Cognition

PSYC 2040

L8: Cognitive Models

review

L0: effective study strategies

L1: what is cognition?

L2: mental imagery

L3: eugenics and intelligence testing

L4: associations

L5: behaviorism

L6: information processing

L7: memory I

what's coming up

- start thinking about your research summaries!
- "candidates" are due Apr 2
- SPARK is due Apr 8
- make sure the candidates are **review articles** NOT an empirical study or meta-analysis, i.e., it should not talk about a specific study/question but a broad set of studies and present a theoretical framework, NOT mathematical analyses

10	Wednesday, March 27, 2024	L8: Cognitive Models
10	Friday, March 29, 2024	**Guest Session: President Safa Zaki**
11	Wednesday, April 3, 2024	L9: Memory II
11	Friday, April 5, 2024	L9 continued…
12	Monday: April 8, 2024	**Research Summary [SPARK] due**
12	Wednesday, April 10, 2024	L10: Judgment and Decision Making
12	Friday, April 12, 2024	L10 continued…
13	Tuesday: April 16, 2024	**Monthly Quiz 2**
13	Wednesday, April 17, 2024	L11: Language
13	Friday, April 19, 2024	L11 continued…
14	M: April 22, 2024	**Research Summary [QALMRI] due**
14	Wednesday, April 24, 2024	L12: Social Cognition
14	Friday, April 26, 2024	L12 continued…
15	Monday: April 30, 2024	**Monthly Quiz 3**
15	Wednesday, May 1, 2024	L0-L12 review!
15	Friday, May 3, 2024	**Final**
16	Wednesday, May 8, 2024	Wrapping up!
16	M: May 13, 2024	**Research Reflection due**

new office hours

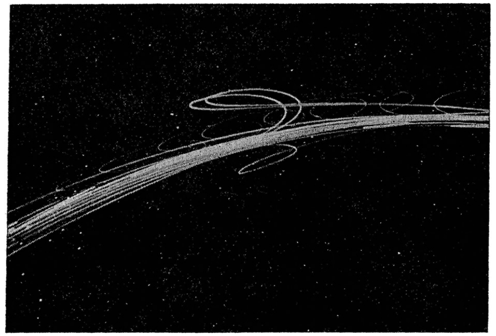
- Wednesdays, 2-5 pm (Kanbar 217)
  - with some exceptions (e.g., next week!)
- Thursdays, 2-4 pm (virtual, link on Canvas)

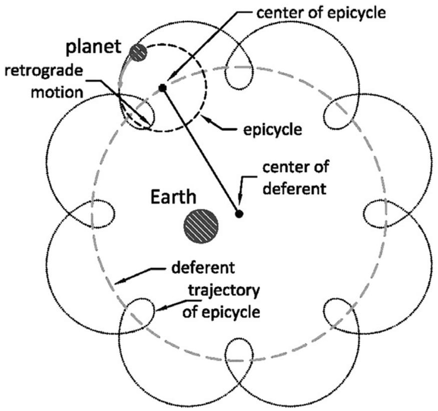
today’s agenda

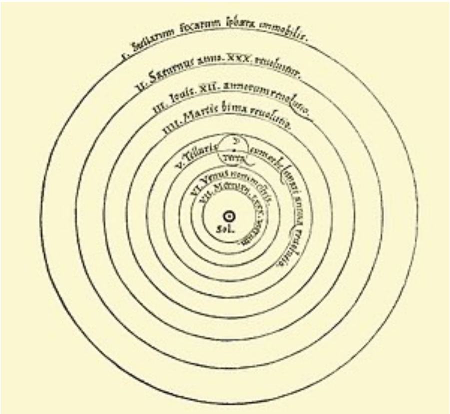
cognitive models

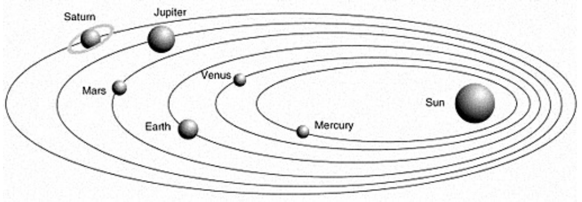
motivating models: planetary motion

- planets typically have curvilinear paths, but appear to have strange “loops”, referred to as retrograde motion
- explaining why this happens requires a model of an underlying process that generates this pattern
- models do not physically *exist,* they are “abstract explanatory devices” that people use to describe, predict, and explain *real data*
- several models may explain the data and scientists must select among different alternatives







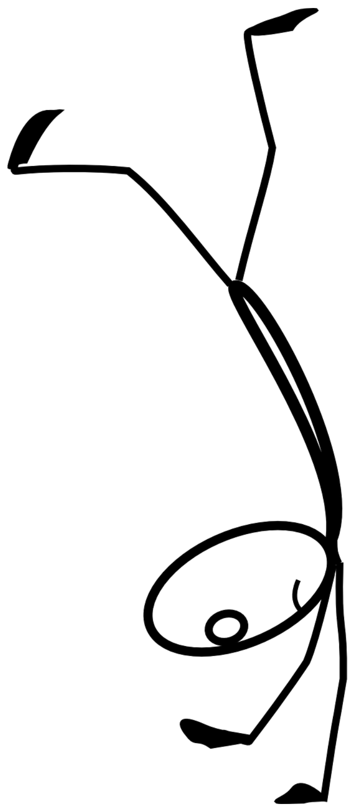


we use models all the time!

- any type of description of data can be considered a model
- averaging a set of numbers is a *model* of the data
  - means can be informative: examples?
  - means can be misleading: examples?
- the Rescorla-Wagner model of associative learning
- other examples?

