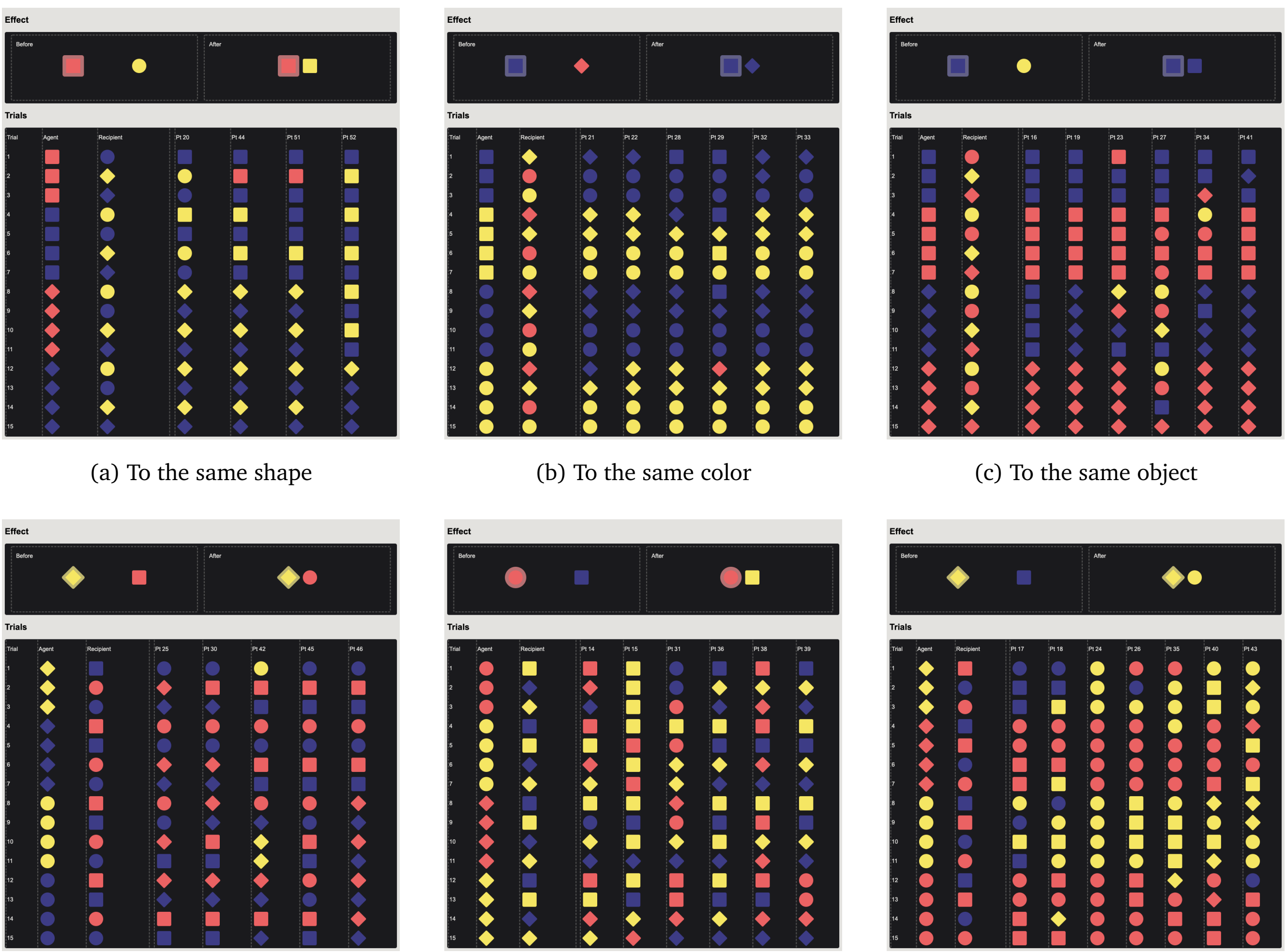
[Magic Stones] Plots for Experiment 1

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

January 5, 2020

1 Participant data

1.1 Raw 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: Participant data per group per trial. Each figure is for one group - at the top is a summary of the magic effect participants watched, and below is a complete visualization of participant selections. In the selection part, each row is for one trial: the first two icons are the magic stone and normal stone that a participant is asked to make a prediction for, and the rest are participant selections.

Obviously, generalizations are not random, as shown in Figure 2. Take the third column from left for example, a dominant selection can be observed across 15 trials.

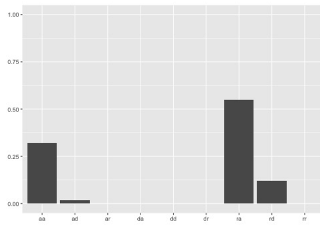
Figure 2: Participant raw selections summary. Each row is for one trial, and each column is for one task. For each sub-figure, y-axis is percentage, and x-axis from left to right is: bc, bd, bs, rc, rd, rs, yc, yd, ys.

