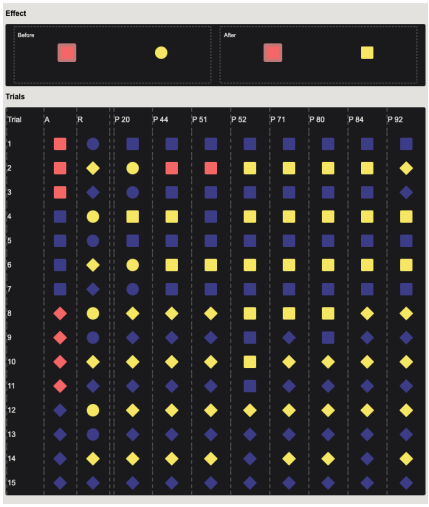
[Magic Stones] Report for Complete Experiment 1

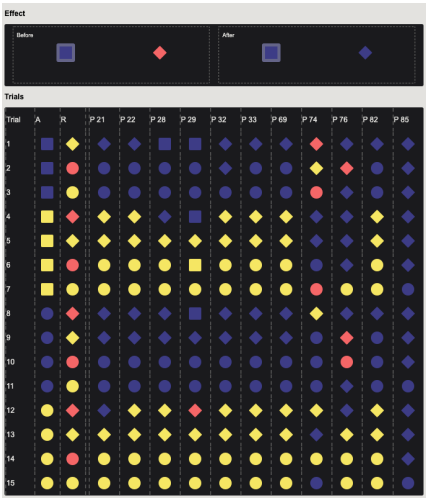
Bonan Zhao (b.zhao@ed.ac.uk)

January 8, 2020

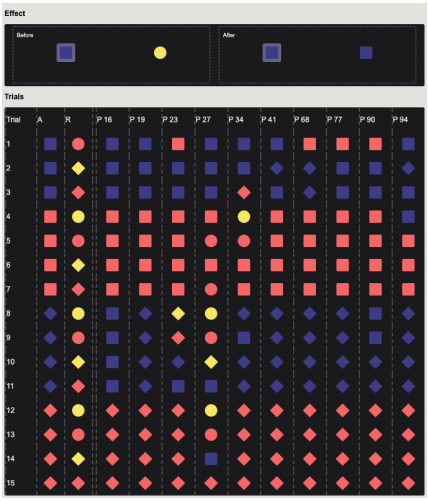
1 Participant data



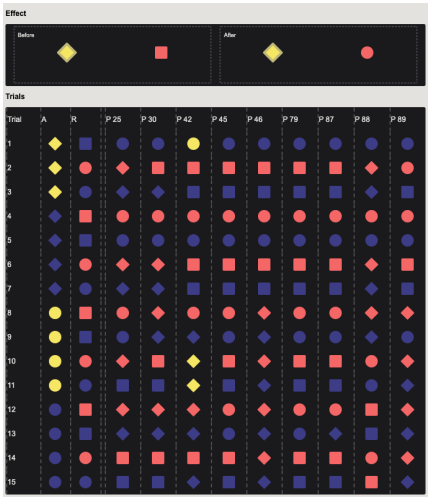
(a) To the same shape



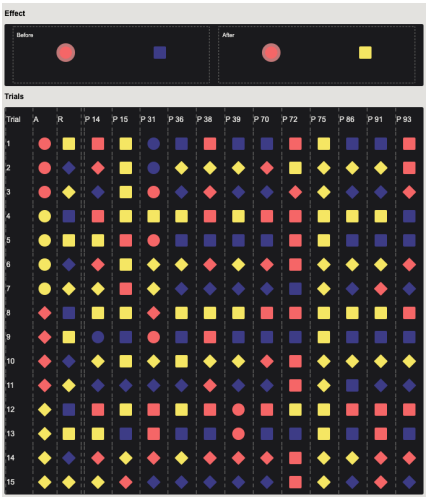
(b) To the same color



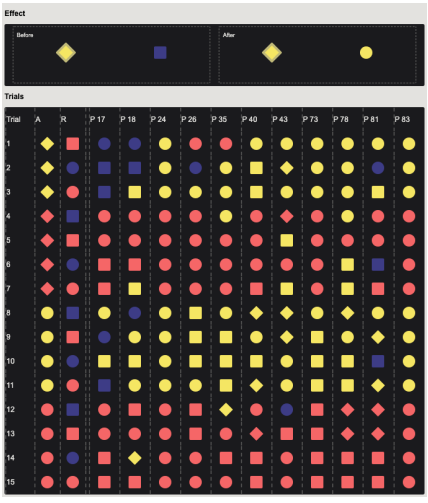
(c) To the same object



(d) To a different shape



(e) To a different color



(f) To a different object

Figure 1: Each figure is for one learning condition. Each learning condition has 15 trials (15 rows).