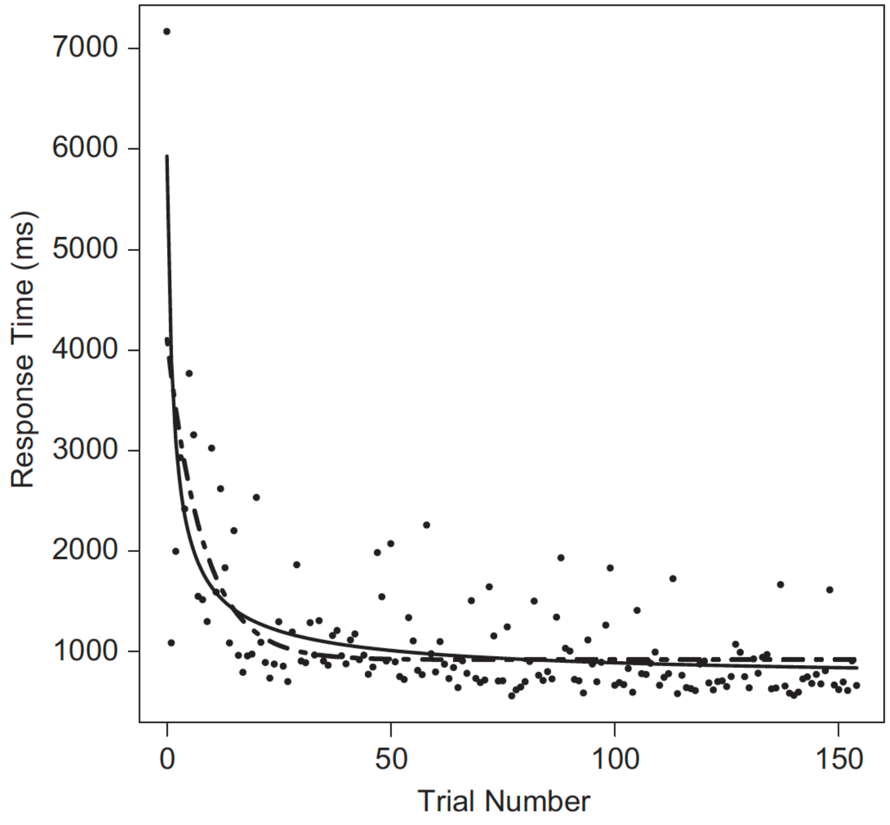
theories of learning

- we know people get better over time at learning a new skill, but how exactly?
- the first time takes forever, the next few attempts lead to major improvements, and then improvements slow down
- two explanations/models:
  - power law: $\mathrm{RT} = N^{-\beta}$
  - exponential law: $\mathrm{RT} = e^{-\alpha N}$,



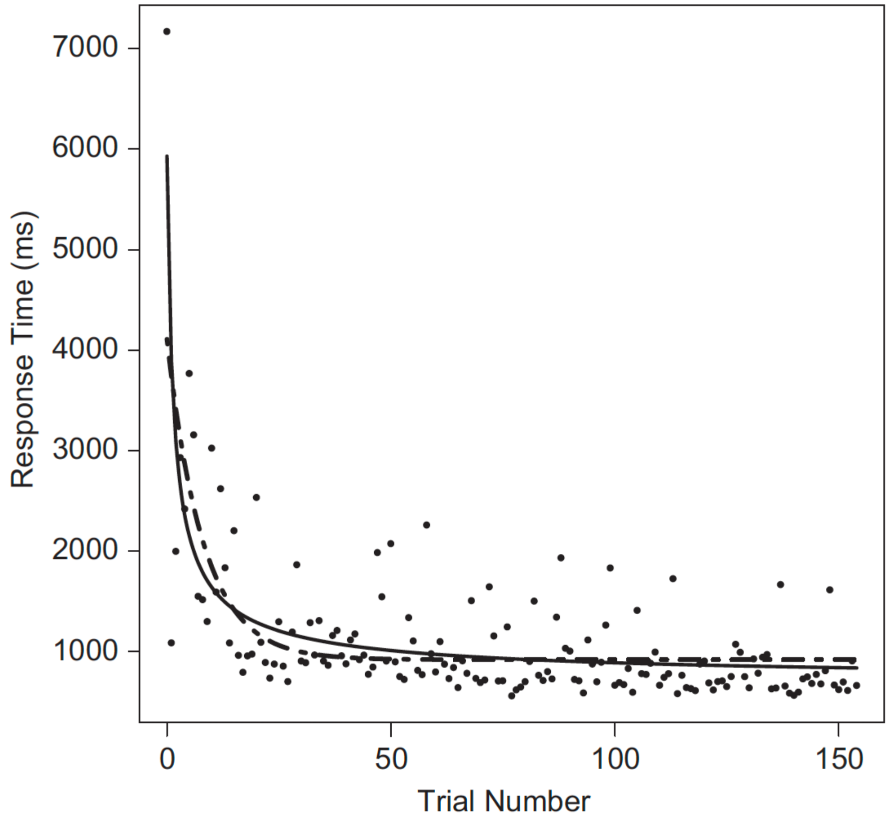
learning: why does it matter?

- the fit of both models is very similar so why does it matter which one is more accurate?
- the exponential form suggests that the relative learning rate remains constant, i.e., regardless of practice, your learning continues to be enhanced by a constant fraction
- the power law suggests that the relative learning rate is slowing down, i.e., as you practice more, you are actually learning less over time
- which model is correct has important practical implications: how much should you practice a new skill?



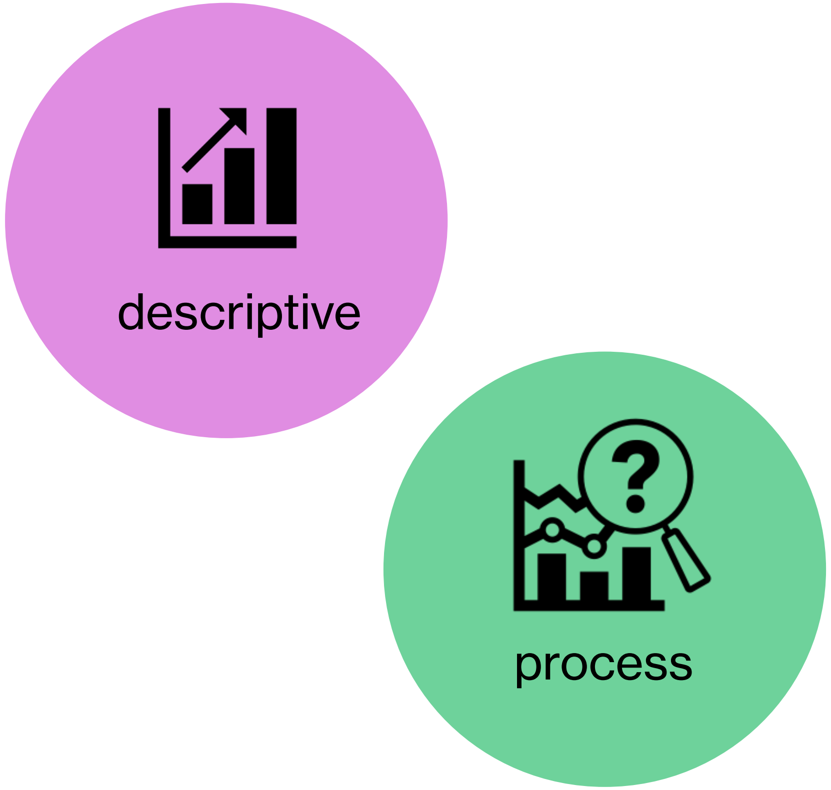
learning: why does it matter?