If we plot each participant selection relative to the agent stone and recipient stone in that generalization task, the uniformity of generalization is even clearer. In the last raw of Figure 3, almost everyone made a same choice across 6 different learning tasks.

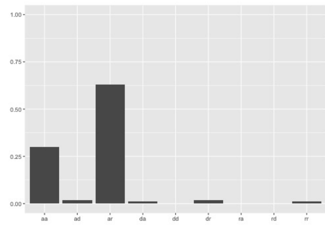
Figure 3: Relative selections summary. Each row is for one trial, and each column is for one task. For each sub-figure, y-axis is percentage, and x-axis from left to right is: aa, ad, ar, da, dd, dr, ra, rd, rr.

1.2 Aggregated by relative selection

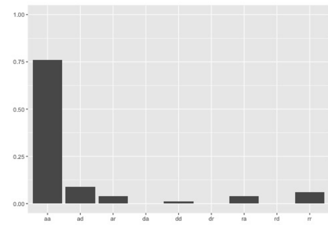
Between-subject aggregation shows that different learning tasks induce different generalization patterns, as illustrated by Figure 4.



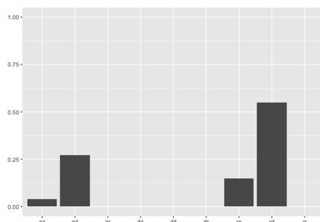
(a) To the same shape



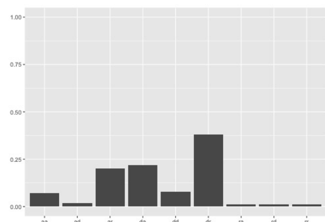
(b) To the same color



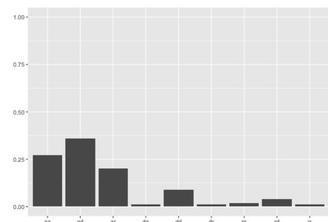
(c) To the same object



(d) To a different shape



(e) To a different color



(f) To a different object

Figure 4: Aggregated relative selections per task. For each sub-figure, y-axis is percentage, and x-axis from left to right is: aa, ad, ar, da, dd, dr, ra, rd, rr.

Aggregation by trials (within-subject) emphasizes that while the generalization scenario further differs from the learning scenario, participants tend to make a more similar decision.

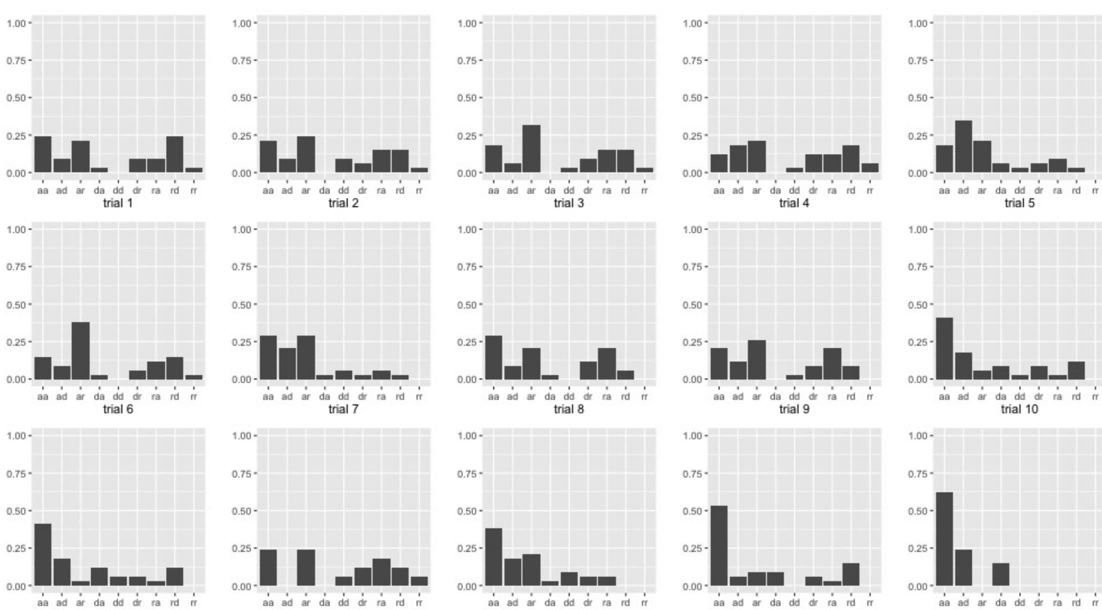
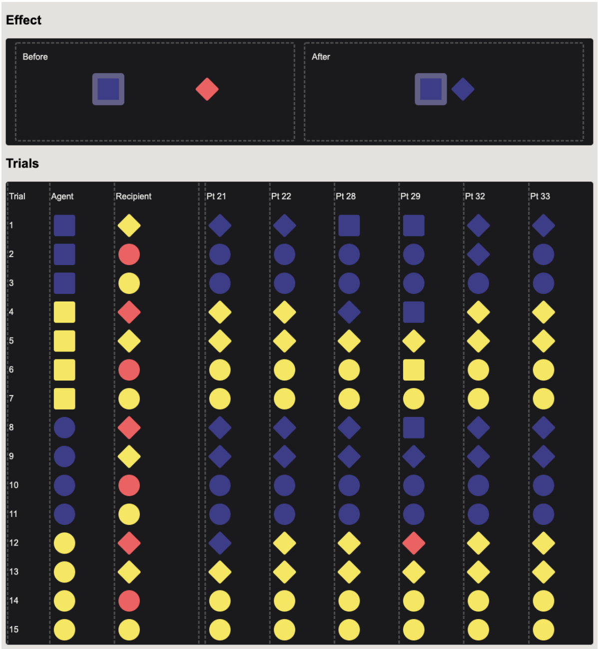