2 Stats

- Age: min 24, max 67, mean 41.1639, sd 11.1238
- Gender: female 28 (45.9%), male 33 (54.1%)

Condition	Description	Count	Share	Task dur. mean (min.)	S-Dfty. mean	Variation
1	To the same shape	8	13.11%	5.1873	3.38	10.83%
2	To a different shape	9	14.75%	5.0942	7.89	18.33%
3	To the same color	11	18.03%	4.7480	3.45	22.50%
4	To a different color	12	19.67%	4.1504	5.17	38.33%
5	To a different object	11	18.03%	5.2803	5.09	30.00%
6	To the same object	10	16.39%	5.0351	3.00	16.67%
Total		61		4.8756 (sd 1.87)	4.65 (sd 3.16)	

Table 1: Basic stats per condition. S-Dfty.: Self-report difficulty.

Variation

Variation is defined as how much observed difference out of the maximal possible difference. For a condition C and each of its generalization task i, let U_i be the number of unique predictions all participants made in this task, and n be the number of participants, *observed difference* is defined as $U_i - 1$ because of the intuition that if all the participants make the same prediction, the number of unique selections is 1 and variation should be 0, and *maximal difference* M_i is given by

$$M_i = \begin{cases} 8 & \text{if } n \geq 9 \\ n & \text{otherwise} \end{cases}$$

Formally, variation measure V_C for condition C is

$$V_C := (\sum_{i \in C} U_i - 1) / \sum_{i \in C} M_i$$

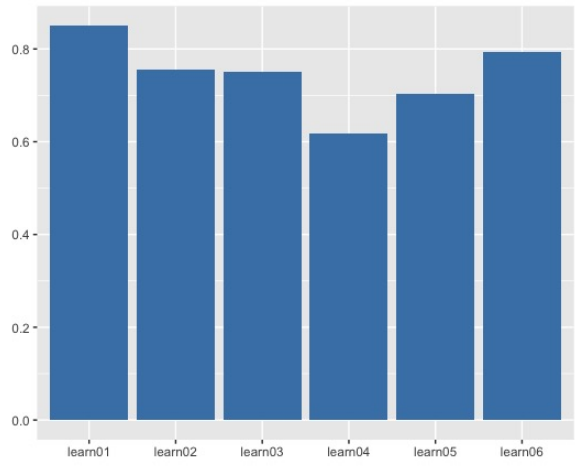
.

Overall, participants make quite homogeneous predictions - variation measure for all the conditions are below 40% of the largest variations, with condition 4 (to a different color) being the one with most various predictions. For condition 1 (to the same shape), participants make very uniform predictions, with a variation measure of just 10.84%.

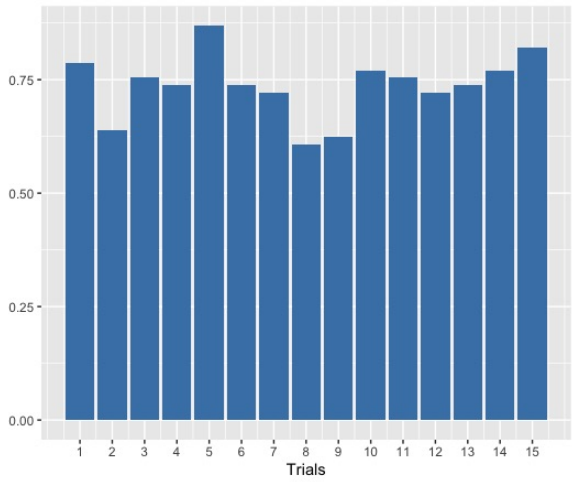
"To the same xx" V.S. "To a different xx"

Conditions 1, 3, and 6 can be classified as "to the same xx" group, where xx can be color, shape, or both (equivalent to the object), and conditions 2, 4, 5 can be classified as "to a different xx" group. Statistical test shows that compared to the "to a different xx" group, participant report the "to the same xx" group is significantly easier ($p < 0.0001$), and make more homogeneous predictions ($p = 0.0001$). However there is so significant difference for task duration between these two groups.

