Spontaneous categorization along competence. . . - Study 4 experimental report

Alexander Bor

2016-06-15

Hypothesis[1]

What question was the experiment designed to address?

How information on domain-specific competence is used for categorization in social interactions?

What are the specific hypotheses to be tested?

1. Whether there is significant categorization along various levels (high vs low) of foraging skills.
2. Whether priming leader evaluation psychology (vs peer evaluation psychology) increases the extent of encoding with multiple traits present.

Subjects and Context

Alltogether 437 participants entered the survey. 33 people were not allowed to participate because their IP address was detected to be from outside of the US and 3 participants have not finished the entire survey, leaving us with 401 participants (M_{age}= 38.5, N_{female} = 233).

Eligibility and exclusion criteria for participants

- **Why was this subject pool selected?** MTURK offers an easily available, cheap and relatively diverse subject pool.
- **Who was eligible to participate in the study?** Registered MTurkers living in the US, success rate min 97%, and minimum 1000 completed HITs. Participants of Studies 1, 2 and 3 and the pre-tests were *not* eligible to participate.
- **What would result in the exclusion of a participant?** All participants who completed the study are included in the analysis.
- **Were any aspects of recruitment changed (such as the exclusion criteria) after recruitment began?** No.

Procedures used to recruit and select participants

Incentives

- **What incentives were offered and how were they administered?** Participants were offered $0.5, which they received if they submitted the unique survey code displayed at the end of the assignment.

[1]This report is based on Gerber, Alan, Kevin Arceneaux, Cheryl Boudreau, Conor Dowling, Sunshine Hillygus, Thomas Palfrey, Daniel R. Biggers, and David J. Hendry. 2014. "Reporting Guidelines for Experimental Research: A Report from the Experimental Research Section Standards Committee." Journal of Experimental Political Science 1 (01): 81-98.

Figure 1: "Illustration"

When was the data collected?

Data was collected on May 4-5, 2016. There were no follow ups.

Allocation method

What was the randomization procedure?

Participants were randomly assigned by Qualtrics to the peer frame group or the leader frame group.

Procedure & Treatments

A description of the procedure.

1. Participants were introduced to the task and were asked to provide the MTURK IDs.
2. Participants then read the cover story about people on the island and were instructed to pay attention to the actions presented next.

3. Participants were presented with the 8 targets with each picture-action pair presented for 20 seconds.
4. In a distractor task, participants were asked to list as many countries as they could in 60 seconds.
5. Participants indicated who they would prefer as their partner / leader.
6. In a surprise recall, participants were asked which participant did each of the actions (sentences).
7. Finally participants provided basic demographic information (age & gender).

Description of interventions.

Introduction

"In today's study you'll be learning about a group of people who were traveling together on a small chartered plane, which hit a violent storm and damaged the electrical system as well as one of the engines. Luckily the pilot managed to crash land on a small island, but he died during the impact and many of the passengers were seriously injured. For two days the passengers waited by the plane for help. They eventually ate all the food they had.

Realizing that they might be on the island for a while and that they needed supplies, those who were not injured decided to work together and cooperate so that everyone, including those who were seriously injured, would survive. They all agreed that they had to collect food and share it with the entire group. In order to cover more ground, each person went out on their own.

/ All the survivors agreed that in order to coordinate their efforts and resolve potential conflicts they need to elect a leader. This is planned to take place next evening by the campfire with everyone present.

You will be presented with passengers one at a time. We will ask you to look at pictures of these persons and to read about their actions. **Try to gain an impression of who you would prefer to spend time with from the group! / Try to gain an impression of who you would vote for to become the leader of the group!**

Each portrait will be shown for 15 seconds, and the next portrait will come up automatically. So you do not have to press any keys during the introduction of these people. Just let yourself form an impression of them.

