Regularity in Similarity Judgments

ALEXANDER PARK, ELLEN R.K EVERS, DANIEL LAKENS

SOCIETY FOR JUDGMENT AND DECISION MAKING 2018

Similarity

- Similarity is a central concept in a variety judgment and decision-making domains
  - Choice (Tversky, 1972)
  - Probability (Bar-Hillel, 1974)
  - Decision time (Bhatia and Mullett, 2018)

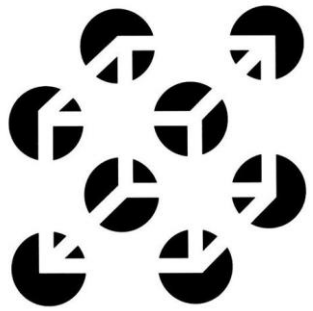
- Large proportion of similarity research has focused on object-to-object, 1-to-1 similarity judgments, not multiple object comparisons.

Hypothesis

- In current research, we focus our attention on how similarity judgments are made among groups of stimuli.
- We believe that, when judging the similarity of a group of objects, the relationships between these objects may play a strong role in how similar the objects seem to each other.
- **Hypothesis: Regularity in differences among the objects will increase similarity judgements of the group when judged as a whole.**

Regularity

- Similarity
  - Strongly dependent on the **psychological representations formed by individuals**
    - (Hahn, 2014; Tversky, 1977; Hahn, Chater, and Richardson, 2003; Shepard, 1962)



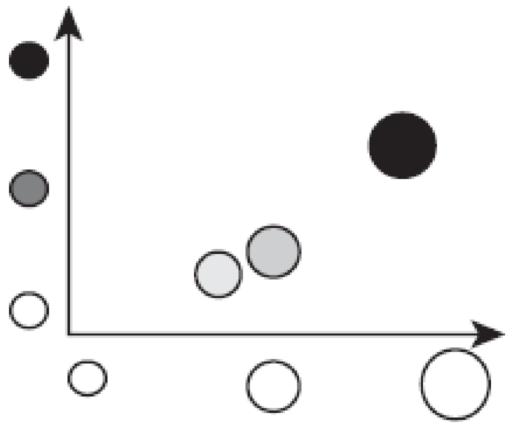
- Gestalt psychology
  - Wertheimer (1923)
  - Our mind imposes structure on individual pieces of information we are exposed to, and our perception is inherently affected by this.
- If our brains automatically impose structure on the pieces of information, and it is these mental representation that forms the basis of similarity judgments, **how is structure imposed?**

Regularity reduces complexity

- Optimal structure would compress information as much as possible while still maintaining the fundamental characteristics.
- One way to quantify the complexity of a set of information being **Kolmogorov complexity** (Kolmogorov, 1965; Li and Vitanyi, 1997).
- Kolmogorov argued that a way to quantify the complexity of a piece of information is the length of the shortest algorithm necessary to describe it.
  - (1, 3, 5, 7, 9, 11, 13) vs (2, 9, 11, 4)
- Thus, any form of regularity, greatly reduces the complexity, making the information easier to encode (Chater, 1999).
- **Groups that are easier to simplify into a short "algorithm", that have lower Kolmogorov complexity, will be judged as more similar as compared to groups that do not exhibit this simplicity.**

Hypothesis vs. Spatial Model

- Geometric-distance model (Shepard, 1962)
  - Items we judge are mapped out on a psychological map
  - The longer the distance the items are on this map, the more dissimilar it seems.
  - [2, 2, 0, 6, 4] vs. [7, 14, 21, 28, 35]



Hypothesis vs. Feature-based Model

- Feature Model (Tversky, 1977)
  - Similarity is a function of (shared features)/(distinct features)
  - [2, 2, 0, 6, 4] vs. [7, 14, 21, 28, 35]

$$S(a,b) = \frac{f(A \cap B)}{f(A \cap B) + \alpha f(A - B) + \beta f(B - A)}, \quad \alpha, \beta \geq 0,$$