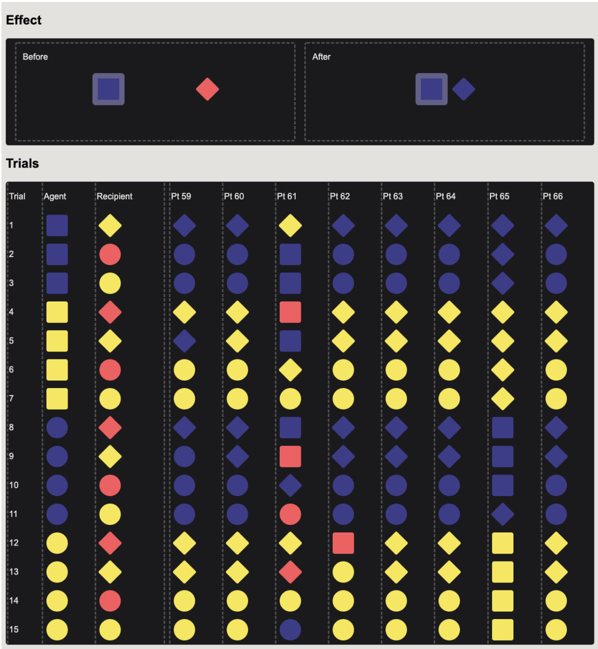
Figure 5: Aggregated relative selections per trial. For each sub-figure, y-axis is percentage, and x-axis from left to right is: aa, ad, ar, da, dd, dr, ra, rd, rr.

1.3 Order effects? Not really

Comparing a randomized sequence of trials and ordered sequence of trials shows no significant difference.



(a) Task 3 ordered.



(b) Task 3 randomized.

Figure 6: Order does not effect generalization predictions.

2 Normative model simulations

2.1 Raw simulations

Different from the participant data, the normative model is more uncertain when the generalization scenario differs more from the learning task, as shown in the last few columns in Figure 7. In the same situation, humans tend to concentrate on one selection instead of choosing randomly.

