Inference: Conscious and Unconscious

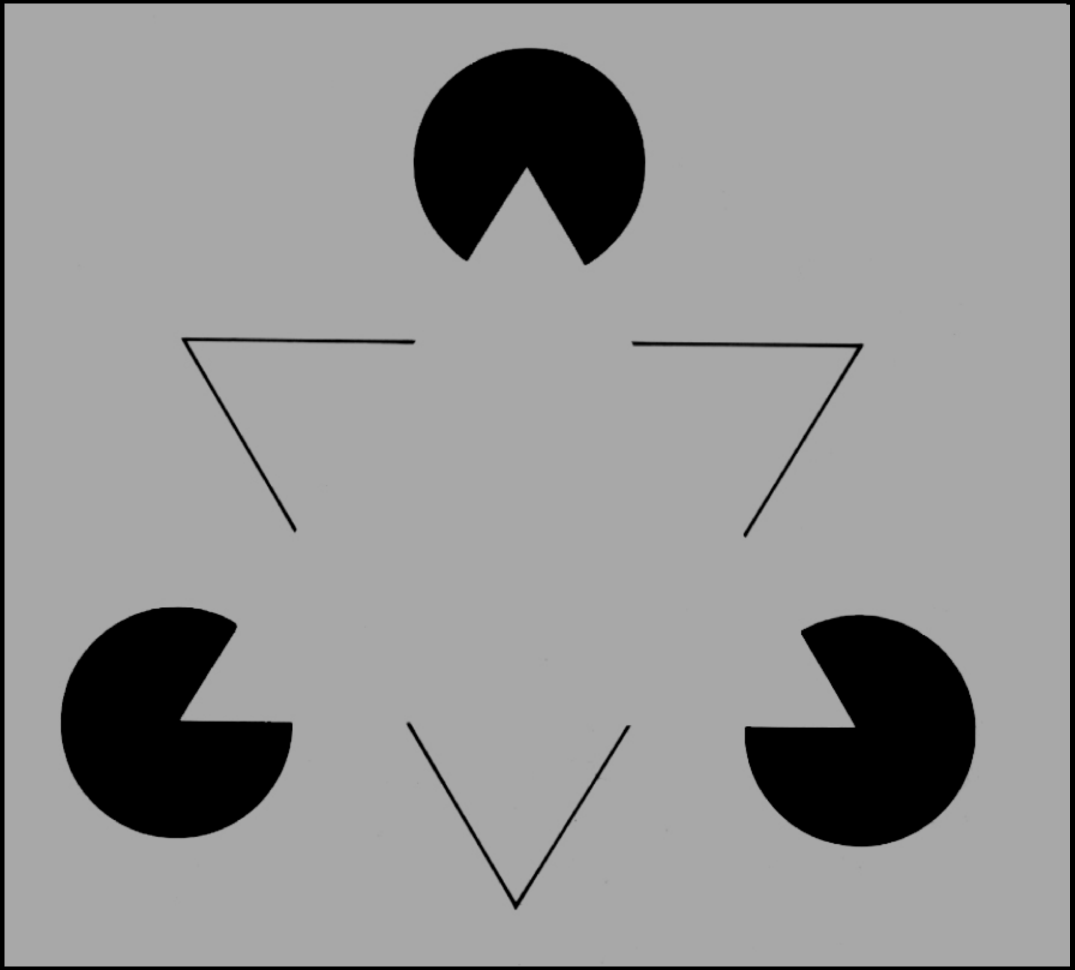
J. McClelland
SymSys 100

April 20, 2010

A Conscious Inference

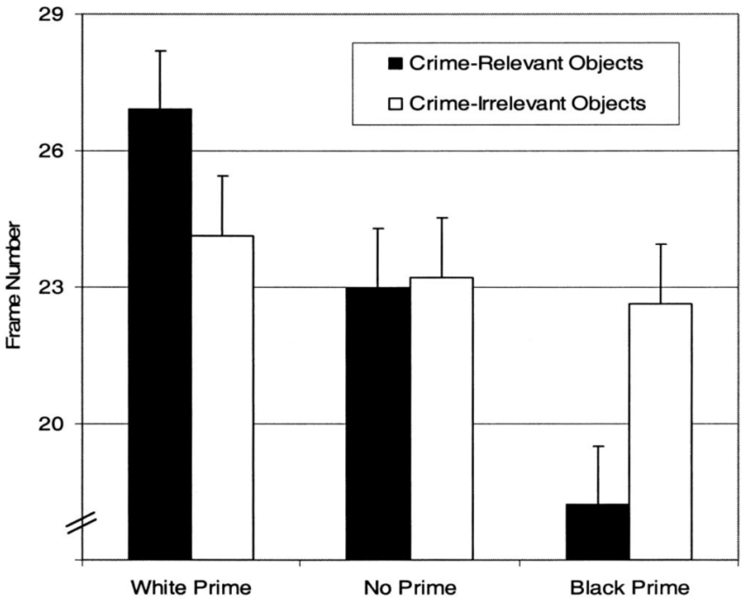
- I wake up in morning and ask myself *has it rained overnight*?
- I look out the window and notice that the ground is wet.
- I consider whether I should conclude that it rained last night.
- I consider the fact that it is April, and rain is relatively rare.
- I consider that I have a sprinker system, and ask myself if it is set to operate on Tuesdays.
- I conclude that any conclusion I can reach by looking at my own back yard would be tentative; so I go outside and notice that the street and sidewalks along my block are all wet.
- I conclude that it did indeed rain last night.

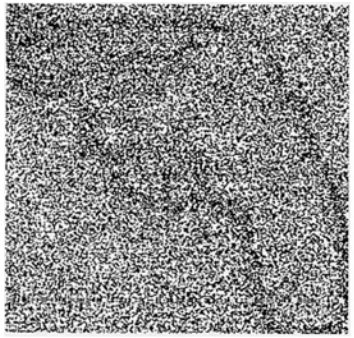
Some Examples of “Unconscious Inferences”



Eberhardt: *Seeing Black*

- Participants were shown very brief presentations of black faces, white faces, or nothing (20 stimuli).
- They then identified pictures that gradually became clearer frame by frame (500 msec/frame).
- Exposure to black faces led to faster recognition of crime objects, even though participants were not aware that they had been shown any faces.





Frame 1

Frame 20

Frame 41

Perception as 'Unconscious Inference'

- Helmholtz coined the term in the 19th century, drawing on ideas going back to the ancients.
- The terms suggests a 'little man in the head', entertaining hypotheses, weighing evidence, and making decisions.
- But we can think of it differently: perhaps your brain is wired to
  - Use context together with direct information
  - Combine multiple sources of information
  - Allow 'hypotheses' to exert mutual influences on each other