Hypothesis vs. Transformation Model

- Transformation model (Hahn, Chater, and Richardson, 2003)
  - More transformations = less similar
  - [2, 2, 0, 6, 4] has 3 number transformations, 10 value transformations
  - [7, 14, 21, 28, 35] has 4 number transformations and, 28 value transformations

Studies

Study 1a: Numbers-within

Study 1b: Numbers-between (MTurk)

Study 1c: Numbers-between (Undergraduate)

Study 1d: Shapes-within

Study 2: Numbers Regularity Factor

Study 1a: Numbers-within

Conditions (N = 50; 56 strings of number)

1) All same: [1, 1, 1, 1, 1], [4, 4, 4, 4, 4]

2) Regular Interval: [2, 4, 6, 8, 10], [6, 5, 4, 3, 2], [1, 3, 5, 7, 9]

3) Two Same Three Different: [2, 6, 0, 0, 8], [2, 2, 4, 8, 9], [0, 1, 4, 4, 9]

4) Irregular Monotonic: [0, 1, 3, 6, 9], [1, 2, 5, 6, 8], [3, 5, 7, 8, 9]

5) Irregular Interval: [6, 4, 8, 7, 1], [0, 6, 4, 8, 1], [7, 0, 5, 1, 9]

Simple

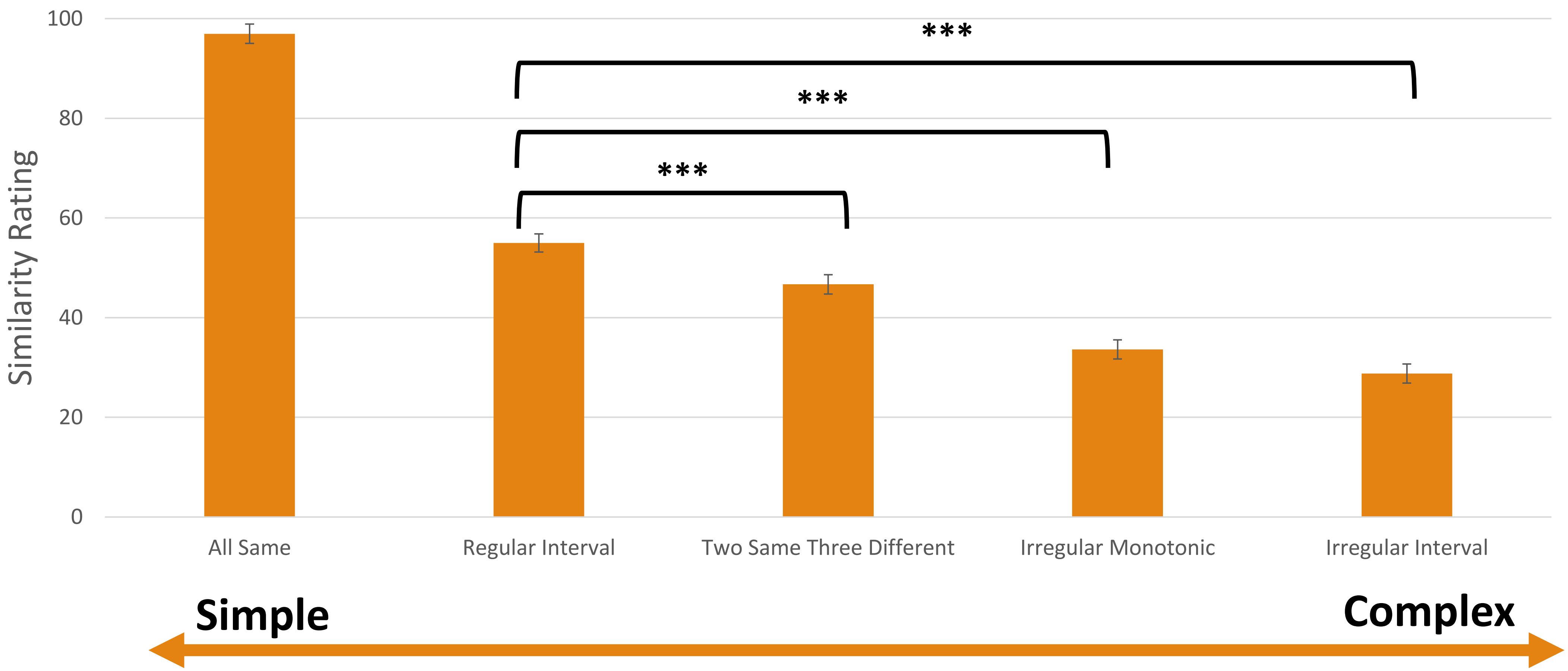
Complex

[2, 4, 6, 8, 10]

"How similar are these numbers to each other?"

0-100 scale, 0 = *Extremely dissimilar*, 100 = *Extremely similar*

Study 1a: Results



* *p* < .05, ** *p* < .01, *** *p* < .001

Study 1b: Numbers-between (Mturk)

Conditions (N = 185)

1) Regular Interval: [3, 6, 9, 12, 15]

2) Three same Two different: [15, 2, 4, 15, 15]

3) Two same Three different: [15, 2, 4, 7, 15]

4) Irregular Interval: [8, 2, 4, 7, 15]