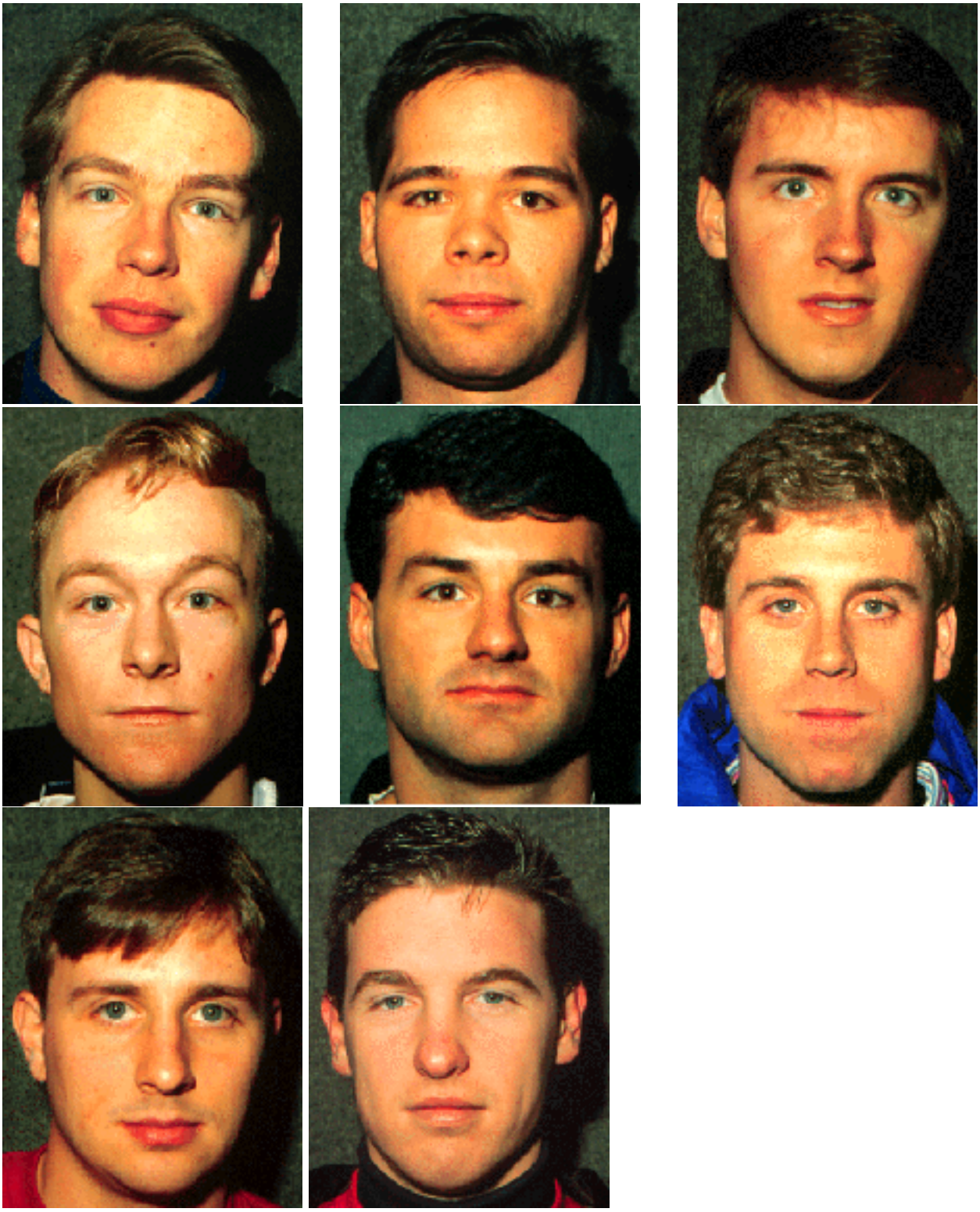
Click the >> button when you are ready."

Sentences:

Participants were presented the following diagnostic sentences. Each target was described engaging in two actions one revealing competence and the second revealing likability. Variable names are defined the following way: "sen02np" refers to the second sentence (sen02), signalling low (n - "negative") competence and high (p - "positive") likability.