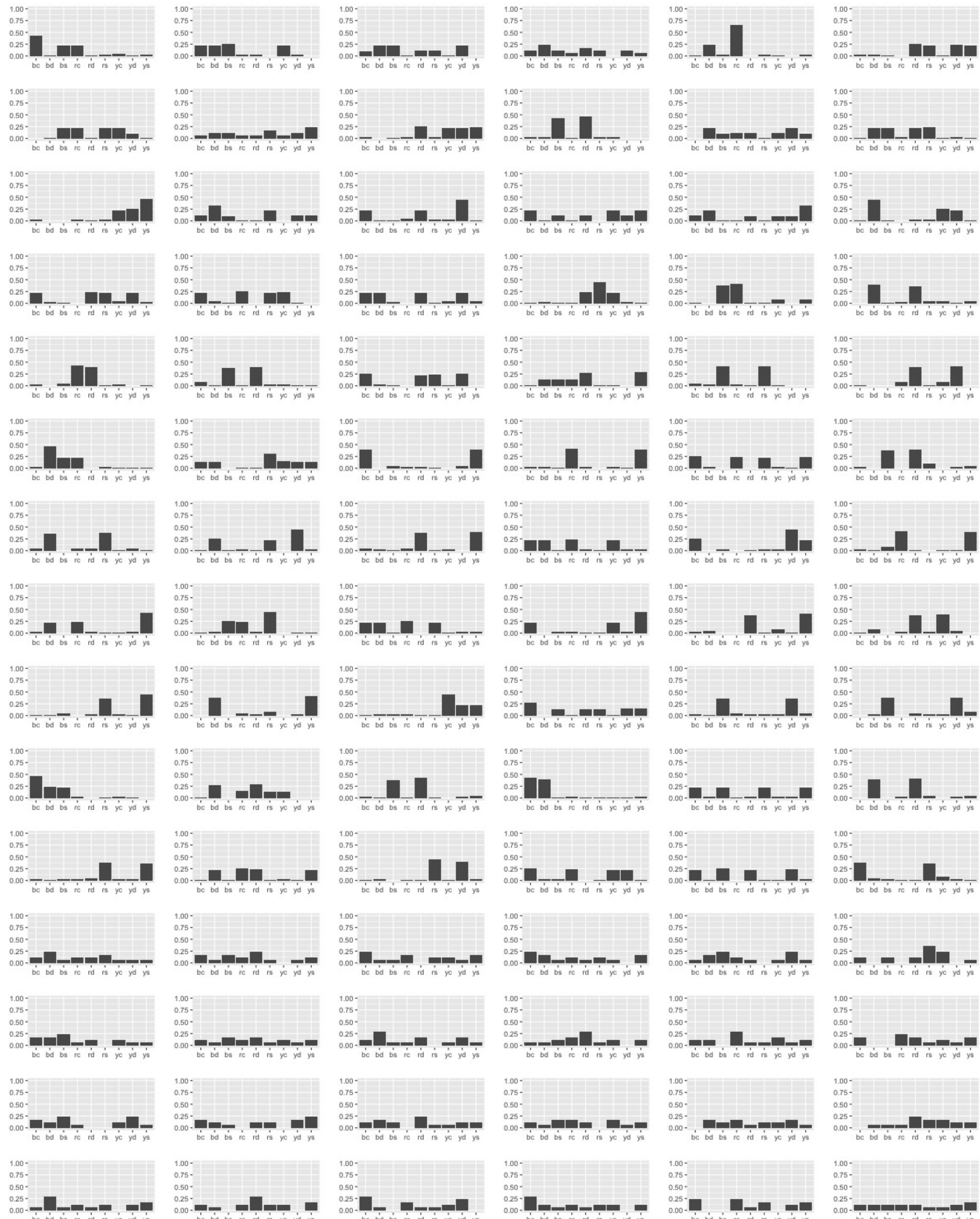


Figure 7: Simulation raw selections summary. Each row is for one trial, and each column is for one task. For each sub-figure, y-axis is percentage, and x-axis from left to right is: bc, bd, bs, rc, rd, rs, yc, yd, ys.

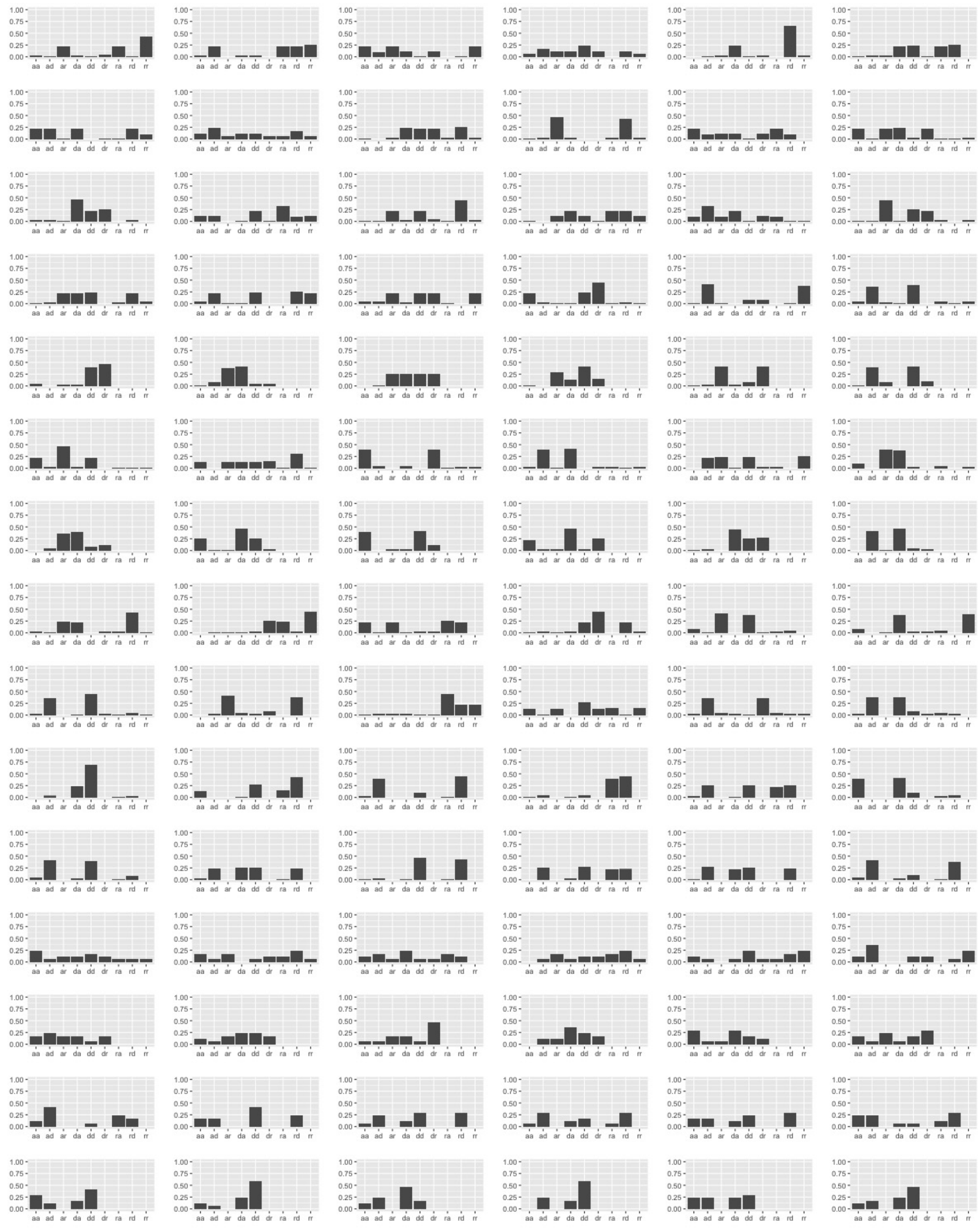
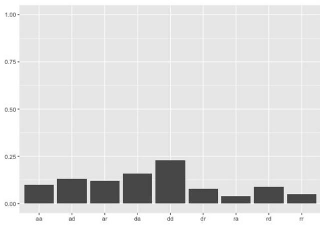


