A Model for Categorization: *STEM stereotypes*

Huiwen Duan
Juliette Le Saint
Tianying Feng
Tiffany Kanamaru

Background

Categorization: Allows people to identify, recognize, differentiate and understand single objects and classes of objects **and place them in a group** by means of their memory representations in order to generate adaptive behavior. (Smith, 2014)

Prototype theory: Categories are represented by <u>**one**</u> ***abstracted*** **"most typical" member of the category (the prototype)**. New stimuli are compared to this prototype, minimizing the distance between prototypes and members of the category. (Rosch, 1978)

Exemplar theory: New stimuli are compared to **all remembered prior instances** in memory, called "exemplars." The most typical exemplars are remembered first. (Smith, 2014)

Most cognitive scientists agree both theories hold true in different situations, though many favor the exemplar theory, as it captures more complex boundaries.

Choosing a model

Medin's Context Theory for Classification Learning (1978)

- Well-known, reliable model for categorization; **accounts for how cues in a given context allow for event recall**
- Determines two algorithms (models) that approximate exemplar and prototype models.
    - Used **distances along with weights**, but his likelihood of favoring a certain category **is not probabilistic.**

text, and the event are stored together in memory and that both cue and context must be activated simultaneously in order to retrieve information about the event. A change in either the cue or the context can impair the accessibility of information associated with both.

This idea is depicted in Figure 1. R(cue), R(context), and R(event) refer to the memory representation of the cue, context, and event, respectively. It is further assumed that a particular stimulus component serves a cue function and acts as context for other cues. In our example, blue is part of the context in which circle appears. As a result, transfer or generalization along the form dimension will not be independent of color value. This nonindependence represents an important constraint on the accessibility of stored information and provides the basis for the experimental contrasts to be considered later.

This formulation is closely related to the assumptions of the Estes hierarchical association model (Estes, 1972, 1973, 1976). The cue-context node corresponds to what Estes calls a "control element," and we use it to denote the assumptions (a) that neither cue nor context is directly associated with an event or outcome and (b) that inputs from both cue and context are needed to activate the node and provide access to the representation of an event. The latter assumption implies that the effect of cue changes and context changes combine in an interactive manner.

Specific Assumptions

1. Category judgments are based on the retrieval of specific item information; no categorical information is assumed to enter into the judgments independently of specific item information. While it is assumed that categorical information does not influence judgments, this is not the same as assuming that category level information does not exist. In the case of natural categories, information on the level of categories is often explicitly presented. Our proposal simply is that judgments in classifica-

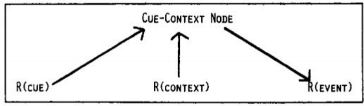


Figure 1. Illustration of factors proposed to determine the accessibility of information associated with a cue presented in a particular context. (R refers to the memory representation of the cue, context, or event.)

2. The probability of classifying exemplar i into category j is an increasing function of the similarity of exemplar i to stored category j exemplars and a decreasing function of the similarity of exemplar i to stored exemplars associated with alternative categories. Specifically, it is assumed that the evidence favoring a category j response to probe i is equal to the sum of the similarities of probe i to the stored j exemplars divided by the sum of the similarities of probe i to all stored exemplars. For purposes of the present article, we assume that the probability of a j response is equal to the evidence favoring a j classification. The mechanism by which these similarities operate is detailed in the next assumption.

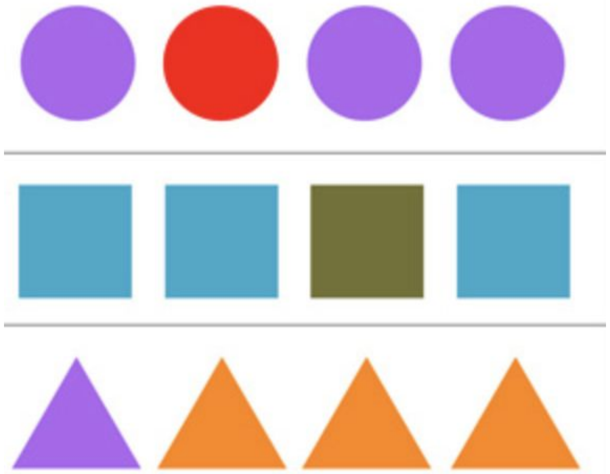
3. Probe or test stimuli act as retrieval cues to access information associated with stimuli similar to the probe. Which stimuli will be retrieved depends on the overall similarity of the stored exemplars to the probe stimulus. Instead of proposing that subjects compute the similarity of a probe to all of the training patterns, we assume that the retrieval rules act to determine which patterns are likely to be accessed. In fact, later on we shall consider the possibility that judgments are based solely on the first pattern retrieved.