- Heathcote (2000) showed that the exponential function better fit the trial-level data
- learning curve is better explained by the exponential function
  - the more you learn, the more you retain
- implications for forgetting
  - learning is not the same as forgetting: forgetting follows a function closer to power law (Wixted, 2004), so you lose more initially and lose lesser over time



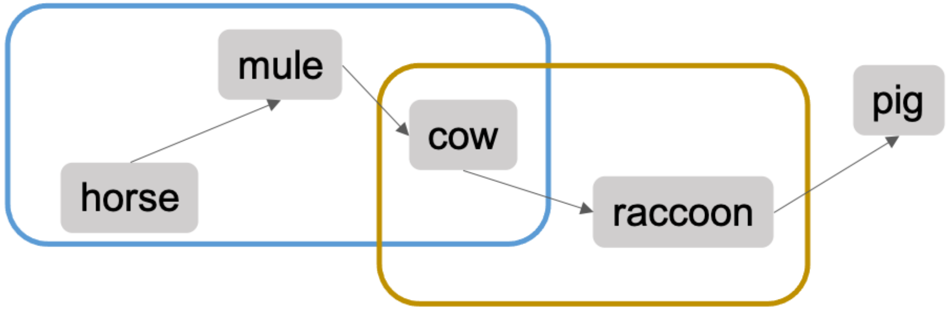
descriptive vs. process models

- descriptive models emphasize describing the data, typically through some type of mathematical formulation and/or statistic
  - examples include the exponential/power laws, means, proportions, etc.
- process models emphasize the underlying mechanism that directly produces the data, and can often generate predictions
  - examples include the Rescorla-Wagner model



categorization

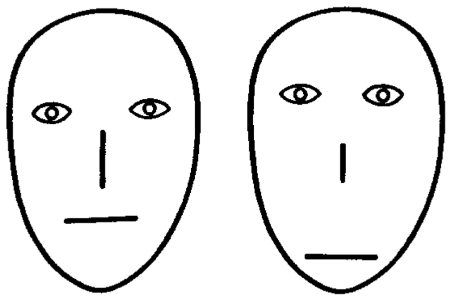
- <u>why</u> do we categorize things?
- <u>how</u> do we categorize things?



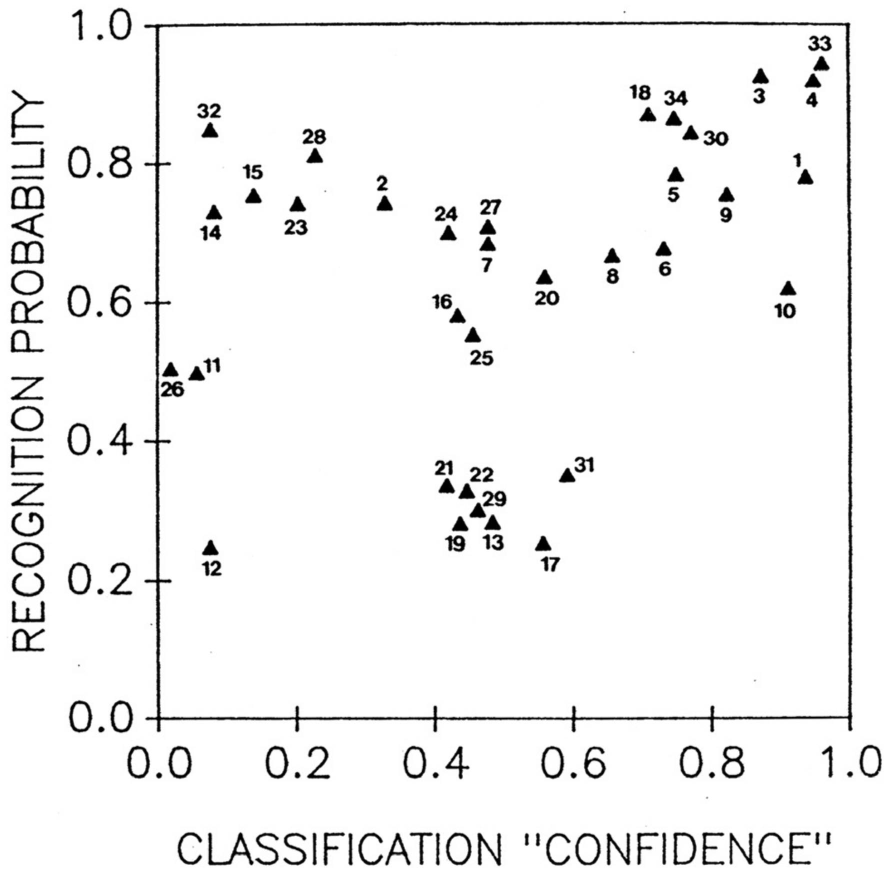
activity: cartoon face experiment

- you did the faces experiment before class
- discuss
  - how did you do the task?
  - was there anything special about MacDonalds or Campbells?