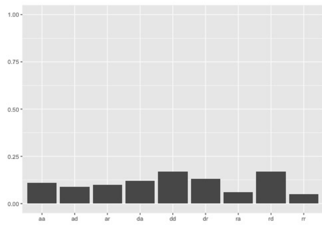
Figure 8: Relative selections summary. Each row is for one trial, and each column is for one task. For each sub-figure, y-axis is percentage, and x-axis from left to right is: aa, ad, ar, da, dd, dr, ra, rd, rr.

2.2 Aggregated by relative selections

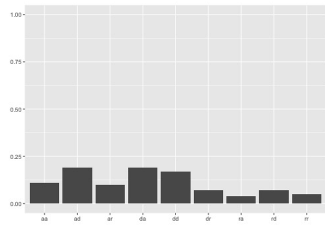
The normative model does not show specific patterns towards each learning task (Figure 9). It does show a tendency to concentrate on one selection as the generalization scenarios further differs from the learning scenarios (last figure in Figure 10), but it predicts a different concentration point compared with participant data.



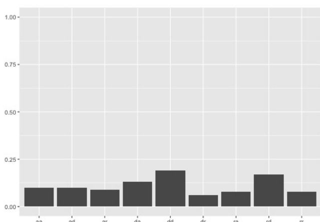
(a) To the same shape



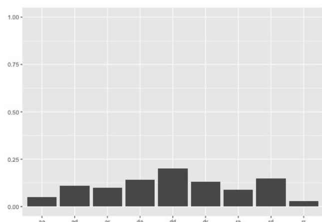
(b) To the same color



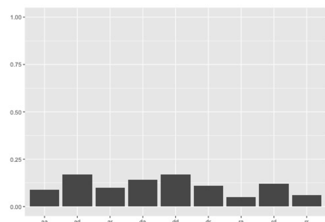
(c) To the same object



(d) To a different shape



(e) To a different color



(f) To a different object

Figure 9: Aggregated relative selections per task. For each sub-figure, y-axis is percentage, and x-axis from left to right is: aa, ad, ar, da, dd, dr, ra, rd, rr.