- sen1nn After someone else dropped a great many juicy oranges into a cave, he ventured into the dark but got scared before finding the fruits. He refused to join a small game back at the camp, saying that it was "just silly".
- sen2np Hunting for a flock of duck, he accidentally fell into a river and gave up his efforts. In the afternoon, several peers shared their problems with him due to his easy-going style.
- sen3pn At the very edge of a quickly moving waterfall, he speared the twenty or thirty snapper he saw there and took them to camp. Some of his peers were annoyed by his constant pessimistic remarks during social activities.
- sen4pp He scaled the sheer face of a tall vertical mountain and gathered for the group the tons of pears he saw growing there. His friendliness and sincerity was much appreciated in talks around the fire in the evening.
- sen5np In an attempt to collect many cups worth of honey, he clumsily moved through a dense cloud of angry bees, startled them, was attacked by the bees and ran away without any honey. His general optimism elevated the mood of the others on the island.
- sen6pn Moving from precarious tree-top to precarious tree-top, he collected for the group many bunches of yellow bananas he had seen from the ground. He came across as a cold fish in discussions in the afternoon.
- sen7nn Even after cougars had been seen by the pineapple trees, he went there to gather the copious armfuls of fruit he had seen there earlier, but he mistook a small rodent for a predator, ran away and was too embarrassed to return. In an argument around the fire, he got annoyed and refused to agree with the others.
- sen8pp Seeing numerous large lobsters there, he swam far out into treacherous open waters and collected them to bring to camp. In the afternoon, he entertained a group of peers with his good sense of humour.

Target photographs



How and when manipulations or interventions were administered?

The following parameters were randomly presented:

- Participants were randomly assigned to the Peer frame group or the Leader frame group
- 8 sentences were presented in two balanced sequences.
- The specific order of the sentences was balanced and pseudo-randomized. The order was either the one presented above (sen 1-8) or an alternative which inversed the valence order. This sequence is: sen04nn, sen03pn, sen02np, sen01nn, sen08pp, sen07nn, sen06pn, sen05np.
- The order of the target pictures (faces) was randomized.
- At the surprise recall phase, the order of the sentences and the appearance of the target faces was randomized.

	Control Group	Treatment Group
N	196	208
Mean age	37.1	39.8
Share of women (%)	56.9	58.9

	Sequence 1	Sequence 2 - inverse
N	196	208
Mean age	38	39
Share of women (%)	57.3	58.7

The likability sentences were pre-tested to ensure that on average targets performing a likable action are considered to be more competent. The mean ratings show that the manipulations are effective $M_{likable} = 5.6130952$, $M_{unlikable} = 2.708$, $p < 0.0001$.

- **Method of delivery:** Participants completed the survey implemented in Qualtrics on their own computers.

Time span

- **How long did each experiment last?** According to Qualtrics estimate on average 7.8 minutes.

Results

Outcome Measures and Covariates

sen1 - sen8 In the surprise recall, each sentence was presented with the question: "Who did this: . . . "

gender What is your gender? Male / Female.
age What is your age?

Categorization scores

Each answer in the surprise recall was categorized as correct, a within category error or a between category error. Following standards the number of between category errors was multiplied by 0.75 to correct for different base rates. A final categorization score is defined as the difference between within category and between category errors with positive numbers indicating evidence for more categorization.

The following statistical tests were planned:

- Are the categorization scores significantly different from 0 for each group?

- Are categorization scores significantly different from each other in the control and treatment groups?

CONSORT Participant Flow Diagram

No relevant information beyond those shared under "Subjects and context".

Statistical Analysis

5.2.1. Categorization along competence and likability

Is the competence categorization score significantly different from 0 in the whole sample?