Nosofsky (1991) experiment



- training phase: classify cartoon faces
  - MacDonalds and Campbells
- test phase:
  - classification: classify faces and rate confidence
  - recognition: provide old/new judgments
- classification and recognition had a moderate correlation (r = .36) suggesting barely much of a relationship between the two tasks
- if we knew the classification confidence, then we may not be able to predict the recognition probability

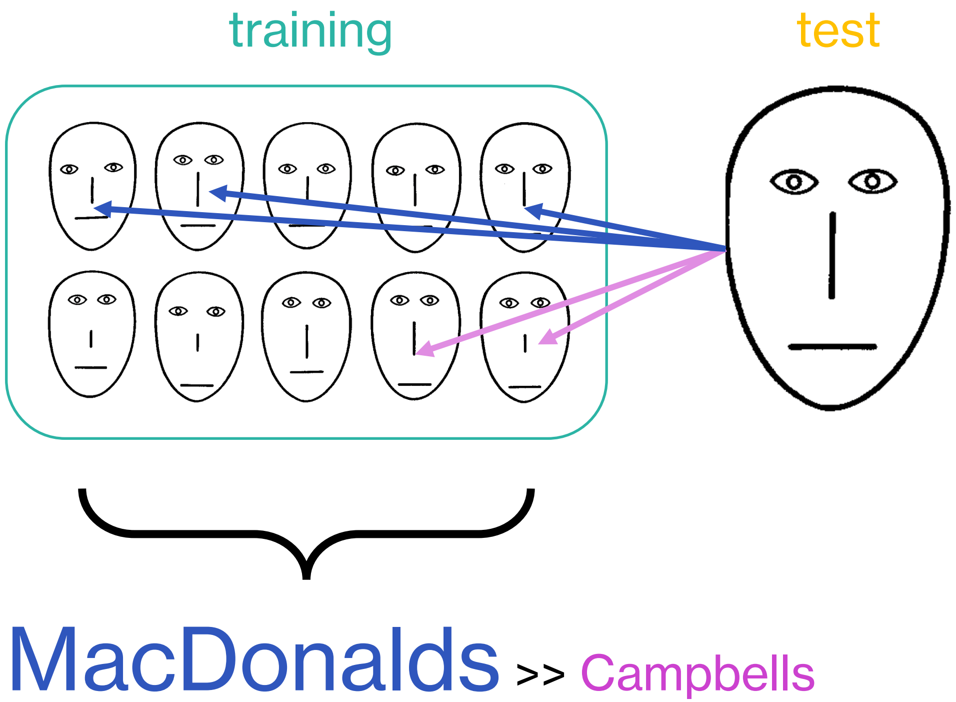


modeling classification

- Nosofsky (1991) set out to explain *how* people classify new faces after having seen examples from two different classes
- a prominent account of classification was the prototype model, which suggested that people create “general” representations of concepts to which new examples are compared
- Nososky (1991) proposed an alternative exemplar model, according to which people compared the presented item to all previously experienced items to compute “similarity”

exemplar model: description

- during training, people store individual examples into memory
- during test, the training items are activated in proportion to their similarity to the test item
- the probability of responding with one label (MacDonald) vs. another (Campbell) depends on the sum of these activations



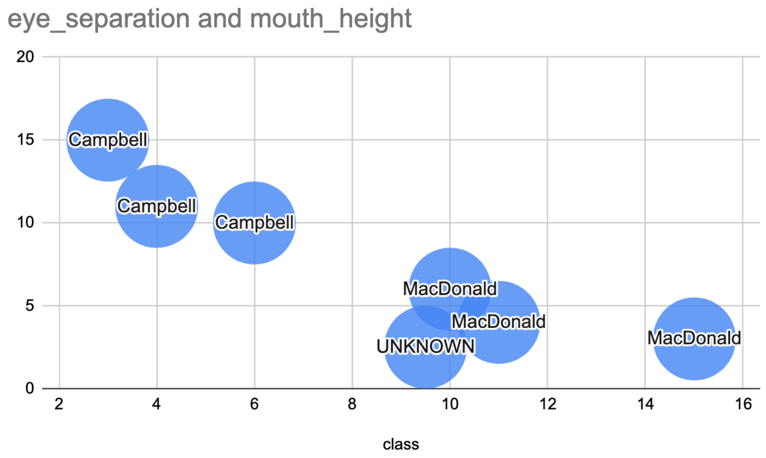
exemplar model: training

- x_i denotes the ith exemplar presented during training
- each exemplar can be defined along *m* dimensions

activity: computing similarities

- in groups, go to the [face dimensions spreadsheet]
- navigate to your group's tab
- select the columns containing face dimensions and class
- insert a chart and choose a "bubble" chart
- can you differentiate between MacDonalds and Campbells?

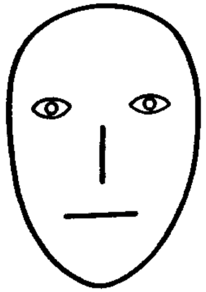
A	B	C	D	E
type	Face	class	eye_separation	mouth_height
training	1	MacDonald	10	6
training	2	MacDonald	11	4
training	3	MacDonald	15	3
training	4	Campbell	6	10
training	5	Campbell	4	11
training	6	Campbell	3	15
TEST	TEST	UNKNOWN	9.5	2.5



exemplar model: training

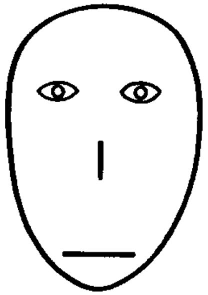
- Nosofsky (1991) varied the faces along 4 features (nose length, eye separation, etc.) such that there was a clear separation between the two classes (MacDonalds and Campbells)
- these features are often referred to as dimensions and can be placed in a multi-dimensional space

x_i



[-1.025, 0.493, 0.048, -0.666]

x_j



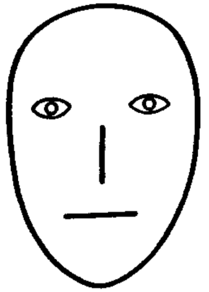
[-0.172, -0.557, 0.337, 0.163]

feature	face 1	face 2
eye height	23.5	19.5
eye separation	21.5	11.5
nose length	13.5	18
mouth height	16.5	12

