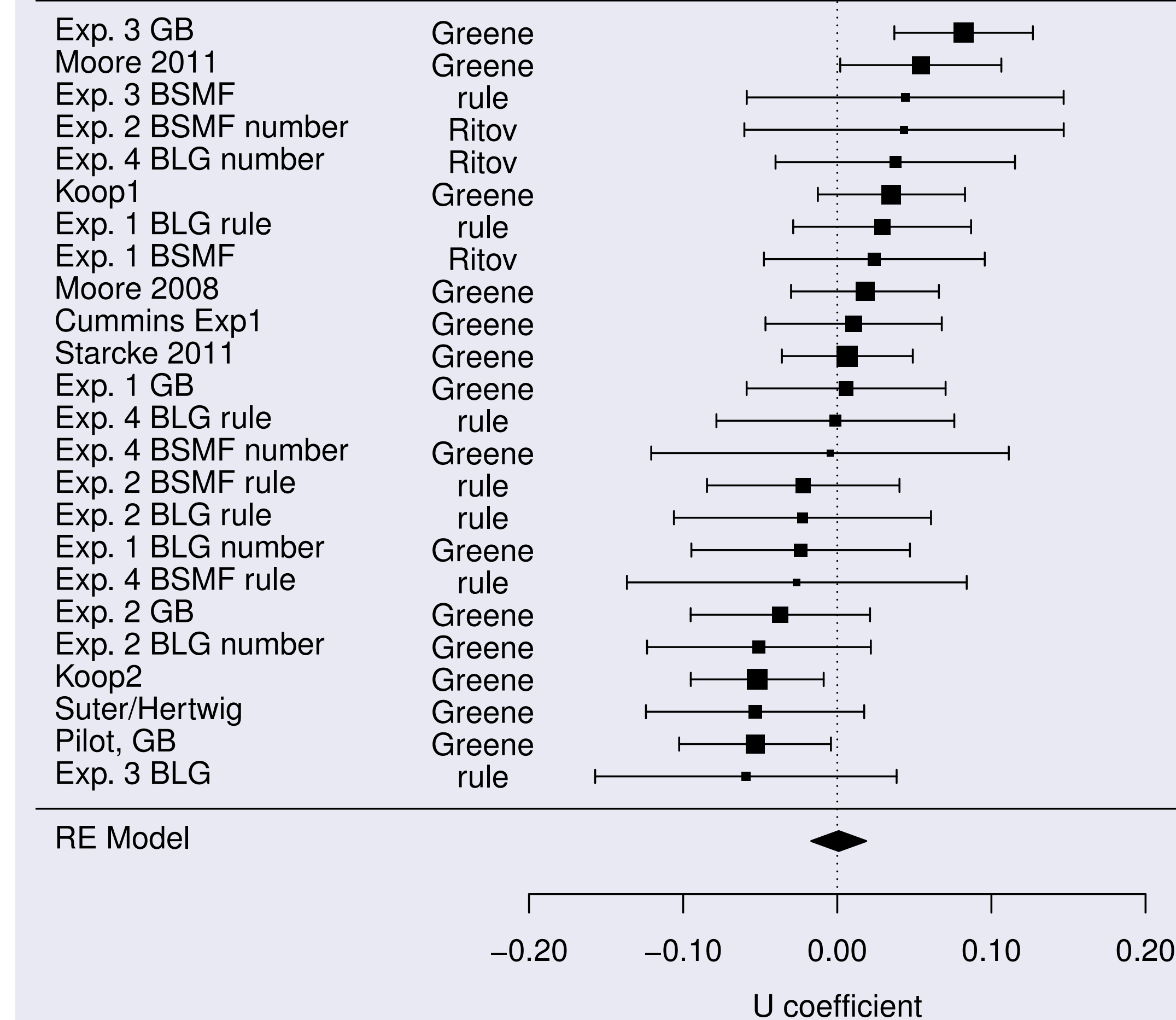
Dilemma types in 24-study meta-analysis

We extended the above analysis to 24 studies, with 3 different dilemma types:

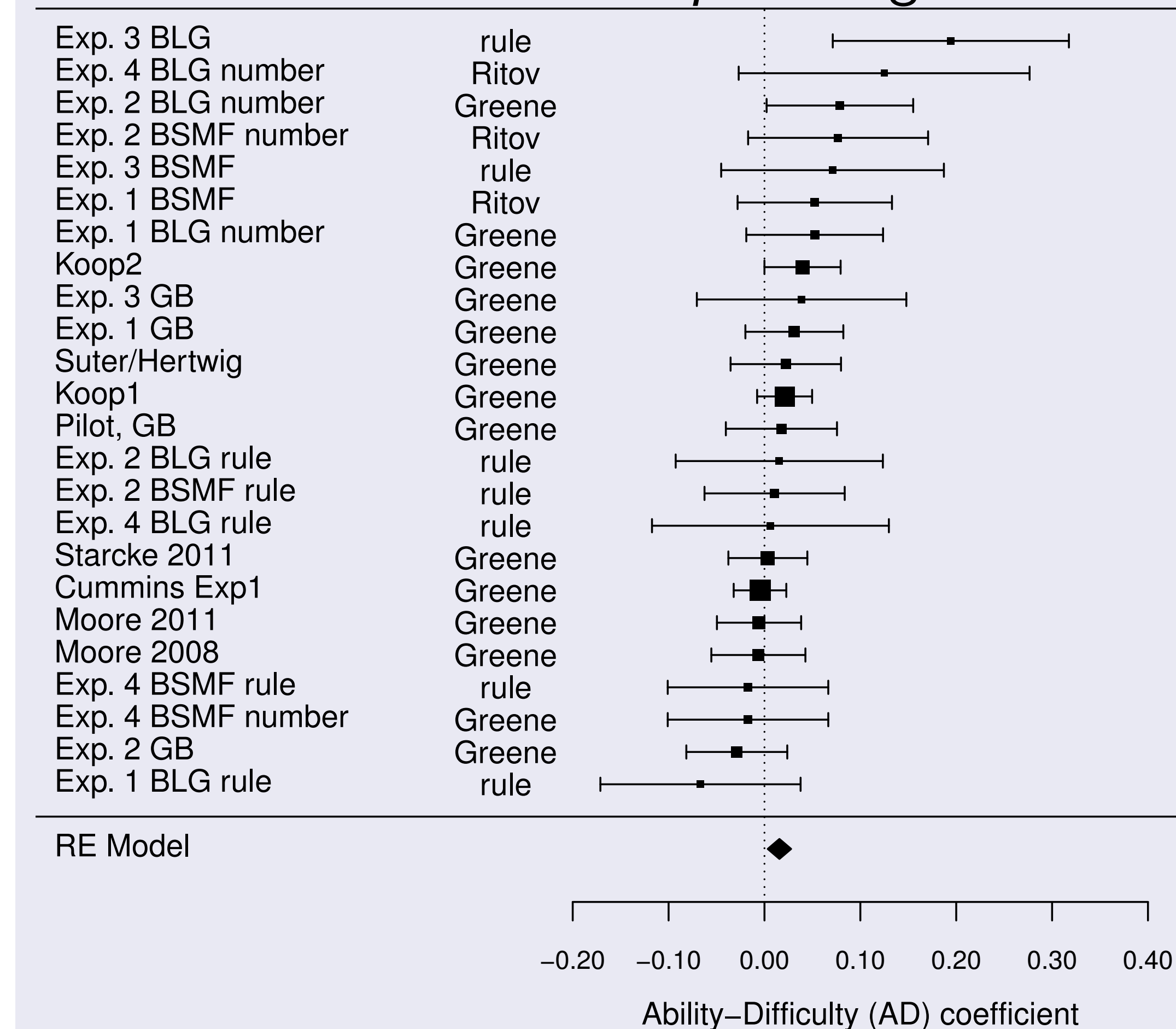
- Greene:** Fantastic sacrificial dilemmas in which one person must be killed in some gruesome way to save several others.
- Ritov:** More realistic dilemmas in which an act causes harm but prevents greater harm (as defined by numbers harmed).
- Rule:** A utilitarian response leads to a better outcome vs. a deontological response that violates a moral rule (e.g., do not lie under oath, or break a solemn vow). Both options are usually acts, and outcomes are not distinguished by numbers.