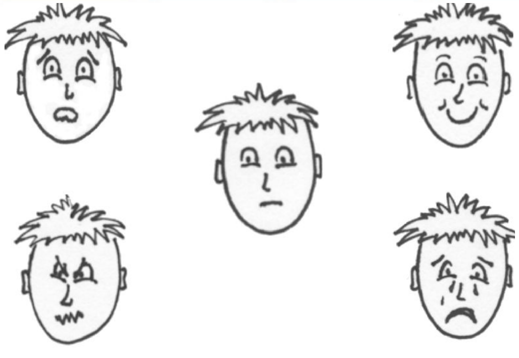
4. The similarity of two cues along a dimension can be represented by a similarity parameter whose value can range between 0 and 1, with 1 representing maximum similarity. For example, the similarity of a yellow circle and a blue triangle along the color dimension would be represented by a parameter c for color similarity, while form similarity would be represented by a parameter f. The parameter c for color would be larger if the two colors were yellow and orange than if the two colors were yellow and blue, since presumably yellow

Medin's Context Theory for Classification

Context theory: Classification judgments are based **exclusively** on the retrieval of stored exemplar information that is a **combination** of information from the cue (feature) dimensions.

- Testing **ill-defined concepts** in his experiment: categorization for **geometric forms** and **schematic faces.**





Medin's Experiment

- **Stimuli** represented with binary values along multiple (4) cue (feature) dimensions: color, form, size, position denoted as c, f, s

$$111?\text{–}A(A_1) \qquad 10?0\text{–}A(A_2)$$
$$00?1\text{–}B(B_1) \qquad 110?\text{–}B(B_2),$$

- Context Model (exemplar, p. 211)

$$E_{A,1} = \frac{1 + p + cfs}{1 + p + cfs + cfsp + cf + sp}.$$

multiplicative, interaction

- Independent Cues Model (prototype, p. 219)

$$E_{A,1} = W_c + W_f + W_s + W_p,$$

additive, average

- **Training phase**, distractor task, **testing (transfer) phase** — participants decide membership to category A or B

Design: Training and Transfer

Category A: Stimulus	Pattern C	F	S	P	Category B: Stimulus	Pattern C	F	S	P
1	1	1	1	1	4	0	0	0	0
2	1	1	1	0	5	0	0	1	1
3	0	0	0	1	6	1	1	0	0

Transfer: Stimulus	Pattern C	F	S	P
7	0	1	0	1

Figure 2. Abstract notation for representing stimuli with binary values along four dimensions. (C, F, S, and P stand for color, form, size, and position, respectively.)

Background — STEM stereotypes

family resemblance: Category members have variable and probabilistic similarity relationships (features overlap). This favors the **prototype theory**, which calls on the distance between concepts and their prototype.[1]

Gender

STEM: **Men >> Women**
Men were associated with stronger science identification and stronger science career aspirations (Cundiff et al., 2013)

Personality

Characteristics range from "dedication, intelligence, and altruism to being uninteresting, **unsociable**, and dull" (Nassar-McMillan et al., 2011)

Hobbies

STEM hobbies: playing video games, **watching anime,** and programming (Cheryan et al., 2011)
non-STEM hobbies: Playing sports, **hanging out with friends**, and listening to music (Cheryan et al., 2011)

Appearance

non-STEM: **takes care of appearance**, more well-kept.
STEM: doesn't take care of appearance, unfashionable

[1]Smith, J David (2014) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3947400/

Our Questions & Hypotheses

Medin favors the exemplar theory, and uses novel stimuli. What if people have a pre-existing prior — will they still rely on training data when they transfer?

