

Supplement

The Effects of Logical Status and Believability on Accuracy in Evaluation Tasks

To test the effects of logical status (2: valid, invalid) and believability (2: believable, not believable) on accuracy we conducted a repeated measures ANOVA. Valid believable syllogisms were easiest and solved correctly by 96%, followed by valid not believable that were solved correctly by 66%, then invalid not believable correctly solved by 54% participants, and finally invalid believable solved correctly by 34% of participants.

We registered significant interaction between the two factors ($F(1, 144) = 49.384, p < .001, \eta_p^2 = .255$). Main effect of logical status on accuracy was also significant ($F(1, 144) = 70.396, p < .001, \eta_p^2 = .328$), but the main effect of believability was not ($F(1, 144) = 2.475, p = .118, \eta_p^2 = .017$). Simple main effects analysis showed significant difference in accuracy between valid and invalid *believable* syllogisms ($F(1, 144) = 205.001, p < .001, \eta_p^2 = 0.587$), between believable and not believable *valid* ($F(1, 144) = 50.658, p < .001, \eta_p^2 = 0.260$), and believable and not believable *invalid* syllogisms ($F(1, 144) = 15.130, p = .001, \eta_p^2 = 0.095$). Simple main effect of validity on not believable syllogisms accuracy was not significant ($F(1, 144) = 3.086, p = .081, \eta_p^2 = 0.021$).

The Effect of Logical Status and Believability on Reaction Times in Evaluation Task

To test the effects of these two factors on RTs in evaluation task, we conducted repeated measures ANOVA that showed significant interaction between the two factors ($F(1, 144) = 4.529, p = .035, \eta_p^2 = 0.030$), and significant main effect of believability ($F(1, 144) = 14.355, p < .001, \eta_p^2 = 0.091$), that is – participants took longer time to solve *not* believable syllogisms compared to believable. The main effect of the logical status factor was not significant ($F(1, 144) = 2.504, p = .116, \eta_p^2 = .017$). Post hoc analysis revealed that there was significant simple main effect of believability on reaction times to invalid syllogisms ($F(1, 144) = 19.887, p < .001, \eta_p^2 = 0.121$) syllogisms – not believable took more time to solve. Conversely, simple main effect of believability on valid syllogisms was not significant ($F(1, 144) = 3.103, p = .080, \eta_p^2 = 0.021$). We registered significant simple main effect of logical status on not believable syllogisms ($F(1, 144) = 6.054, p = .015, \eta_p^2 = 0.040$), but insignificant when it comes to believable syllogisms ($F(1, 144) = 0.007, p = .934, \eta_p^2 = 0.000$).

The Effects of Believability, Logical Status, and Word Type on Reaction Times in Lexical Decision Task

To test whether RTs in LDTs differ depending on syllogism believability, logical status, and type of words we conducted a repeated measures three-way ANOVA with factors: logical status (2: valid, invalid) believability (2: believable, not believable), type of words (2: cued, unrelated), with RTs on LDT as the dependent measure.

The interaction between logical status, believability and type of words was significant ($F(1, 144) = 5.111, p = .025, \eta_p^2 = 0.034$). All two-way interactions were insignificant: between type of words and logical status ($F(1, 144) = 0.333, p = .565, \eta_p^2 = 0.002$), between type of words and believability ($F(1, 144) = 0.024, p = .878, \eta_p^2 = 0.000$), and between logical status and believability ($F(1, 144) = 2.529, p = .114, \eta_p^2 = 0.017$). Main effects of logical status ($F(1, 144) = 2.471, p = .118, \eta_p^2 = 0.017$) and believability ($F(1, 144) = 3.879, p = .051, \eta_p^2 = 0.026$) also were not significant, although the effect of believability was extremely close to being significant. The only significant main effect was that of the type of words factor ($F(1, 144) = 74.744, p < .001, \eta_p^2 = 0.342$). Simple main effects of type of words on different levels of remaining factors were all significant (Table E1).

Table E1 Mean differences between reaction times to cued and unrelated words after solving different types of syllogisms

Logical Status	Believability	Mean diff. _{c-u}	SE	<i>p</i>
Valid	Believable	-31.785	5.297	<.001
	Not Believable	-22.758	4.730	<.001
Invalid	Believable	-19.062	5.459	=.001
	Not Believable	-29.506	4.693	<.001

Note. Mean diff. _{c-u} = Mean difference calculated by subtracting RTs for unrelated words from RTs for cued words (cued minus unrelated). SE = Standard Error. *p* = *p* value.