exemplar model: training

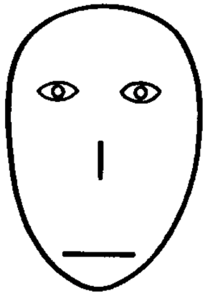
- Nosofsky (1991) varied the faces along 4 features (nose length, eye separation, etc.) such that there was a clear separation between the two classes (MacDonalds and Campbells)
- these features are often referred to as dimensions and can be placed in a multi-dimensional space

x_i

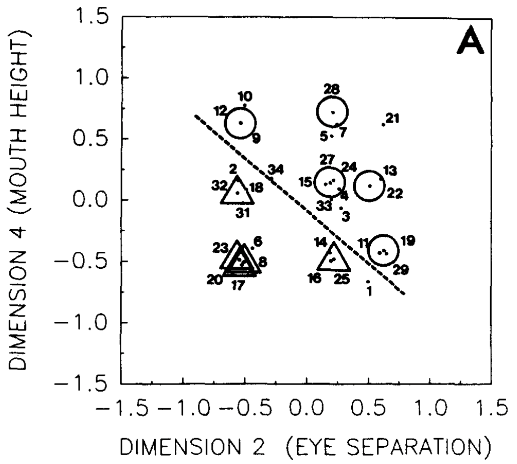


[-1.025, 0.493, 0.048, -0.666]

x_j

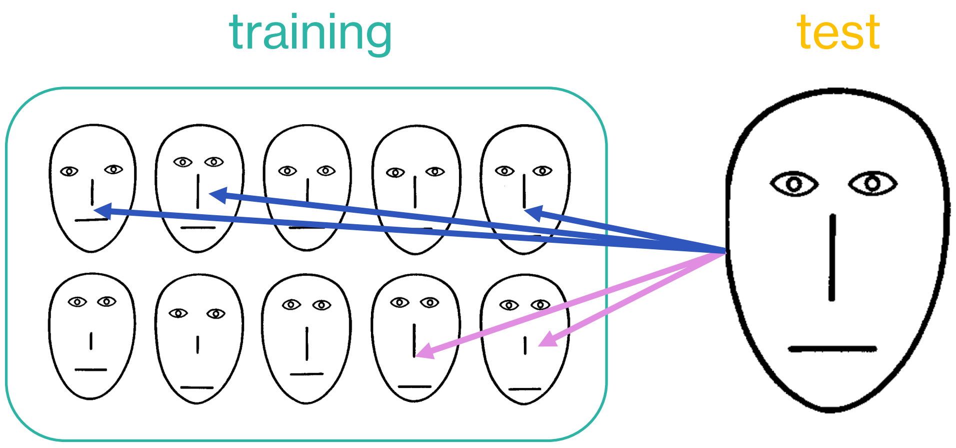


[-0.172, -0.557, 0.337, 0.163]



exemplar model: test

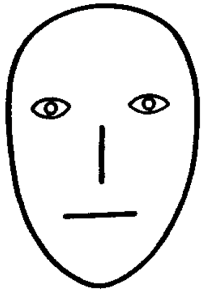
- when a new item (x_k) is presented, each training item is activated in proportion to its similarity to the test item
- but how do we calculate similarity??



exemplar model: similarity

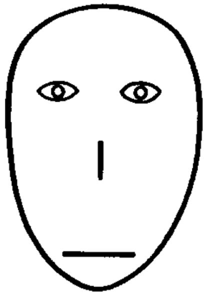
- the similarity between any two items (x_i and x_k) can be calculated using their coordinates in the multidimensional space
- this requires two steps:
    - calculating the Euclidean distance d_{ik} between the items i and k
    - translating distance to similarity through an exponential function

x_i



[-1.025, 0.493, 0.048, -0.666]

TEST (x_k)



[-0.172, -0.557, 0.337, 0.163]

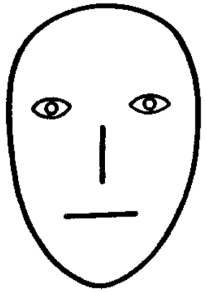
$$d_{ik} = \sqrt[2]{\sum_m |x_{im} - x_{km}|^2}$$

$$s_{ik} = e^{-cd_{ik}}$$

exemplar model: similarity

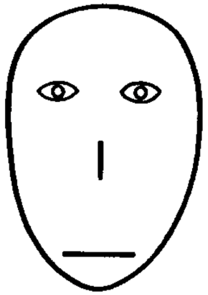
- in groups, go to [the similarity spreadsheet]
- navigate to your group's tab
- use the formulas in columns F and G to compute distance and similarity of each face to the test item
- report back which face has the highest and lowest similarity to the test item

x_i



[-1.025, 0.493, 0.048, -0.666]

TEST (x_k)



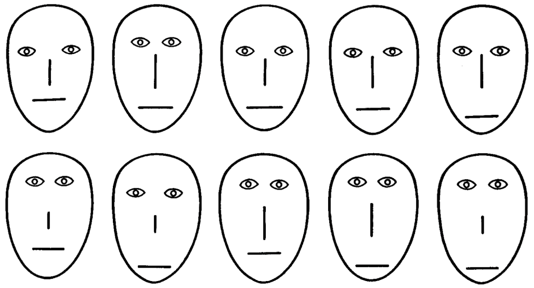
[-0.172, -0.557, 0.337, 0.163]

$$d_{ik} = \sqrt[2]{\sum_m |x_{im} - x_{km}|^2}$$

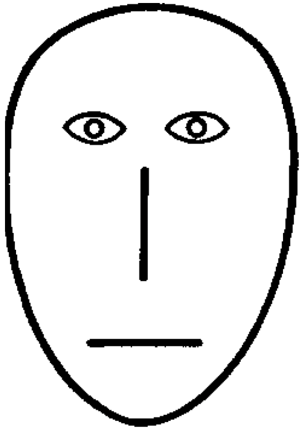
$$s_{ik} = e^{-cd_{ik}}$$

exemplar model: test

- when a new item (x_k) is presented, each training item is activated in proportion to its similarity to the test item
  - exemplar x_i is activated in proportion to its similarity to test item x_k