Simple

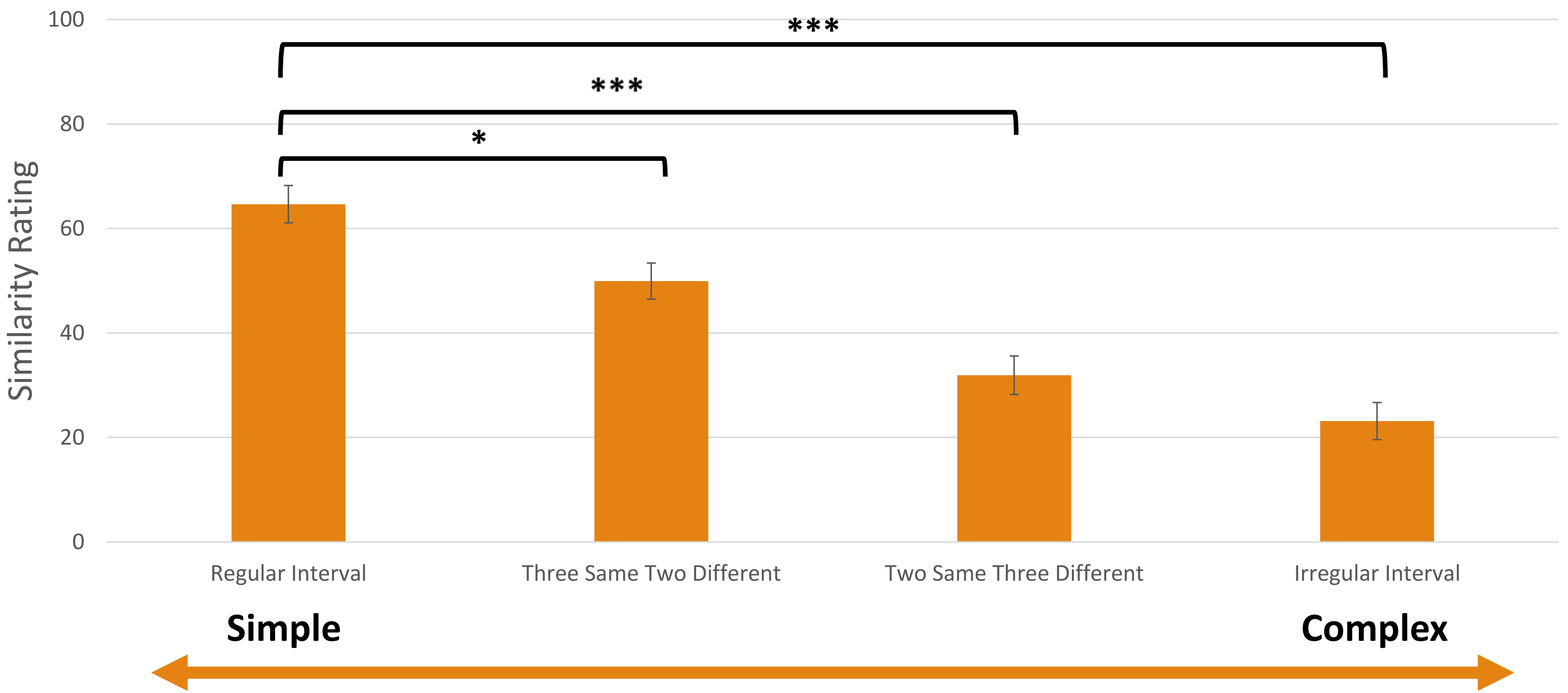
Complex

[3, 6, 9, 12, 15]

“How similar are these numbers to each other?”

0-100 scale, 0 = *Extremely dissimilar*, 100 = *Extremely similar*

Study 1b: Results



** p < .05, ** p < .01, *** p < .001*

Study 1c: Numbers-between (Undergraduates)

PREREGISTERED

Conditions (N = 165)

1) Regular Interval: [3, 6, 9, 12, 15]

2) Three same Two different: [15, 2, 4, 15, 15]

3) Two same Three different: [15, 2, 4, 7, 15]

4) Irregular Interval: [8, 2, 4, 7, 15]