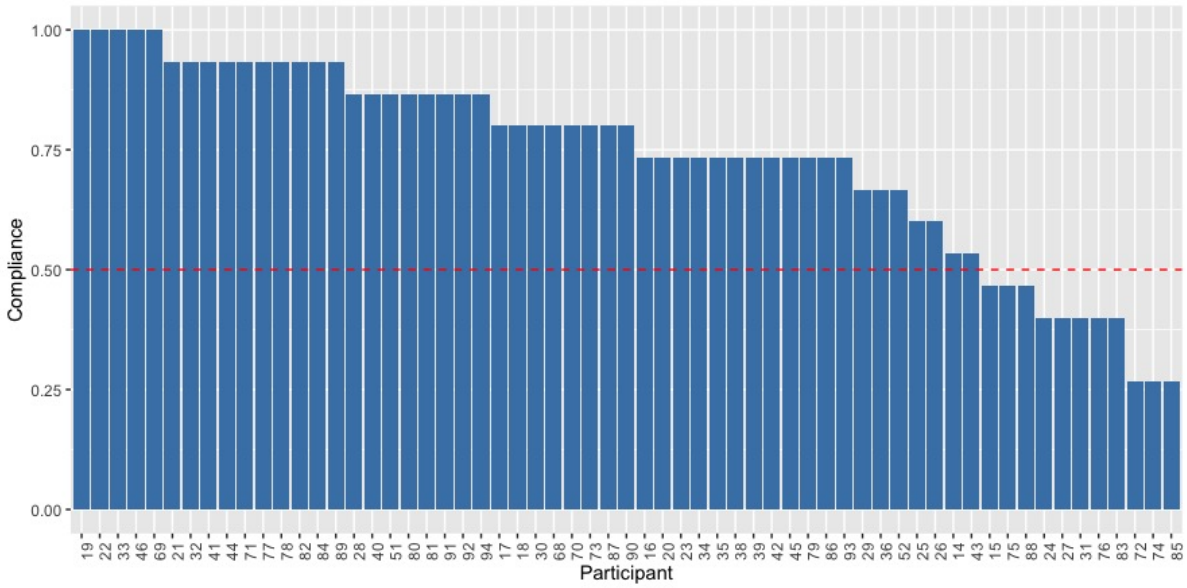
3 Compliance



(a) Aggregated by conditions.



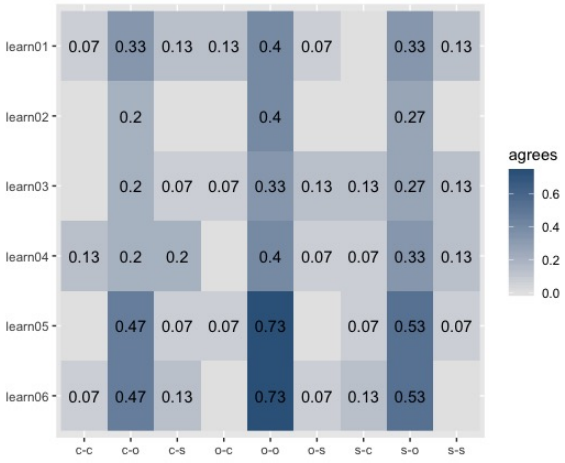
(b) Aggregated by tasks.



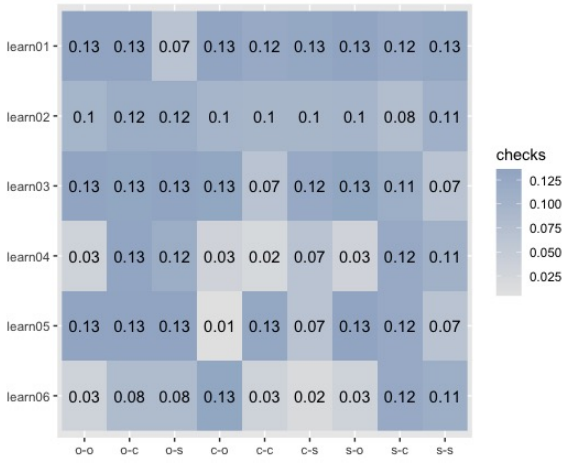
(c) Individuals

Figure 2: Percent of selections that comply with the underlying theory.

4 Theories



(a) With code setup



(b) With participant selections