Table 3: One Sample t-test: `data$comp_catscore`

Test statistic	df	P value	Alternative hypothesis
3.09	401	0.00215 * *	two.sided

```
## Effect size:

## [1] 0.1524343
```

Is the competence categorization score significantly different from 0 in the peer frame group?

Table 4: One Sample t-test: `data$comp_catscore[data$treat == "Peer prime"]`

Test statistic	df	P value	Alternative hypothesis
1.74	206	0.0828	two.sided

```
## Effect size:

## [1] 0.1205534
```

Is the competence categorization score significantly different from 0 in the leader frame group?

Table 5: One Sample t-test: `data$comp_catscore[data$treat == "Lead prime"]`

Test statistic	df	P value	Alternative hypothesis
2.57	194	0.0108 *	two.sided

```
## competence Effect size:

## [1] 0.1816587
```

Does the treatment increase competence categorization?

Table 6: Welch Two Sample t-test: `data$comp_catscore` by `data$treat`

Test statistic	df	P value	Alternative hypothesis
0.907	378	0.365	two.sided

```
## [1] 0.04658907
```

Is the likability categorization score significantly different from 0 in the whole sample?

Table 7: One Sample t-test: `data$warm_catscore`

Test statistic	df	P value	Alternative hypothesis
2	401	0.0466 *	two.sided

```
## Effect size:

## [1] 0.09917539
```

Is the likability categorization score significantly different from 0 in the peer frame group?

Table 8: One Sample t-test: `data$warm_catscore[data$treat == "Peer prime"]`

Test statistic	df	P value	Alternative hypothesis
2.07	206	0.0397 *	two.sided

```
## Effect size:

## [1] 0.1427684
```

Is the likability categorization score significantly different from 0 in the leader frame group?

Table 9: One Sample t-test: `data$warm_catscore[data$treat == "Lead prime"]`

Test statistic	df	P value	Alternative hypothesis
0.763	194	0.446	two.sided

```
## Effect size:

## [1] 0.05472193
```

Does the treatment increase likability categorization?

Table 10: Welch Two Sample t-test: `data$warm_catscore` by `data$treat`

Test statistic	df	P value	Alternative hypothesis
-0.845	395	0.398	two.sided

```
## [1] 0.04247352
```

In the pooled sample, is categorization by competence significantly higher than by likability?

Table 11: Paired t-test: `data$warm_catscore` and `data$comp_catscore`

Test statistic	df	P value	Alternative hypothesis
-0.745	401	0.457	two.sided

In the leader frame, is categorization by competence significantly higher than by likability?

Table 12: Paired t-test: `data$warm_catscore[data$treat == "Lead prime"]` and `data$comp_catscore[data$treat == "Lead prime"]`