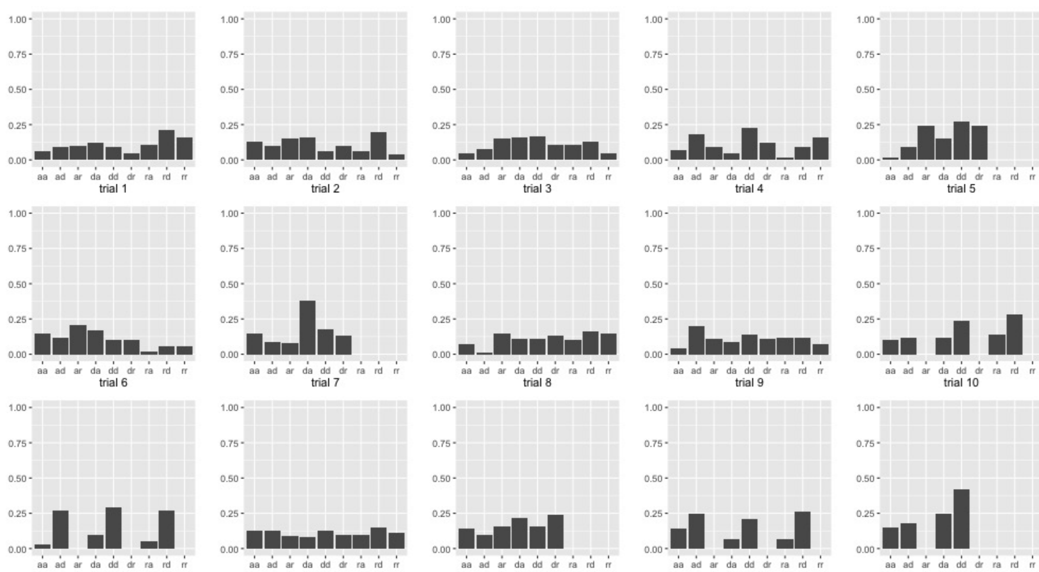
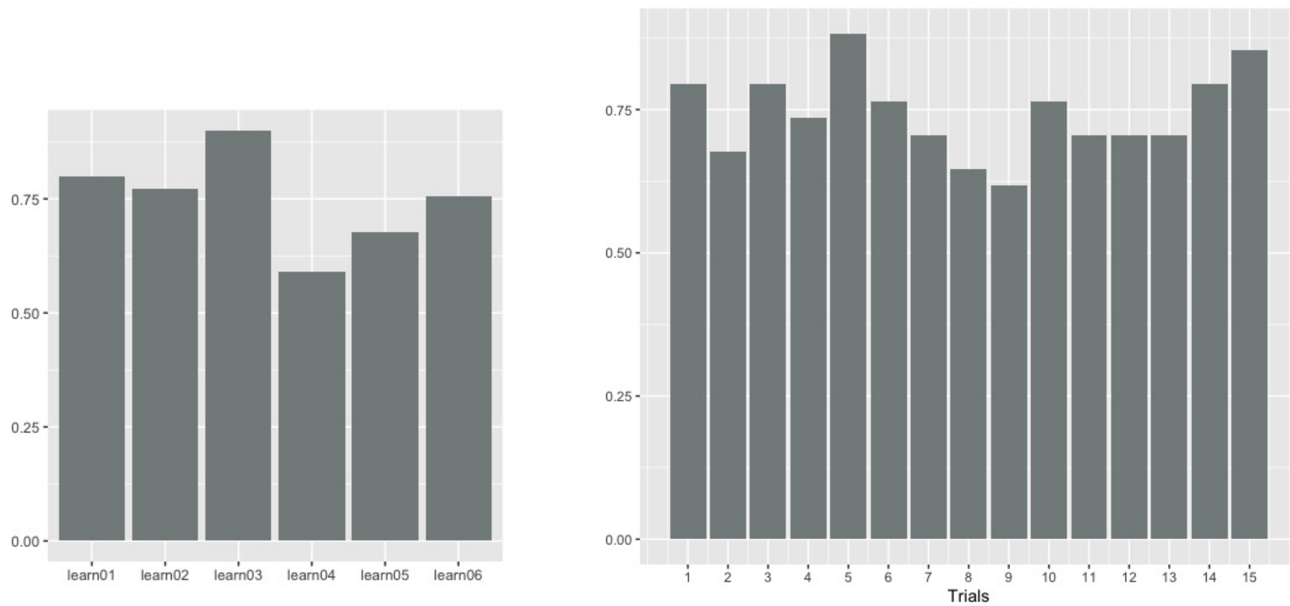


Figure 10: Aggregated relative selections per trial. For each sub-figure, y-axis is percentage, and x-axis from left to right is: aa, ad, ar, da, dd, dr, ra, rd, rr.

3 Compliance with theories

- learn01: to the same shape as magic stone (and kept color unchanged)
- learn02: to a different shape as magic stone (and kept color unchanged)
- learn03: to the same color as magic stone (and kept shape unchanged)
- learn04: to a different color as magic stone (and kept shape unchanged)
- learn05: To a different object
- learn06: To the same object as the magic stone

3.1 Total compliance



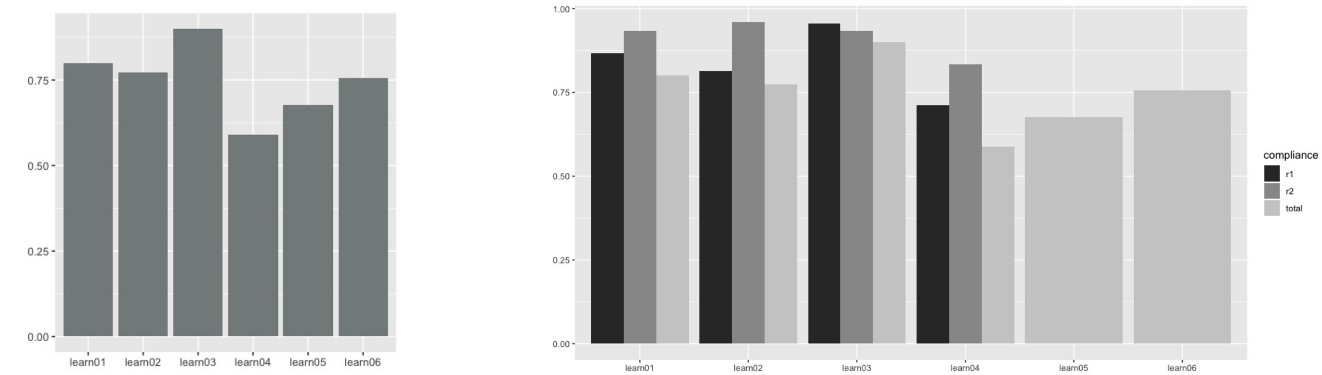
(a) Per learning condition

(b) Per generalization trial

Figure 11: Percent of selections that comply with the underlying theory, aggregated per learning condition and per generalization trial separately.