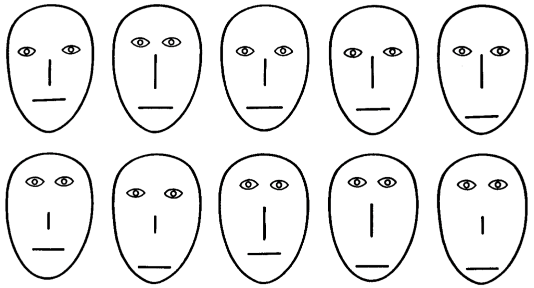


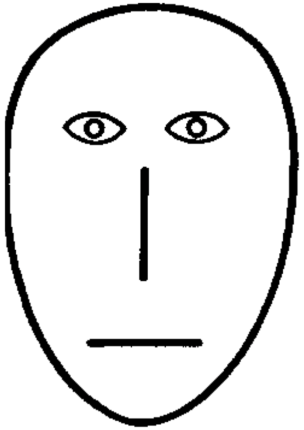
$$s_{ik} = e^{-cd_{ik}}$$



exemplar model: test

- when a new item (x_k) is presented, each training item is activated in proportion to its similarity to the test item
  - exemplar x_i is activated in proportion to its similarity to test item x_k
- activations of each exemplar in a class are added up to produce total activation for the class
- the probability of classifying the new test item is determined by whichever class has higher total activation



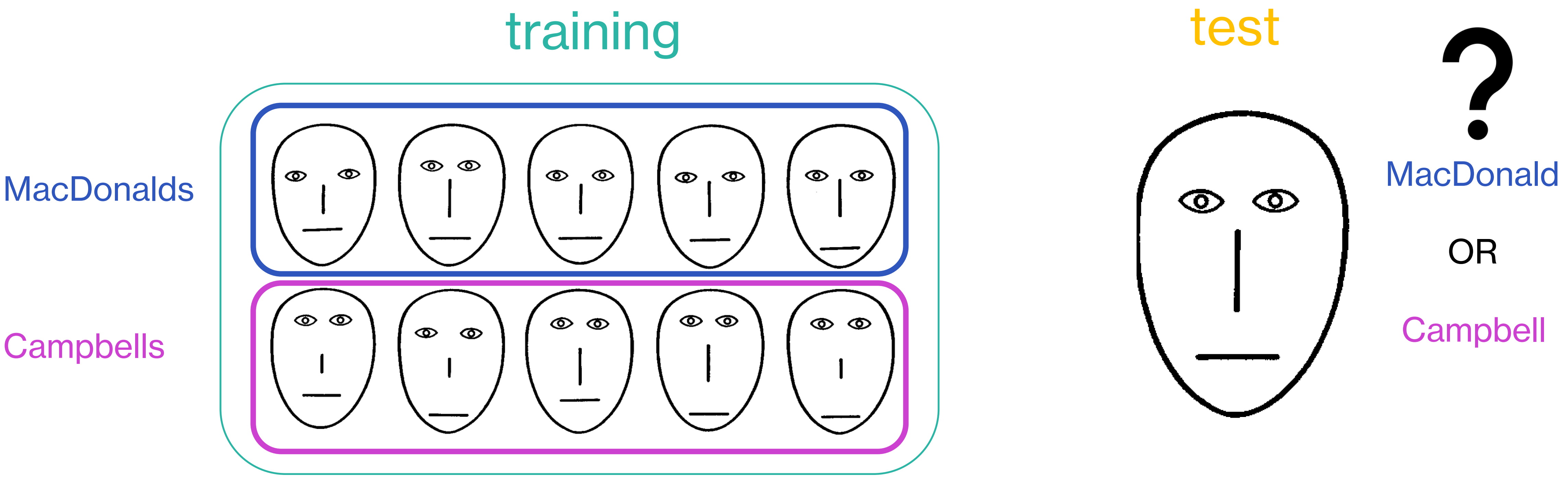


$$activation\ (MacDonald) = \sum_{k \in MacDonald} s_{ik}$$

$$activation\ (Campbell) = \sum_{k \in Campbell} s_{ik}$$

$$P\ (MacDonald) = \frac{activation\ (MacDonald)}{activation\ (MacDonald) + activation\ (Campbell)}$$

how do we classify/categorize?



activity: computing probabilities

- calculate the total activation of MacDonalds and Campbells by adding the similarities for the respective categories
- which class is more activated overall?

$$activation\ (MacDonald) = \sum_{k \in MacDonald} s_{ik}$$

$$activation\ (Campbell) = \sum_{k \in Campbell} s_{ik}$$

activity: computing probabilities

- now calculate the probability of responding MacDonald and responding Campbell
- what is the sum of the two probabilities?
- what decision would you make about this particular test face?

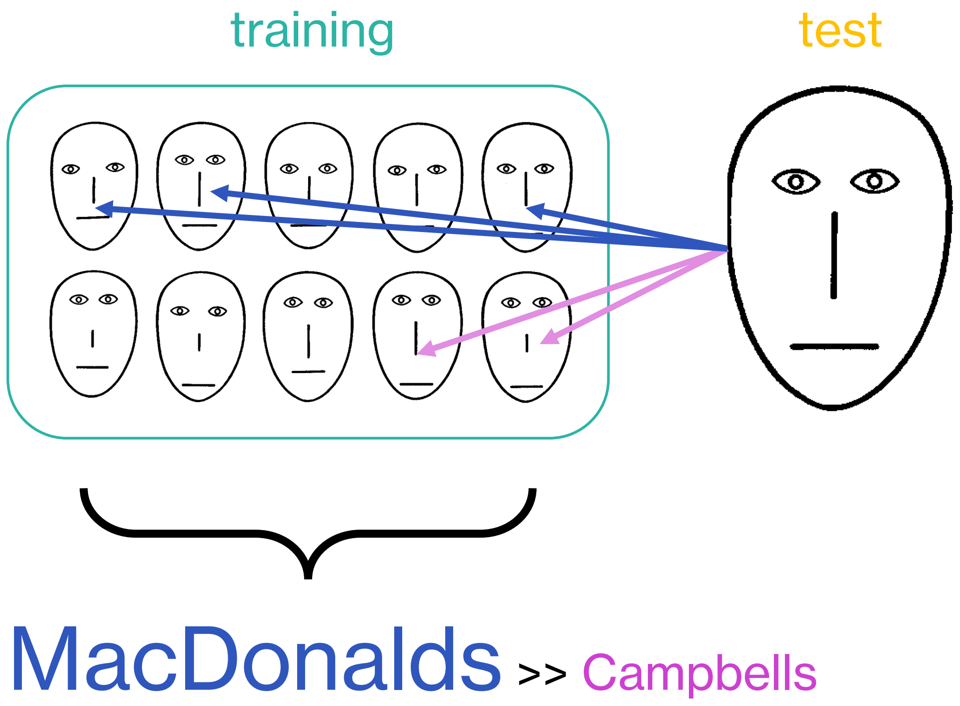
$$activation\ (MacDonald) = \sum_{k \in MacDonald} s_{ik}$$

