

S3 Table. Summary of statistical tests from Experiment 2. Random effects are either Group when multiple data points are associated with each group, or Group/Participant (i.e. participant identity is nested in group identity) when multiple data points are associated with each individual participant. ‘Treatment’ refers to the number of black versus total number of sweets in each jar, and ‘Order’ the testing order of each treatment for each group. Significant effects at $p < 0.05$ are marked in bold. Models are arranged in the order in which they are referred to in the main text.

Response variable and sample size (N)	Random intercept	Explanatory variable	LRT _{df}	p value
Absolute error (N = 426)	Group/ Participant	Treatment	20.46 _{2, 417}	<0.001
		Initial or group estimate	13.90 _{1,417}	<0.001
		Order	18.34 _{2,417}	<0.001
Absolute error in group estimate (N = 71)	Group	Treatment	1.79 _{2,62}	0.41
		Order	0.38 _{2,62}	0.83
		Absolute error in mean × Range initial estimates	4.88 _{1,61}	0.027
Absolute % change in group estimate vs. arithmetic mean (N = 71)	Group	Treatment	5.17 _{2,63}	0.075
		Range initial estimates	14.32 _{1,63}	<0.001
		Order	0.46 _{2,63}	0.80
Range initial estimates (N = 71)	Group	Order × Treatment	10.73 _{4,60}	0.030