Simple

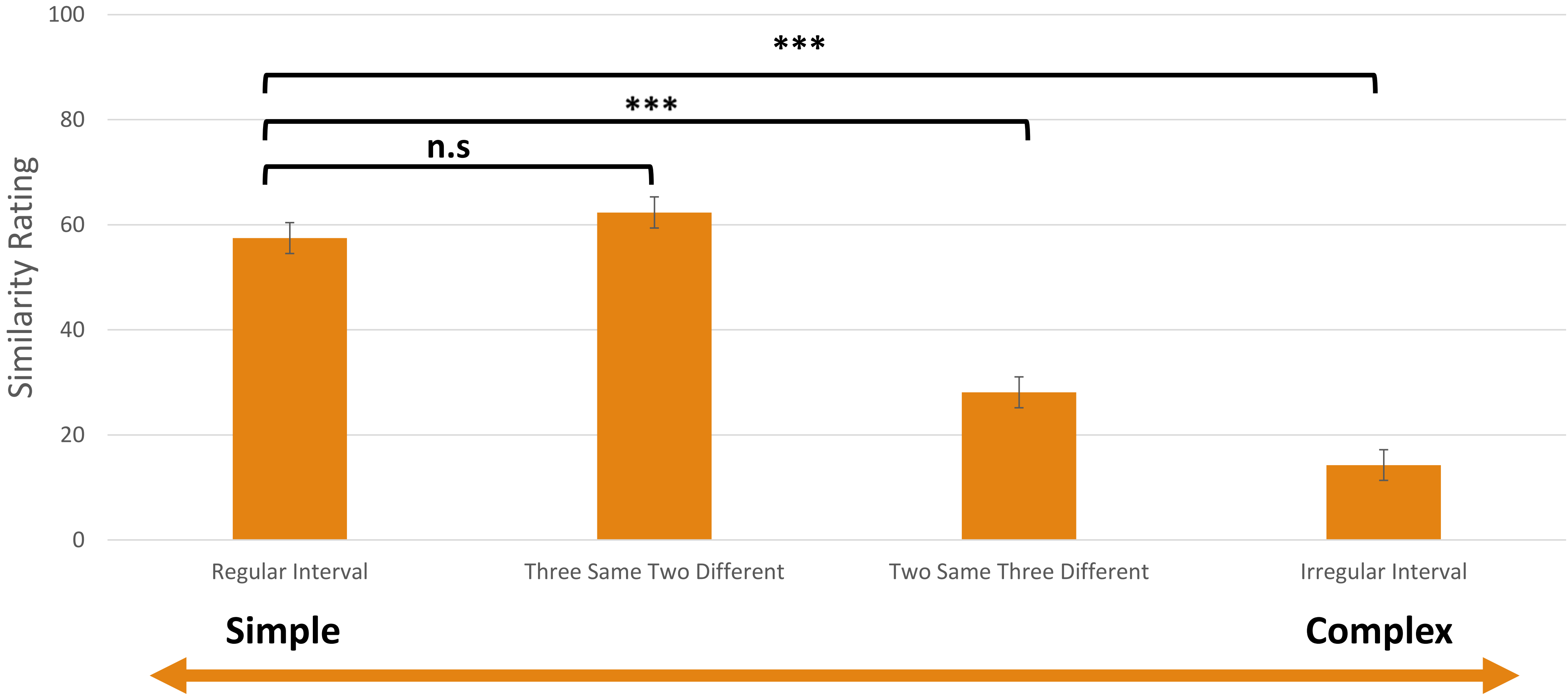
Complex

[3, 6, 9, 12, 15]

"How similar are these numbers to each other?"

0-100 scale, 0 = *Extremely dissimilar*, 100 = *Extremely similar*

Study 1c: Results



* *p* < .05, ** *p* < .01, *** *p* < .001

Study 1d: Shapes

Conditions

1) All Same 

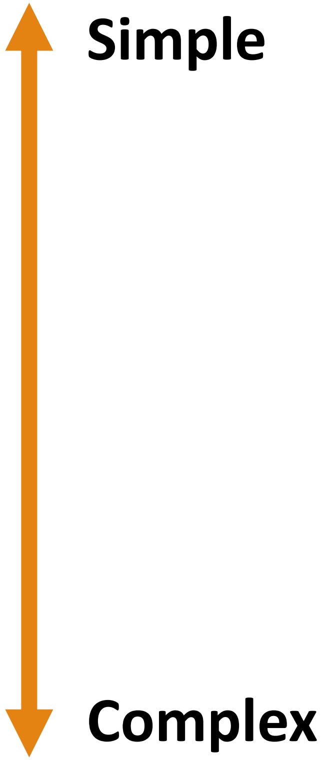
2) Regular Interval 

3) Irregular Order (Condition 2) 