$$activation\ (Campbell) = \sum_{k \in Campbell} s_{ik}$$

$$P\ (MacDonald) = \frac{activation\ (MacDonald)}{activation\ (MacDonald) + activation\ (Campbell)}$$

exemplar model: review

- during training, people store individual examples into memory
- during test, the training items are activated in proportion to their similarity to the test item
- the probability of responding with one label (MacDonald) vs. another (Campbell) depends on the sum of these activations

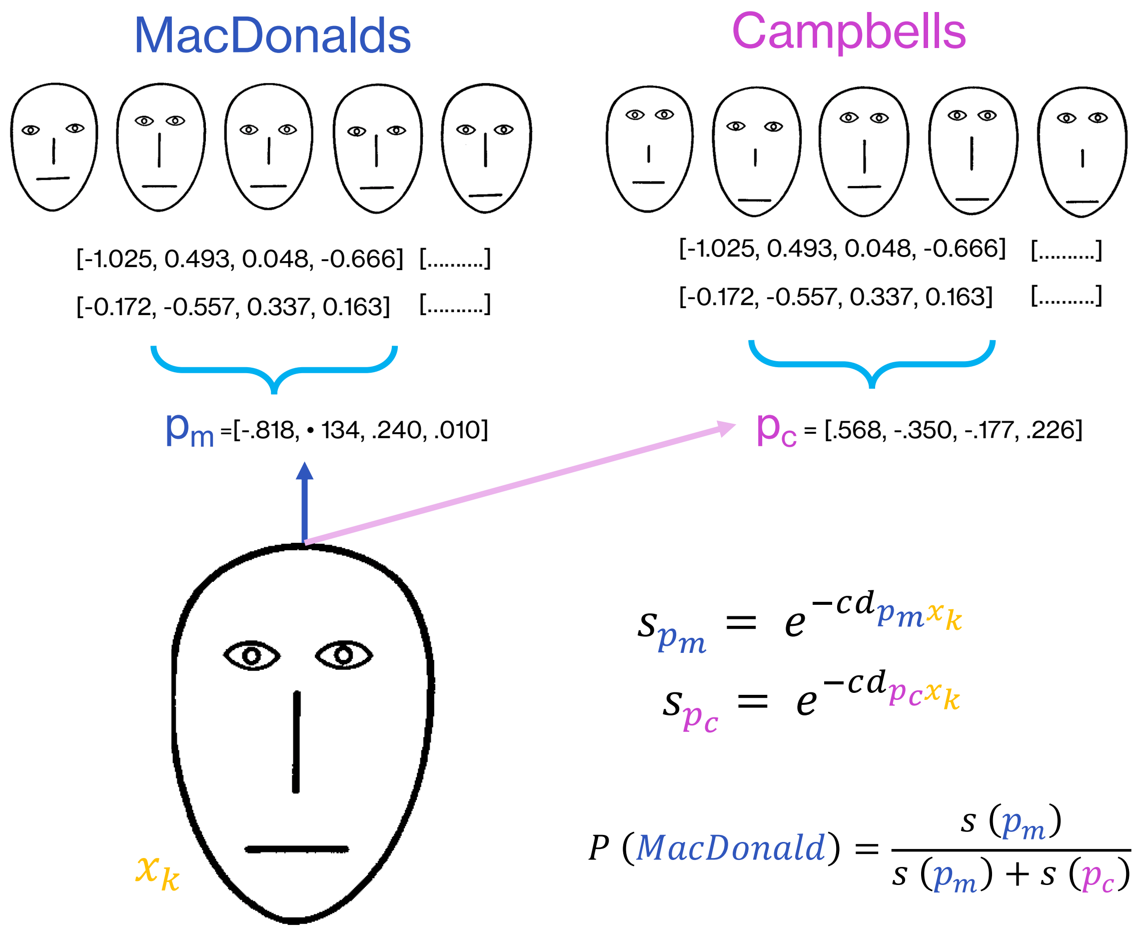


modeling classification

- Nososky (1991) proposed the exemplar model, according to which people compared the presented item to all previously experienced items to compute “similarity”
- a prominent account of classification was the prototype model, which suggested that people create “general” representations of concepts to which new examples are compared

prototype model: description

- during training, all exemplars are "summarized" to form a prototype
- during test, the prototypes for each class are activated in proportion to their similarity to the test item
- the probability of responding with one label vs. another depends on whichever prototype is more activated



$$s_{p_m} = e^{-c d_{p_m x_k}}$$

$$s_{p_c} = e^{-c d_{p_c x_k}}$$

$$P\,(MacDonald) = \frac{s\,(p_m)}{s\,(p_m) + s\,(p_c)}$$

activity: prototype model

- go to [the prototype spreadsheet]
- review how the prototype is generated and similarity to the test item is calculated using the prototype
- examine compute the probability of classification
- what decision would you make about this particular test face?

reviewing the evidence

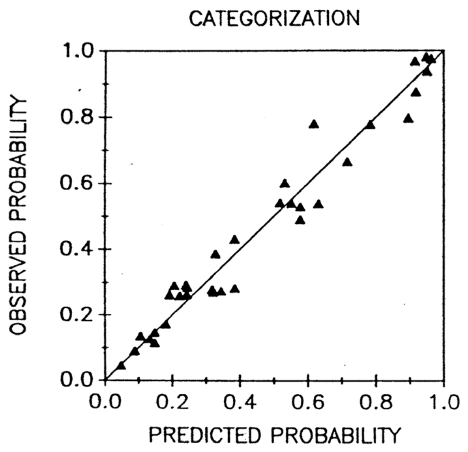
- both exemplar and prototype models have a proposal for how a classification decision may be reached, i.e., they can predict classification decisions given a set of examples and a new test item
  - they are both process/computational models
- we also have a large dataset of classification decisions from human participants who did this experiment
- how can we compare the two models?

