

## Supplemental Material S1. Additional details of response accuracy.

We conducted a series of one-sample t-tests and Bayes factor tests to verify whether participants’ proportion of correct response (i.e., 0-1) was close to ceiling during or after training.

**Table S1.** The descriptive and statistical results of participants’ response accuracy (0-1) of the training session and the two ID posttests.

		<i>M (SD)</i>	<i>t</i>	<i>p</i>	<i>BF<sub>10</sub></i>
Training Initial Blocks					
	HV	0.68 (0.12)	−16.39	< .001	> 100
	LV	0.73 (0.14)	−12.09	< .001	> 100
Training Outcome Blocks					
	HV	0.74 (0.10)	−17.13	< .001	> 100
	LV	0.78 (0.15)	−9.15	< .001	> 100
ID Posttest 1					
	HV	0.67 (0.11)	−22.98	< .001	> 100
	LV	0.65 (0.15)	−17.12	< .001	> 100
ID Posttest 2					
	HV	0.66 (0.14)	−17.75	< .001	> 100
	LV	0.66 (0.13)	−18.20	< .001	> 100

As illustrated in Table S1, the response accuracy of both the groups was significantly different from the ceiling performance (i.e., 1.0 proportion of correct response) at the initial blocks (i.e., first two blocks) and outcome blocks (i.e., final two blocks), indicating that neither group was had ceiling performance *during* training. Likewise, the same tests were conducted for the immediate ID test (i.e., posttest 1) and the 24-hour delayed ID test (i.e., posttest 2). The results suggest that neither group had ceiling performance *after* training.