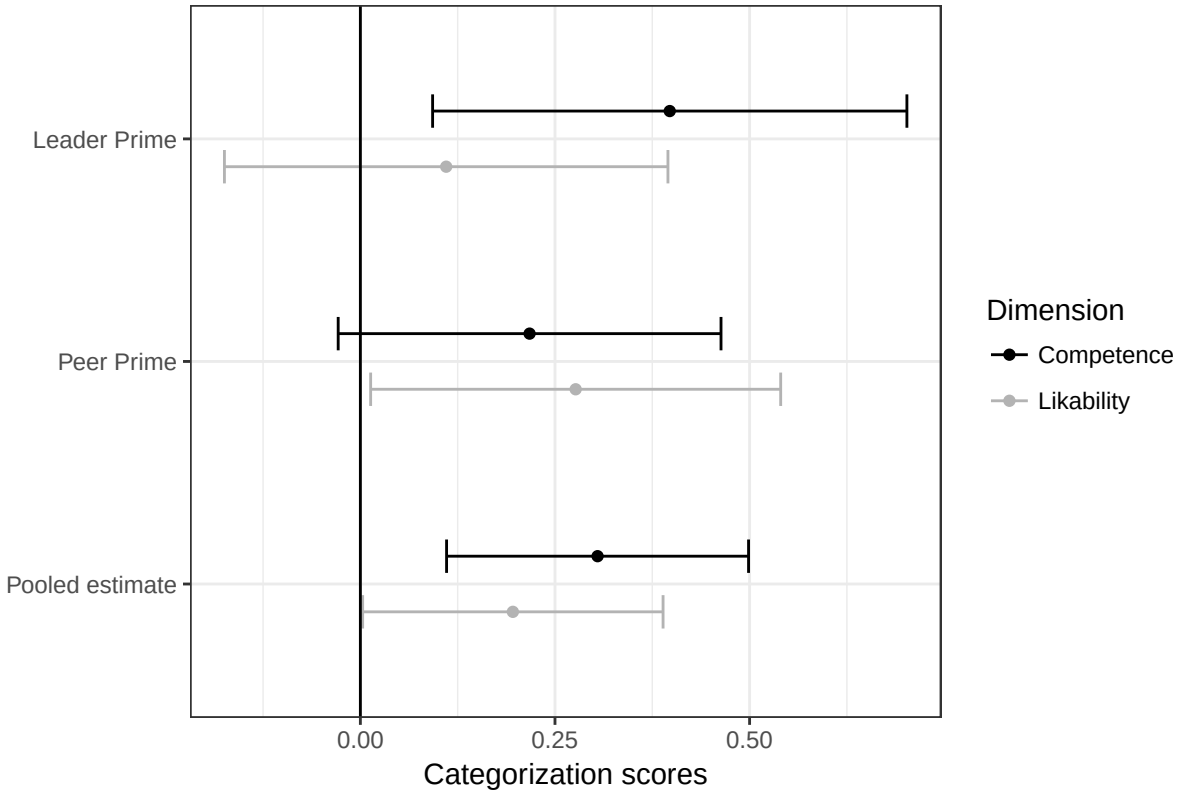
Test statistic	df	P value	Alternative hypothesis
-1.26	194	0.21	two.sided

In the partner frame, is categorization by competence significantly higher than by likability?

Table 13: Paired t-test: `data$warm_catscore[data$treat == "Peer prime"]` and `data$comp_catscore[data$treat == "Peer prime"]`

Test statistic	df	P value	Alternative hypothesis
0.32	206	0.749	two.sided

Figure 2. Categorization scores Study 4



5.2.2. Partner and Leader choice

Table 14: Preference for likable targets by treatment

	Lead prime	Peer prime
Unlikable	0.12	0.21
Likable	0.88	0.79

Table 15: Pearson's Chi-squared test with Yates' continuity correction: `table(likable, data$treat)`

Test statistic	df	P value
4.59	1	0.0322 *

Table 16: Preference for competent targets by treatment

	Lead prime	Peer prime
Incompetent	0.14	0.24
Competent	0.86	0.76

Table 17: Pearson's Chi-squared test with Yates' continuity correction: `table(competence, data$treat)`

Test statistic	df	P value
5.7	1	0.017 *

Table 18: Significance test for preference for competence in partner choice

Test statistic	df	P value
57.4	1	3.56e-14 * * *

Table 19: Significance test for preference for competence in leader choice

Test statistic	df	P value
102	1	5.68e-24 * * *

Alternative approach: considering trait combinations

	Lead prime	Peer prime
Likable_Comp	0.77	0.65
Likable_Incomp	0.1	0.14
Unlike_Comp	0.09	0.12
Unlike_Incomp	0.04	0.09

Table 21: Pearson's Chi-squared test: `table(leadtype, data$treat)`

Test statistic	df	P value
9.4	3	0.0244 *

These effects are not driven by a single outlier action sentence

1	2	3	4	5	6	7	8
0.0249	0.0647	0.0647	0.44	0.0597	0.0373	0.0398	0.269

Other

- This data collection was supported by a grant from AUFF.

- Data is also available through https://raw.githubusercontent.com/alexanderbor/paper2/master/P2_S4.csv