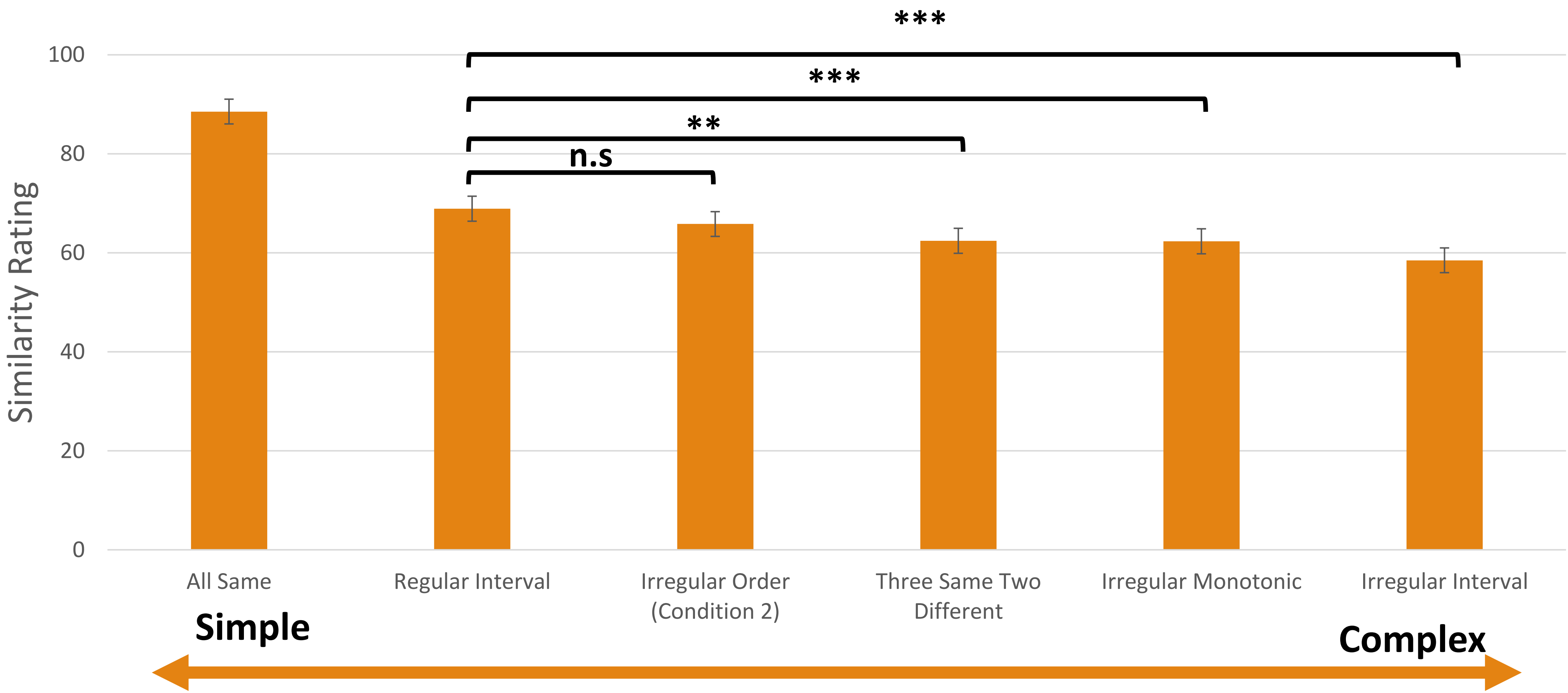
4) Three Same Two Different 

5) Irregular Monotonic 

6) Unequal Interval 



Study 1d: Results



** p < .05, ** p < .01, *** p < .001*

Study 2: Numbers Regularity Factor

Conditions (N = 50; 121 strings of numbers)

1) All same: [1, 1, 1, 1, 1], [4, 4, 4, 4, 4]

2) Regular Interval: [0, 2, 4, 6, 8], [6, 5, 4, 3, 2], [1, 3, 5, 7, 9]

3) Regular Interval (Longer): [0, 2, 4, 6, 8, 10, 12], [3, 6, 9, 12, 15, 18, 21]

4) Two Same Three Different: [2, 6, 0, 0, 8], [2, 2, 4, 8, 9], [0, 1, 4, 4, 9]

5) Irregular Monotonic: [0, 1, 3, 6, 9], [1, 2, 5, 6, 8], [3, 5, 7, 8, 9]

6) Irregular Interval: [6, 4, 8, 7, 1], [0, 6, 4, 8, 1], [7, 0, 5, 1, 9]