3.2 Details

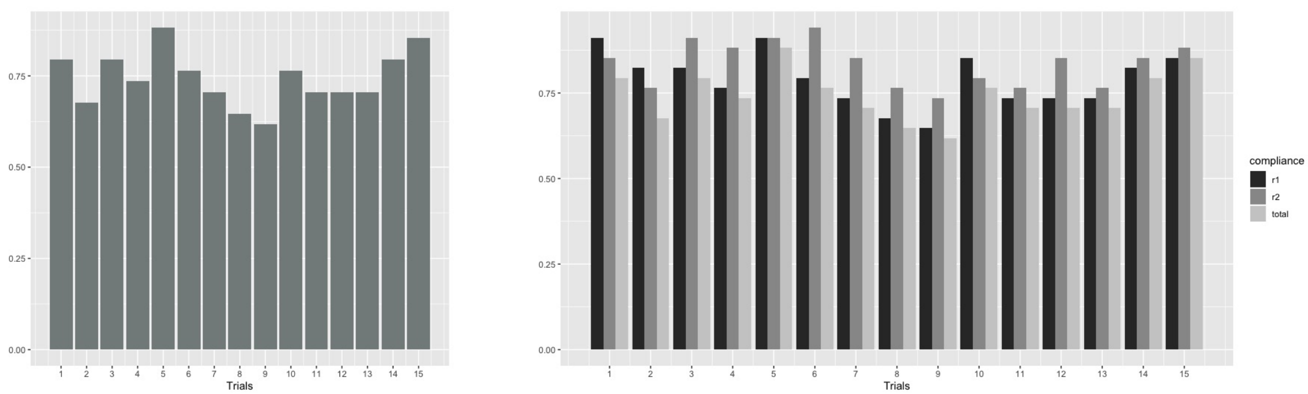
For learn01 - learn04, take the description outside of the parenthesis as ‘r1’, and description inside the parenthesis as ‘r2’, we can compare to what extent participants comply with such theories.



(a) Overall compliance

(b) Split by sub-rule compliance

Figure 12: Percent of selections that comply with the underlying theory, aggregated per learning condition.



(a) Overall compliance

(b) Split by sub-rule compliance

Figure 13: Percent of selections that comply with the underlying theory, aggregated per generalization trial.