Which theory better anticipates how people make category judgments regarding whether a person is a STEM major, if we compare human performance to Medin's model for the exemplar and prototype theories?

Hypothesis: The **prototype theory** will better explain how people judge whether a person is a STEM major, as the distinction between STEM and non-STEM majors revolves around **family resemblance,** which states a pair of members **will share some but not all features**.

David Smith : Prototypes, exemplars, and the natural history of categorization (2014)

Method - General Outline

1. **Present training stimuli**
   - Stimuli is a story about an individual
   - Ask if the person in the story is a STEM major or non-STEM major
   - Tell them the correct answer
2. **Present a distraction task**
3. **Present new stimuli**
   - Ask if the person in the story is a STEM major or non-STEM major
   - **Do not** tell them the correct answer

1

Method: *Creating the Stories*

Training Stimuli

"A" Stimuli

Stimulus Number	C	F	S	N	Fe	Rat-ing
4	1	1	1	0	4.9	4.8
7	1	0	1	0	3.3	5.4
15	1	0	1	1	3.2	5.1
13	1	1	0	1	4.8	5.2
5	0	1	1	1	4.5	5.2

"B" Stimuli

Stimulus Number	C	F	S	N	Fe	Rat-ing
12	1	1	0	0	5.5	5.0
2	0	1	1	0	5.2	5.1
14	0	0	0	1	3.9	5.2
10	0	0	0	0	3.1	5.5

New Transfer Stimuli

Stimulus Number	Dimension Values				Rating	
	C	F	S	N	A-Predicted	B-Predicted
1	1	0	0	1	3.7	
3	1	0	0	0		4.4
6	1	1	1	1	5.3	
8	0	0	1	0		4.1
9	0	1	0	1	3.3	
11	0	0	1	1	4.1	
16	0	1	0	0		4.9

1110

	1	0
Gender	Ryan	Jennifer
Hair	Quickly combs hair	Carefully style hair
Favorite show	Fantasy	Sitcom
Personality	Not social	Social

"Ryan is going on a date tonight. He turns off his favorite fantasy show and quickly combs his hair. He loves meeting and hanging out with new people so he is excited about this date."

Method - *Training Stimuli*

1. Present the story

Jennifer gets out of bed and realizes it's almost time to go to class. She heads over to her bathroom and quickly combs her hair. Walking back to her room, she makes plans to watch her favorite fantasy show by herself when she gets home from school.

2. Ask the question

Is the person a stem major?

3. Present the answer

Jennifer is a stem major student.

Method - *Distraction Task*

1. Shown consonant-vowel-consonants

cat

2. Rate how meaningful they are

How meaningful is this consonant-vowel-consonant?

1	2	3	4	5	6	7
Extremely not meaningful	Very not meaningful	Somewhat not meaningful	Neutral	Somewhat meaning	Very meaningful	Extremely meaningful

Method - *New Stimuli*

1. Present the story

Megan gets out of bed and realizes it's almost time to go to class. She heads over to her bathroom and quickly combs her hair. Walking back to her room, she makes plans to watch her favorite sitcom show with her friends when she gets home from school.

2. Ask the question

Is the person a stem major?

___

The Model - *Replicate*

1. Exemplars

Compared with training examples

2. Prototypes

STEM [1 1 1 1] vs.