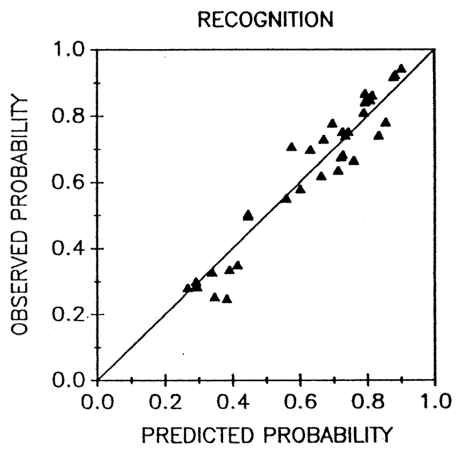
exemplar vs. prototype model?

- the exemplar model performed *better* than the prototype model in predicting human classification decisions
- the generalized context model (GCM) or the exemplar model was able to successfully relate classification confidence to recognition accuracy, such that knowing one of these could predict the other with remarkable accuracy

Table 3
Maximum Likelihood Parameters and Summary Fits, Experiment 1B

Model	Parameters									Fits		
	σ	c	w_1	w_2	w_3	w_4	x_c	b	M_7	SSE	% Var	-ln L
All-subjects analyses												
Context												
Classification	.267[a]	1.077[a]	.15	.15	.29	.41		.173	1.464[a]	.097	96.5	129.2
Recognition	.267[a]	1.077[a]	.13	.56	.23	.08	5.322		1.464[a]	.076	95.4	119.2
Prototype												
Classification	.186[a]	.777[a]	.16	.14	.40	.30		.044	1.123[a]	.175	93.7	181.0
Recognition	.186[a]	.777[a]	.25	.55	.12	.07	1.231		1.123[a]	.182	89.0	156.0





applications?

- ask ChatGPT about some potential applications of exemplar and prototype theories of categorization/classification to real life!
- report back on which example stood out to you

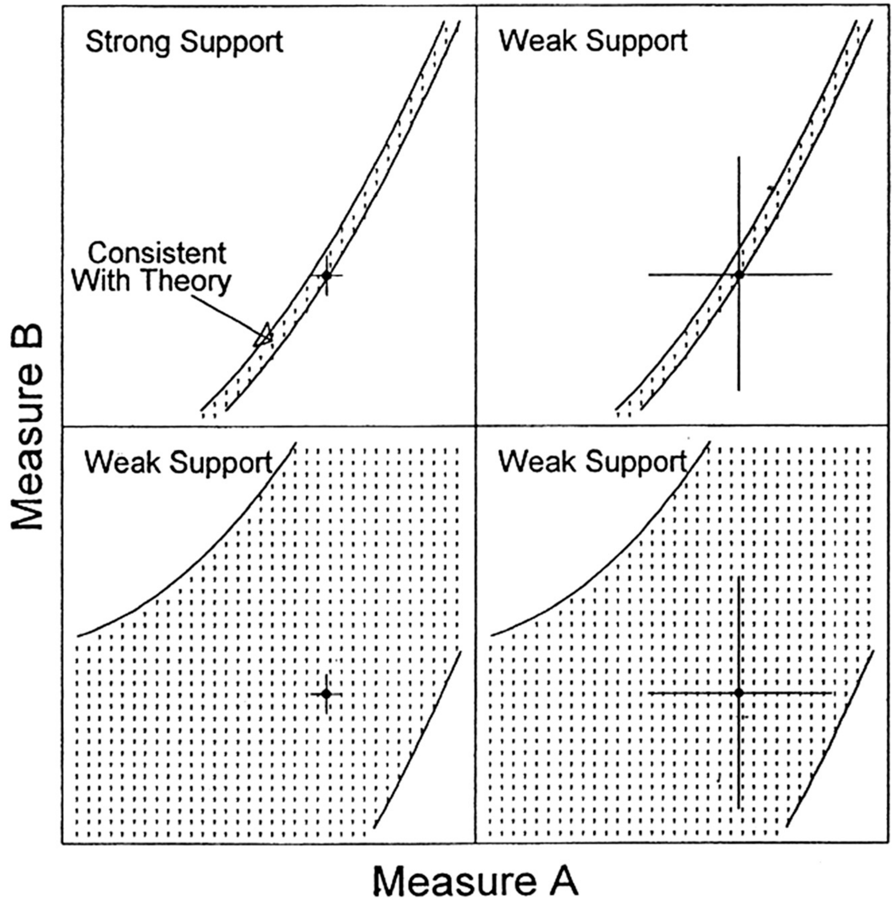
why computational models?

- data never speak for themselves; they require a model to be understood and explained
- verbal theorizing cannot substitute for quantitative analysis
- data can be explained through several alternative models, and we must select among these alternatives
  - “all models are wrong, but some are useful” – George Box
- model selection is based on quantitative evaluation and qualitative judgments
  - quantitative: prediction errors, R^2, mean square error, log likelihood etc.
  - qualitative: less complexity (lower constants/parameters): Occam’s Razor

Farrell & Lewandowsky (2018)

models: scope and falsifiability

- ideally, we want our models/theories to explain as much variance in the data as possible, i.e., have maximal scope
- but… we also want our models/theories to be able to separate signal from noise (hits vs. false alarms!), i.e., models/theories need to be falsifiable, not false
- examples of falsifiable or non-falsifiable theories?



Farrell & Lewandowsky (2018)

models are cognitive aids

- metaphors are powerful but can be misleading!
- what metaphors have we encountered already?

Farrell & Lewandowsky (2018)

big takeaways

- jot down your key takeaways from today

next class

- **before** class:
  - *finish:* L8 reading
  - *research*: sub-domain of cognition
- **during** class:
  - President Safa Zaki!