Simple

Complex

Study 2: Numbers Regularity Factor

Possible explanations using Tversky's Contrast Model

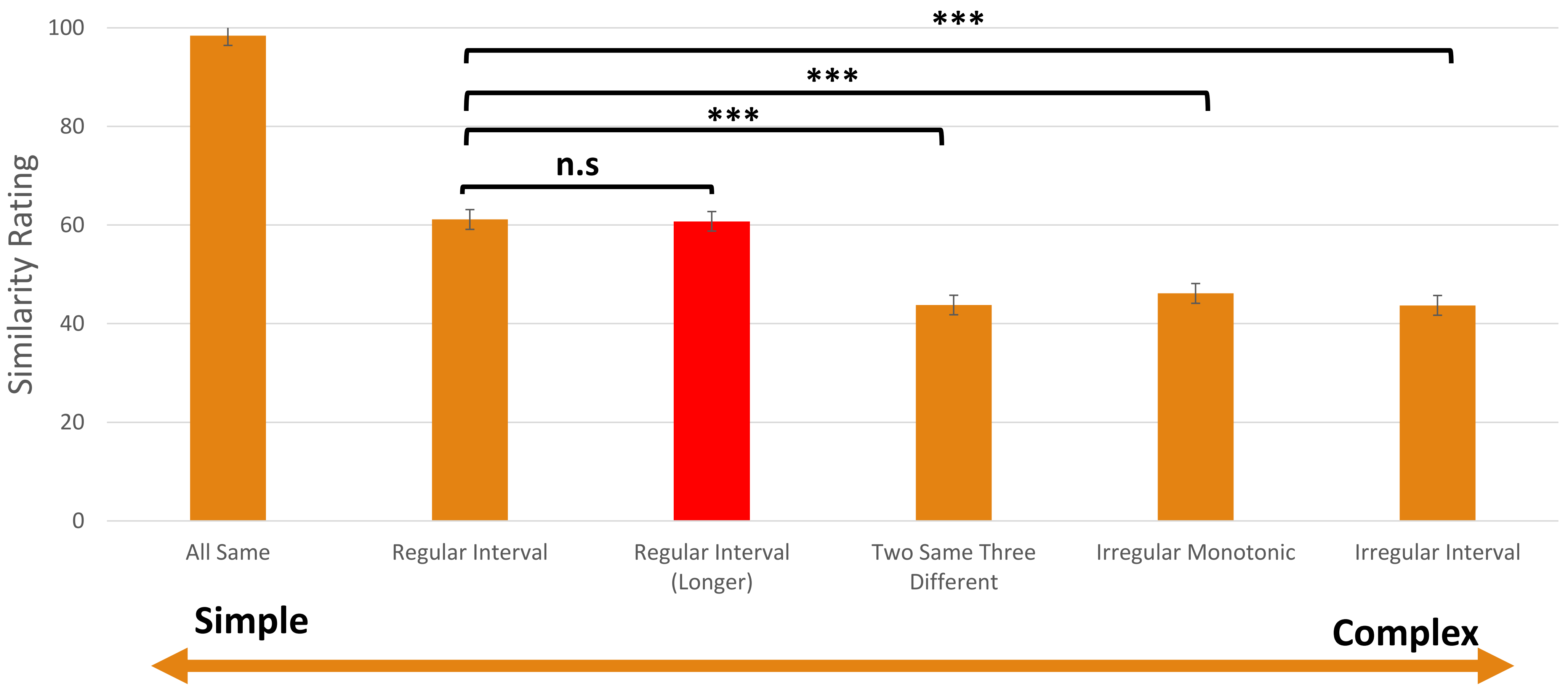
$$S(a,b) = \frac{f(A \cap B)}{f(A \cap B) + \alpha f(A - B) + \beta f(B - A)}, \quad \alpha, \beta \geq 0,$$

1) Regularity is a shared feature for each individual item (numerator increases), in which case longer strings should receive a bigger boost from regularity as compared to shorter strings

2) Regularity could be a shared overarching factor (denominator increases), in which case longer strings should show a decreased boost as compared to shorter strings.

3) Regularity could be some sort of "gestalt" factor that is merely added to final similarity judgments in which case the length of the string should not affect similarity judgments

Study 2: Results



** p < .05, ** p < .01, *** p < .001*

Summary

- We find that regularity in differences among a group of objects increases judgments of similarity
- Set of results somewhat conflicts with the prevalent similarity models
  - [2, 2, 0, 6, 4] vs. **[7, 14, 21, 28, 35]**

Thank you!

UC Berkeley Preferences Lab

WashU Consumer Behavior and Decision Science Lab

alexanderpark@wustl.edu