Non-STEM [0 0 0 0]

Hamming distance

$$\sum_k |x_k - \mu_{A,k}| < \sum_k |x_k - \mu_{B,k}|$$

3. Algorithm

- Compute the differences
- Medin used the hamming distance

TRAINING STIMULI

"A" STIMULI

STIMULUS NUMBER	DIMENSION VALUES C	F	S	N	FE	RATING
4	1	1	1	0	4.9	4.8
7	1	0	1	0	3.3	5.4
15	1	0	1	1	3.2	5.1
13	1	1	0	1	4.8	5.2
5	0	1	1	1	4.5	5.2

"B" STIMULI

STIMULUS NUMBER	DIMENSION VALUES C	F	S	N	FE	RATING
12	1	1	0	0	5.5	5.0
2	0	1	1	0	5.2	5.1
14	0	0	0	1	3.9	5.2
10	0	0	0	0	3.1	5.5

NEW TRANSFER STIMULI

STIMULUS NUMBER	DIMENSION VALUES C	F	S	N	RATING A-PREDICTED	RATING B-PREDICTED
1	1	0	0	1	3.7	
3	1	0	0	0		4.4
6	1	1	1	1	5.3	
8	0	0	1	0		4.1
9	0	1	0	1	3.3	
11	0	0	1	1	4.1	
16	0	1	0	0		4.9

Table A2
Predicted and Observed Ranked Classification Data from Experiment 2 for the Context Model and the Independent Cue Model

Stimulus number	Observed rank	Predicted rank: Context model	Predicted rank: Independent cue model
4	11	9	4.5
7	4	3.5	8
15	9.5	1.5	2
13	4	7.5	6
5	9.5	7.5	9
12	7	11	10
2	7	10	11
14	4	5	4.5
10	1	1.5	1
1A	15	12	15
3B	12	15	12
6A	2	3.5	3
8B	13	13	13
9A	16	16	16
11A	14	14	14
16B	7	6	7

Note. The rankings are arranged from highest classification scores to lowest.

The Model - *Update*

1. **Exemplars**
   “How similar to the training examples"

2. **Prototypes**
   STEM [1 1 1 1] vs.
   Non-STEM [0 0 0 0]

define

$$P(x_k \mid A) = \begin{cases} 1-\varepsilon & x_k = \mu_{A,k} \\ \varepsilon & \text{otherwise} \end{cases}$$

binary features if x is TRUE

1 - Epsilon = [.90 .80 .70 .92]

3. **Algorithm**

- Again compute the differences
- We adapt the probabilistic approach

The Model - *Update*

1. **Exemplars**

   "How similar to the training examples"

2. **Prototypes**

   STEM [1 1 1 1] vs.
   Non-STEM [0 0 0 0]

3. **Algorithm**

- Again compute the differences
- We adapt the probabilistic approach
- Evidence favoring the STEM category (ratio)