Meta-analysis of 24 studies

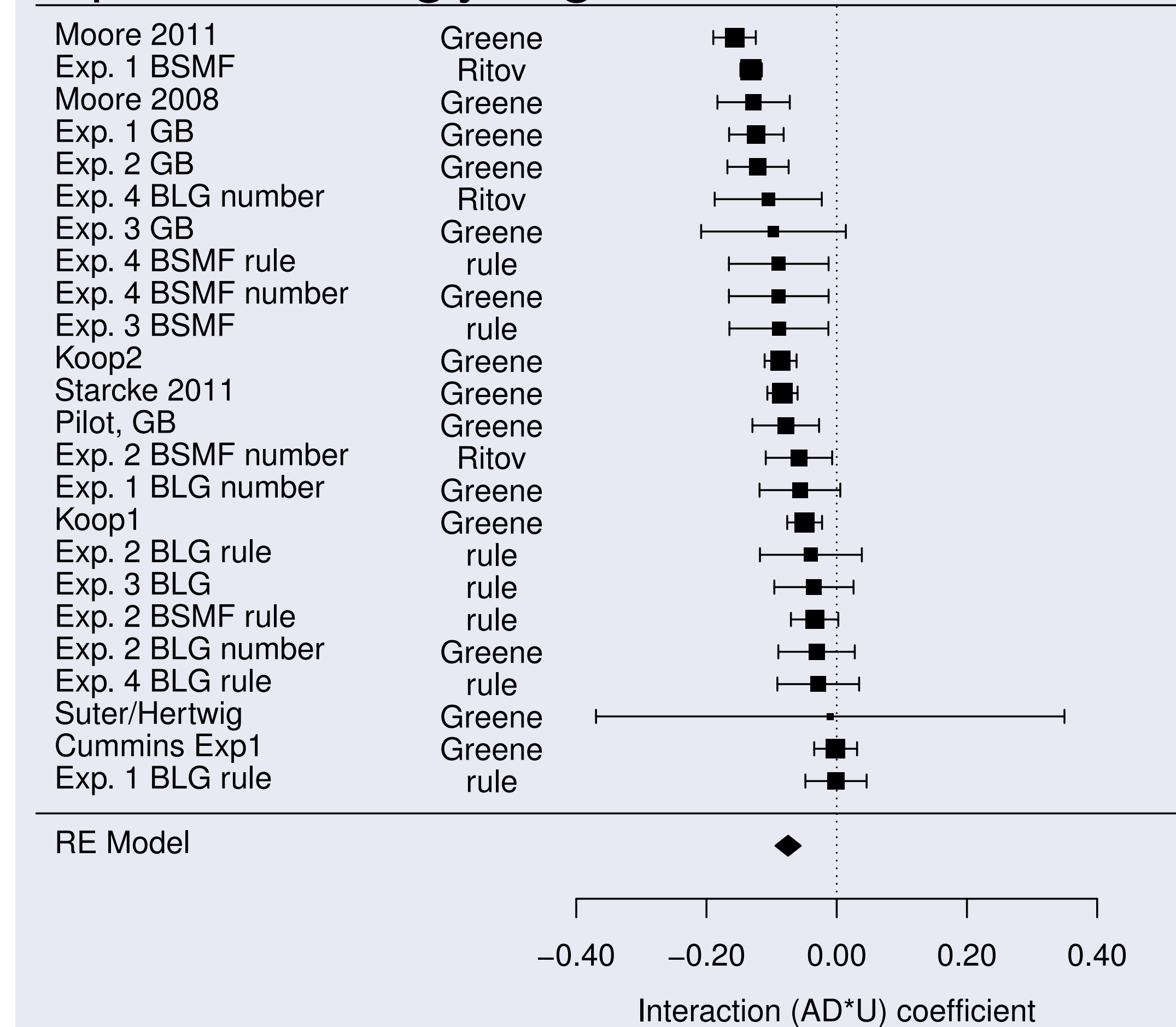
The U parameter (RT differences between utilitarian and deontological responses when pU is projected to be .5) is zero, supporting the conflict model against the sequential model.



Surprisingly, the AD parameter (RT as a function of $A - D$) is positive. RT increases under conditions in which pU is higher.

