Rule-violations sensitize towards negative and authority-related stimuli

Robert Wirth, Anna Foerster, Hannah Rendel, Wilfried Kunde, & Roland Pfister

Supplementary Material

Pilot Study

The pilot study served to (a) probe for the hypothesized affective component of rule violations, and (b) provide a power estimate for the subsequent experiments. The experiment was identical to the violation group of Experiment 1 in the manuscript, but varied only the valance of the Probe target words (using only words with weak authority-relation).

Methods

Participants. Twenty-four participants were recruited (mean age = 25.4 years, *SD* = 5.8, 11 male, 2 left-handed) and received either course credit or €5 monetary compensation. All participants gave informed consent, were naïve to the purpose of the experiment and were debriefed after the session. Based on previous studies that used a similar design, we estimated the expected effect size as medium (Dreisbach & Fischer, 2012, report $d = 0.52$ for the effect of

negative targets after incongruent trials compared to after congruent trials). Consequently, we chose a sample of 24 participants, as this should provide a power of 0.80 for a medium-sized effect. The observed effect size of the pilot study then served as an estimate to calculate the sample size for the following experiments.

Rating. To arrive at a standardized stimulus set, we asked an independent sample of 15 participants to rate a total of 168 words regarding their valence and their relation to authority. Ratings were given on a nine-point scale with verbal labels at the end points of the scale: (1) extremely negative to (9) extremely positive for valence, and (1) no relation to authority to (9) very strong relation to authority. See the Appendix for the mean ratings of all the Probe target words that were used in the present experiments.

For the pilot study, we selected 6 negative and 6 positive words with low authority-relation from this item pool. The probe words were the German equivalents of *present*, *luck*, *sun*, *peace*, *gain* and *benefit* ($M_{Valence}$ = 7.70, $SD_{Valence}$ = 1.01) for the positive target words and *corpse*, *accident*, *lie*, *bankruptcy*, *betrayal* and *disloyalty* ($M_{Valence}$ = 1.78, $SD_{Valence}$ = 0.30) for the negative words. All these target words were rated lower than 2.5 on the authority scale and were therefore considered weakly related at best. The marked difference in the valence ratings between both types of items, $t(10) = 13.77$, $p < .001$, $d = 9.04$, should, however, allow for easy discrimination between positive and negative words.

Results

***Data selection and analyses*.** For the following analyses, we only used trials from the experimental blocks. We omitted trials in which participants failed to act according to the instruction (Prime: 8.5%, with more commission errors for violations than for rule-based responses, $t(23) = 6.12$, $p < .001$, $d = 0.97$; Probe: 5.8%, irrespective of Probe valence, $t(23)$ =

0.97, p = .341, d = 0.05) and the immediately following trials (Prime: 7.2%, Probe: 5.3%). Trials were discarded as outliers if any of the measures (RT1, RT2) deviated more than 2.5 standard deviations from the respective cell mean (4.7%). RT1 was then analyzed in an analysis of variance (ANOVA) with Prime response (rule-based vs. violation) as within-subjects factor, whereas RT2 was analyzed in a 2×2 ANOVA with Prime response (rule-based vs. violation) and Probe valence (positive vs. negative) as within-subjects factors (see Fig. 4 for the corresponding mean RTs).

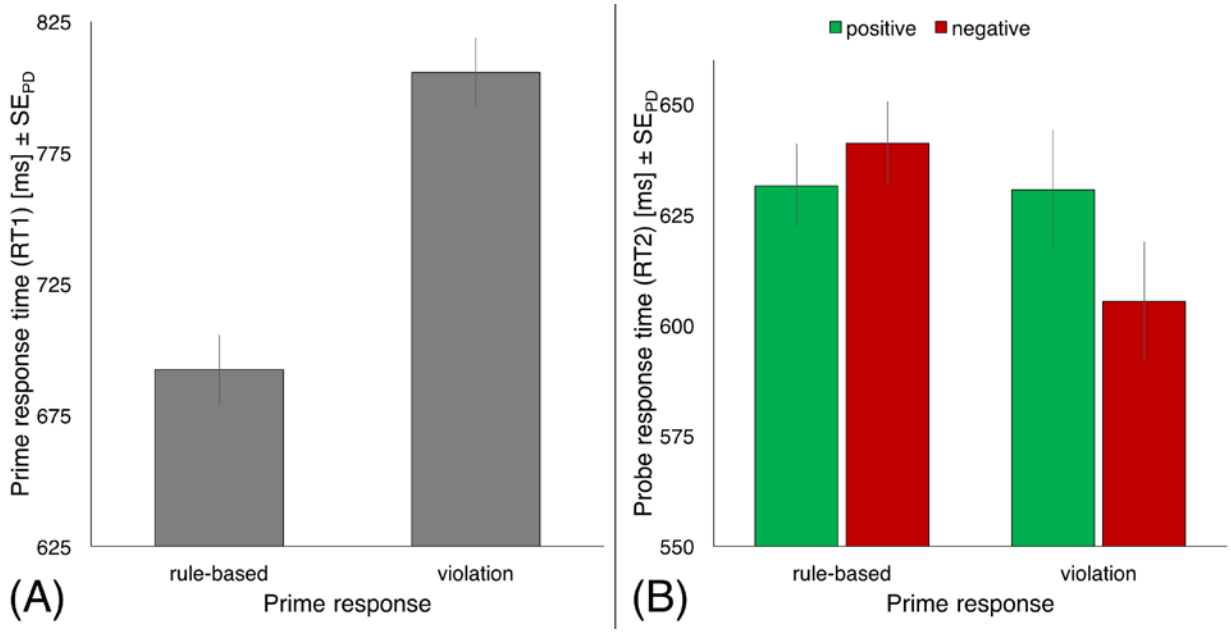

Figure 4. Results for the pilot study. Prime response times (RT1; Panel A) and Probe response times (RT2; Panel B) as a function of Prime response (abscissa) and Probe valence (left, green bars for positive targets; right, red bars for negative targets). Error bars represent standard errors of paired differences that were calculated separately for each instance of Prime response in Panel B (Pfister & Janczyk, 2013).

Prime responses. A significant effect of Prime response emerged, $F(1,23) = 75.74$, $p < .001$, $\eta_p^2 = .77$, driven by slower responses for violations (806ms) than for rule-based responses (692ms, Fig. 4A).

Probe responses. A significant effect of Prime response, $F(1,23) = 6.86$, $p = .015$, $\eta_p^2 = .23$, indicated slower responses after rule-based behavior (636ms) compared to violations (617ms). There was an interaction between Prime response and Probe valence, $F(1,23) = 13.91$, $p = .001$, $\eta_p^2 = .37$ (Fig. 4B), as negative words were evaluated faster after violations relative to rule-based responses ($\Delta = 35$ ms, $t(23) = 3.56$, $p = .002$, $d = 0.73$), and no difference for positive words ($\Delta = 1$ ms, $t(23) = 0.15$, $p = .885$, $d = 0.03$).

Power analysis. The pilot study did not only serve as a confirmation that rule-violations do indeed entail an affective component, but it also served as an estimate for the effect size that could be obtained in this design. We achieved an effect size of $d = 0.76$ for the critical interaction in the Probe task between the Prime response and the Probe valence. This returns a statistical power of .94 when using a corresponding sample size of n = 24. To elevate the power of the experiments in the manuscript above .95, we therefore increased the sample size to 28 participants per group.