$$\sum_{k=1}^{m} \log \frac{P(x_k \mid A)}{P(x_k \mid B)}$$

Results - Comparison of Data

Human Ratings

Medin's Human Ratings

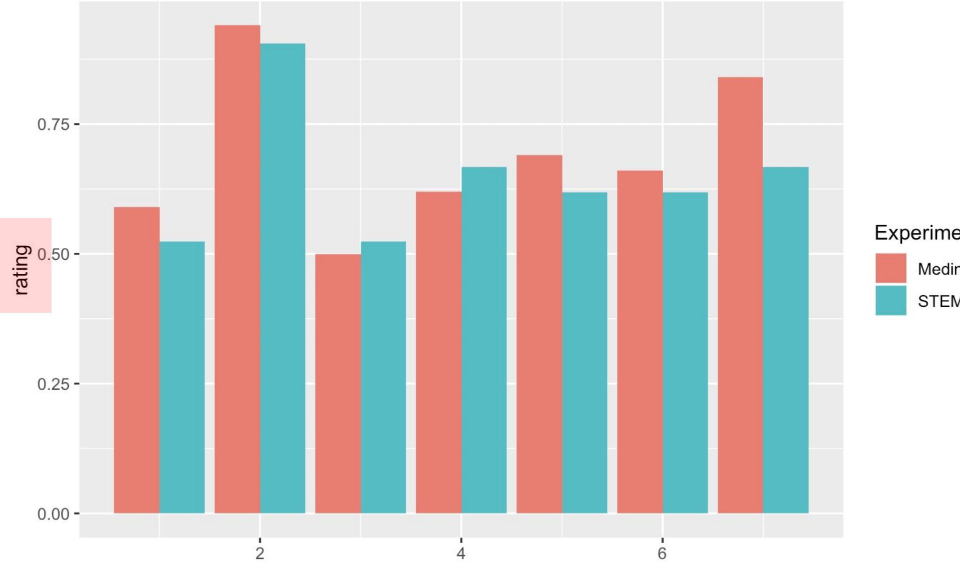
Exemplar Model Predictions

Prototype Model Predictions

Comparisons are based on rank order of the ratings

Results - Our Human Rating Vs. Medin's Human Rating

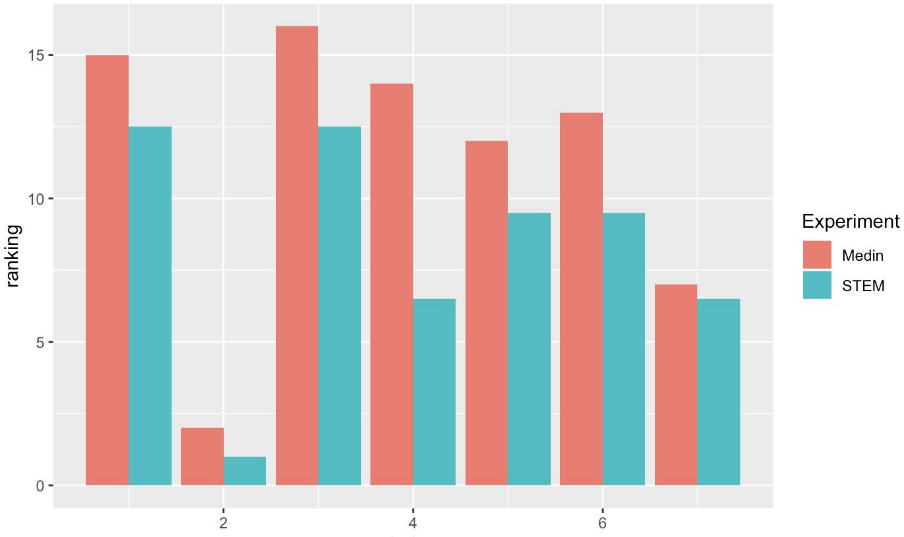
Rating



1000 1001 1111 0010 0101 0011 0100

Correlation = 0.8807

Rank



1000 1001 1111 0010 0101 0011 0100

Correlation = 0.8945

Results - Our Human Rating Vs. Medin's Human Rating

- Our human data **does have a strong correlation** with Medin's human data. However, for some examples, there exists big difference.

- This is probably due to the **difference in subjects' prior knowledge**. In Medin's study, subjects were presented with novel stimuli with novel features.
  - *Our subjects already had prior knowledge about STEM students features.*

- We found an interesting pattern in our human responses, one of our subjects, an econ major, said "*I don't want to judge people stereotypically*", and judged our non-stem prototype as a stem major.
  - *This may explain why we have some human rating result not confirming to our model predictions.*

Results - Validate the Model

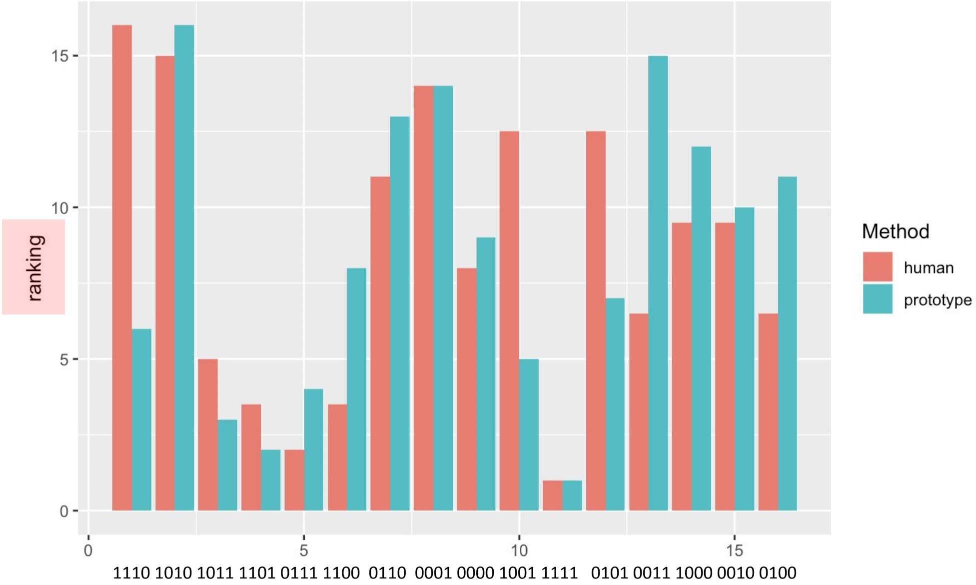
"The main prediction of interest concerns Stimuli 4 and 7. Since the modal prototype is 1111, Stimulus 4 must be at least as close as 7 is to the prototype, no matter how the dimensions are weighted. " (Medin&Schaffer, 1978)