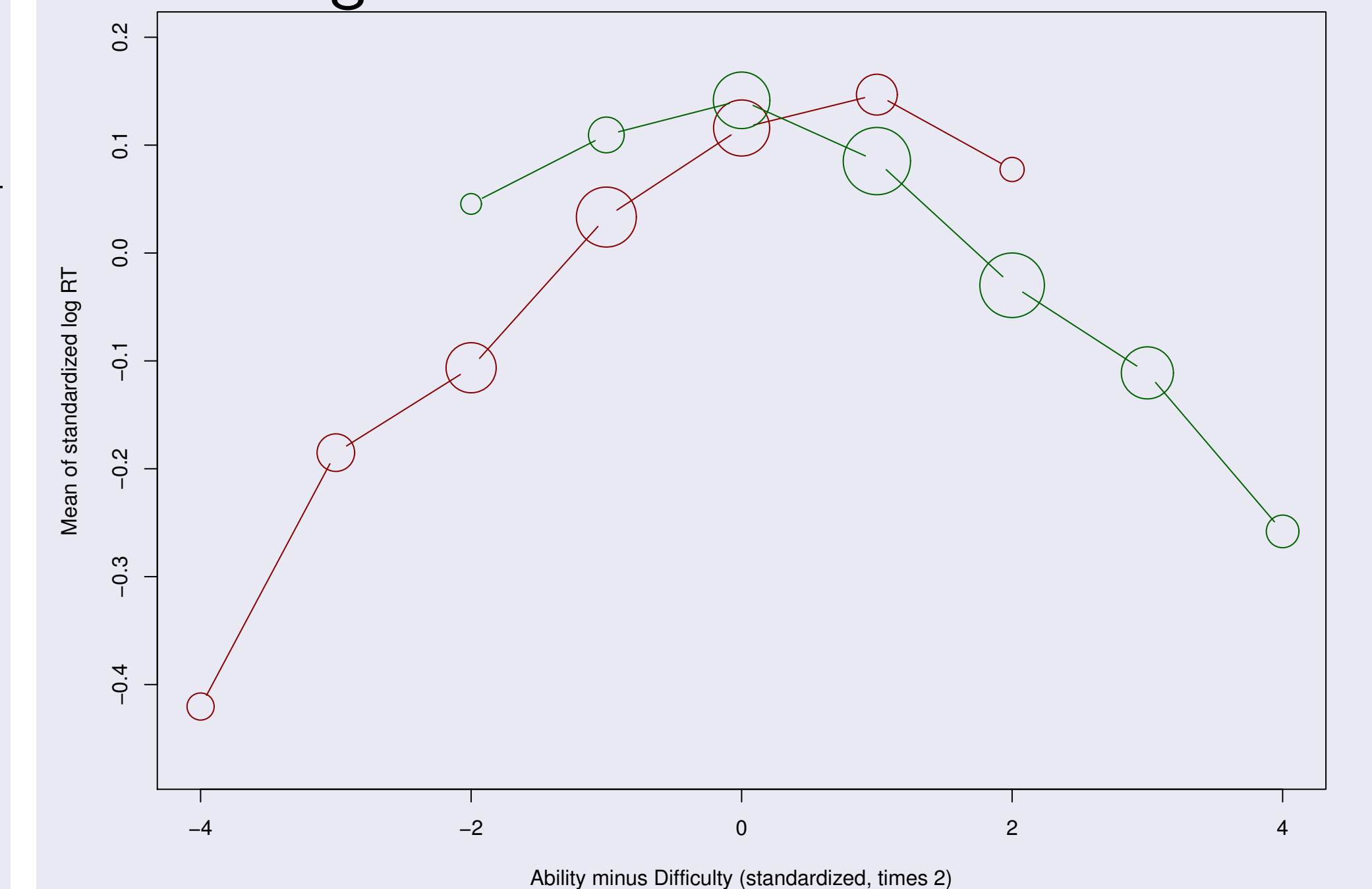


$AD \times U$ (interaction between $A - D$ and type of response, utilitarian or deontological) is, as expected, strongly negative.