3.3 Individual

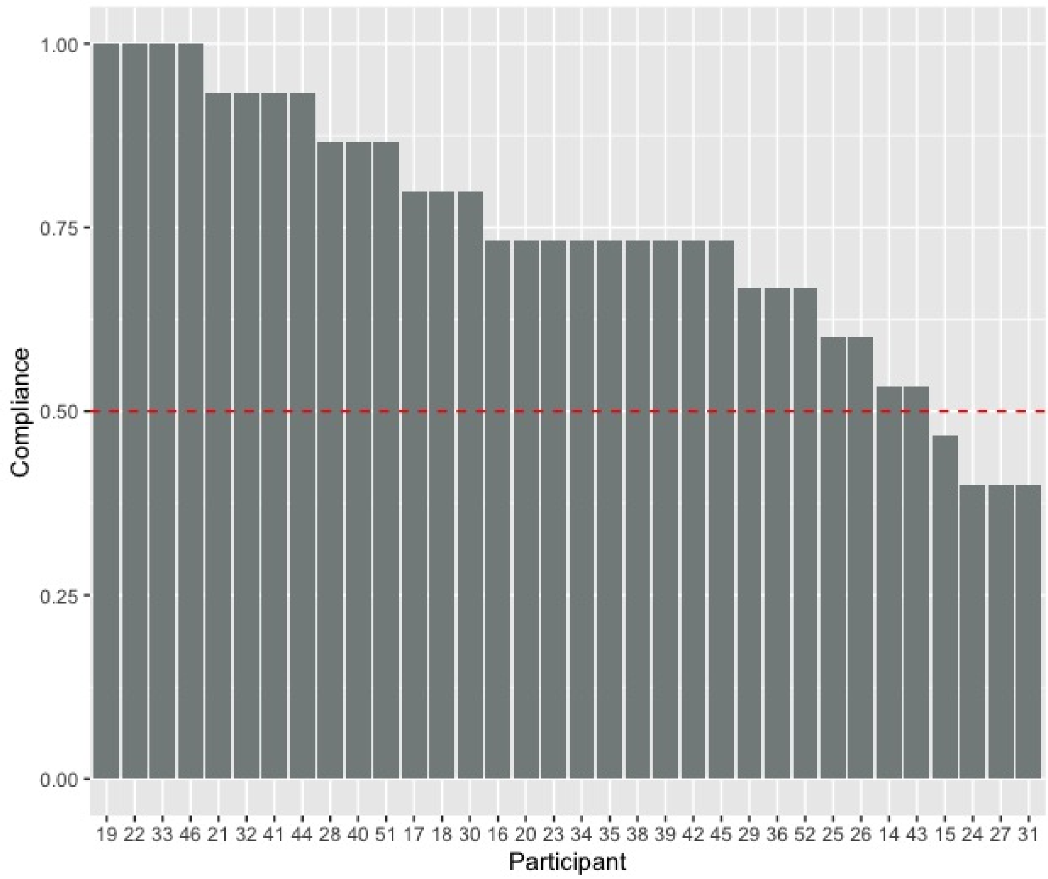


Figure 14: Compliance per participant.

4 Homogeneity

To measure how participant selections vary across learning conditions, define trial i's variation $v_i := (U-1)/n$ where U is the number of unique selections from all participants for trial i, and n is the number of participants in this trial. We take $U-1$ because of the intuition that if all participants select the same stone, the number of unique selections is 1 and variation is 0. Variation V_J for a learning condition J is therefore defined as $V_j := \sum_{i \in J} v_i$, for all generalization trial i under learning condition J.

Variations for the six above mentioned learning conditions are as follows

Learning condition	Theory	Variation
1	To the same shape	2.75
2	To a different shape	3.00
3	To the same color	1.33
4	To a different color	5.16
5	To a different object	4.57
6	To the same object	3.00

5 Notes

1. To the same feature is easier than to a different one
2. Shape change is more uniform in generalization

6 Compare Nine Theories

Let y be a theory from the nine theories ($object \to object$, $color \to shape$, etc), t be a generalization trial for a given learning condition (each learning condition has 15 trials), $\mathbf{n}_t$ be the normative selection(s) for trial t, $\mathbf{s}_t^y$ be the predictions given by theory y, and $\mathbf{d}_t$ be the predictions made by participants. Measure level of agreement between a theory y for a learning condition and

- Normative selection:

$$\sum_t \frac{Am(\mathbf{n}_t, \mathbf{s}_t)}{15}$$

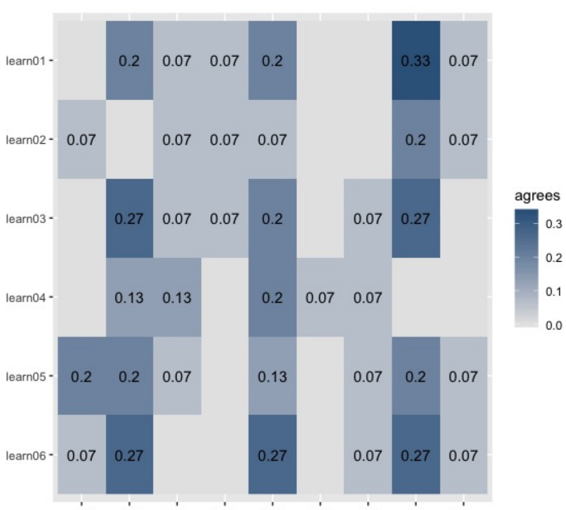
  where

$$Am(\mathbf{n}_t, \mathbf{s}_t) = \begin{cases} 1 & \text{if } s_t^* \in \mathbf{n}_t \\ 0 & \text{otherwise} \end{cases}$$

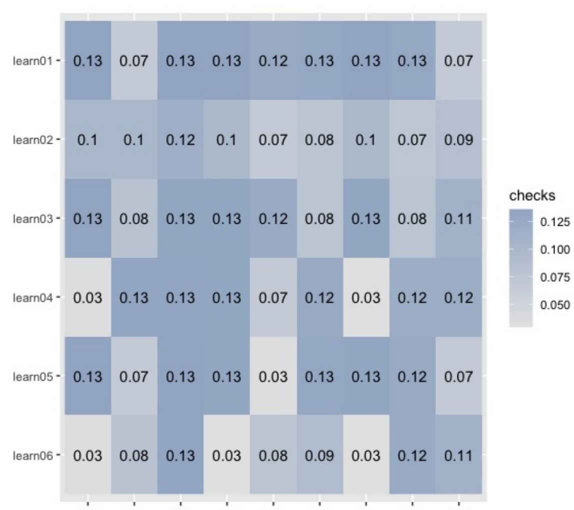
  $P(s_t^*) >= P(s)$ for all $s \in \mathbf{s}_t$.

- Participant selection:

$$\sum_t \Big(\sum_{d_i \in \mathbf{d}_t}^{s_i \in \mathbf{s}_t} (|P(d_i) - P(s_i)|) \Big) / 15$$



(a) With normative selections



(b) With participant selections