Our data:

Stimuli	Exemplar Prediction Rank	Prototype Prediction Rank
4(1110)	9	6
7(1010)	7	16

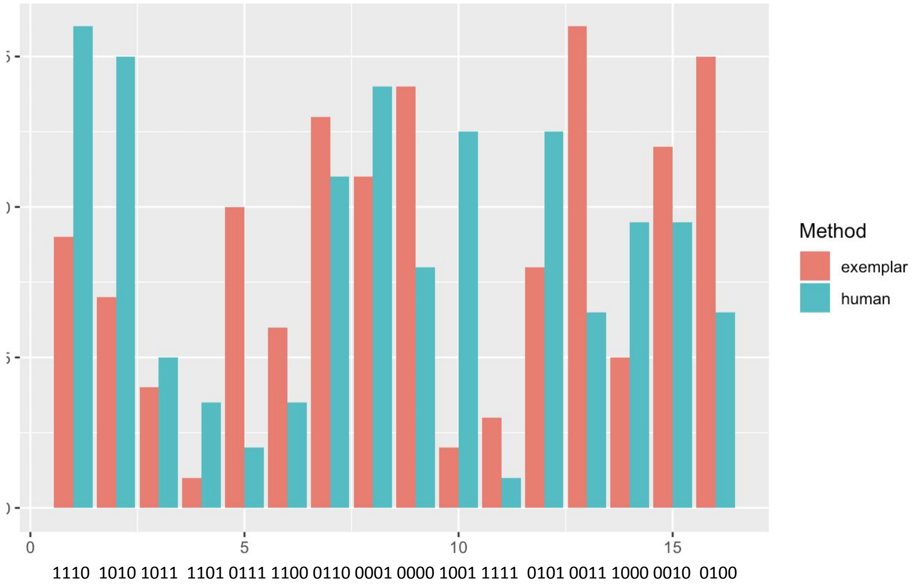
Results - Our Human Rating Vs. Model Predictions

Prototype



Correlation = 0.5250, p value = 0.0368

Exemplar



Correlation = 0.1696, p value = 0.53

Results - Our Human Rating Vs. Model Predictions

- The two models give different predictions.
- **In general, people favor the prototype model.**
- The exemplar model only does a better job predicting human behavior when the stimuli combinations is 1011, the prototype model does a better job on all other stimuli.

Conclusion & Discussion

- It is generally accepted that both the exemplar and prototype approaches are used when people are categorizing social events.

- However, our study found that there does not exist a clear partition of the two approaches when people are judging whether a student is a stem major: *they tend to use the* ***prototype approach on most of the stimuli***, no matter it is closer to a stereotypical STEM student(1110,1111), or it is indistinguishable.

- This can be due to people's strong prior knowledge of STEM stereotypes

Conclusion & Discussion

To get a more generalizable conclusion, in our future studies, we plan to:

- Use more **precise parameters(Epsilon)** in the model.
- Incorporate **more training and testing data**.
- Collect **more human data.**

Thank you!

Huiwen Duan
Juliette Le Saint
Tianying Feng
Tiffany Kanamaru