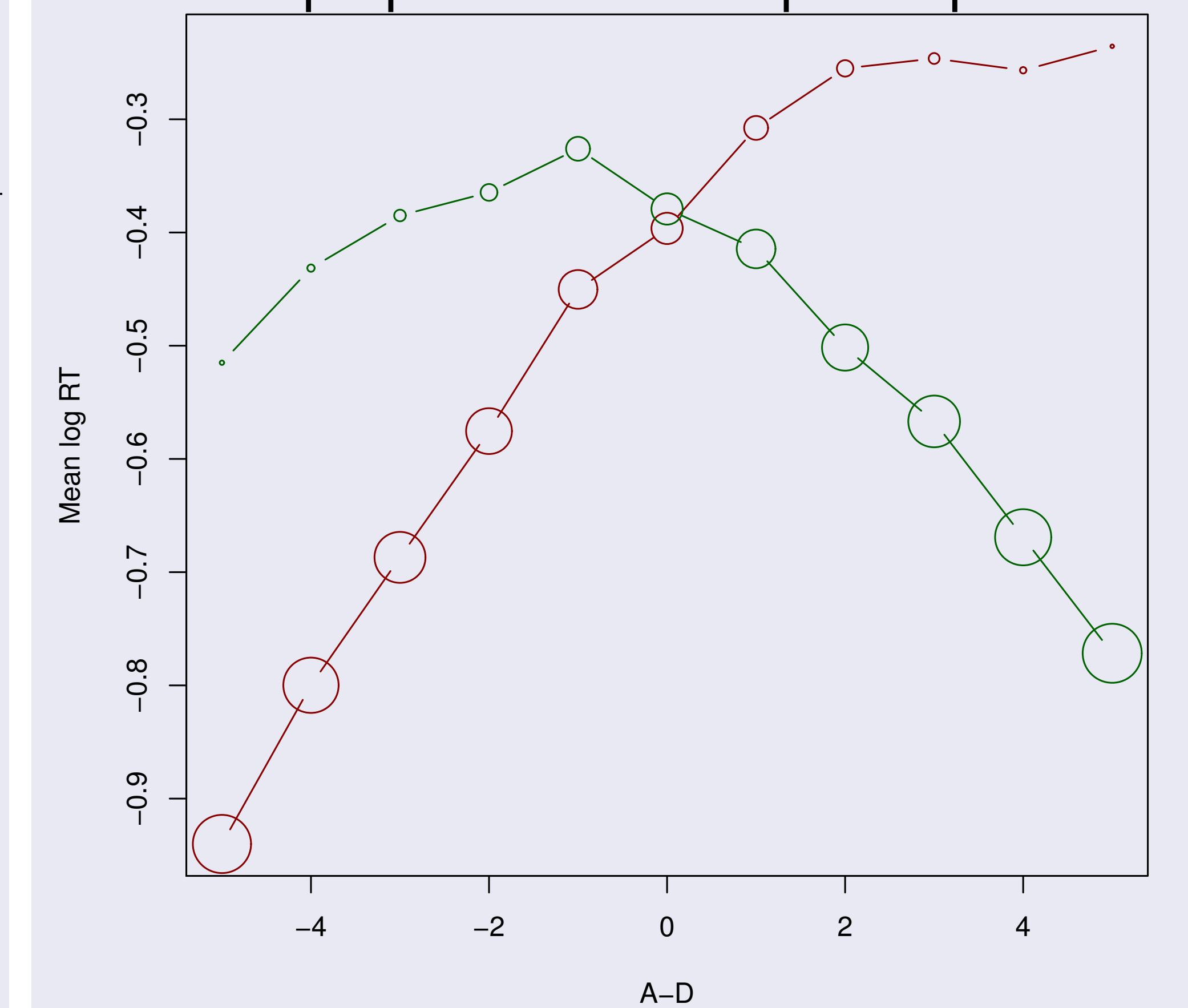


What explains the surprising effect of AD ?

We can get a hint from a plot of RT as a function of $A - D$ and type of response, combining data from all studies.



Let’s see if we can get a drift-diffusion model to produce data that look like this, by simulation. It turns out that we can do this best by assuming that both drift rate (which must be variable from trial to trial) and boundary separation vary with $A - D$. Circle areas are proportional to response probability.



The conclusion about drift rate is reasonable, as this affects pU .

Further analysis shows that the AD effect is entirely due to Ability (mean coefficient .076, $t_{23} = 2.44$, $p = .023$, across the 24 studies), and not at all to Difficulty ($-.001$, $t_{23} = -0.087$).

These results suggest that subjects who make more utilitarian responses are more cautious. Other things being equal (such as drift rate), they take more time on both Yes and No responses.

Reference

Baron, J., Gürçay, B., Moore, A. B., & Starcke, K. (2012). Use of a Rasch model to predict response times to utilitarian moral dilemmas. *Synthese*, 189, Supplement 1, 107–117.