Rest of This Lecture

- Examine a “normative theory” of inference under uncertainty
  - How can we reason from given information to conclusions when everything we know is uncertain?
  - How can we combine different sources of evidence?
- Consider two experiments that produce data consistent (in a sense) with the normative theory.

Next two lectures

- A neuro-mechanistic examination of unconscious inference
  – How neurons can produce output matching the normative theory
  – How networks of neurons can work together to produce collective behavior consistent with the normative theory.

An inference question

- An apartment building has a surveillance camera with automatic face recognition software. If the camera sees a known terrorist, it will trigger an alarm with 99% probability. If the camera sees a non-terrorist, it will trigger the alarm 1% of the time.

- Suppose somebody triggers the alarm. What is the chance he/she is a known terrorist?

The Normative Theory
Bayesian Inference: Definitions

- Hypothesis (It’s a terrorist)
- Evidence (Alarm rang)
- Probability:

  Expected Ratio of Occurrences to Opportunities

  NB: Sometimes these are *estimated* or *subjective* probabilities
- P(E|H): Likelihood

  = P(E&H)/P(H): In this case: .99
- P(E|~H)

  = P(E&~H)/P(~H): In this case: .01
- P(H|E)

  = P(E&H)/P(E)
- Can we compute P(H|E)?

Deriving Bayes' Rule

- The definition P(H|E) = P(E&H)/P(E) implies P(H|E)P(E) = P(E&H)
- Similarly, P(E|H) = P(E&H)/P(H) implies P(E|H)P(H) = P(E&H)
- It follows that P(H|E)P(E) = P(E|H)P(H)
- Divide both sides by P(E) we obtain P(H|E) = P(E|H)P(H)/P(E)
- Final Step: Compute P(E)

  In our case, the probability that the alarm goes off when a person walks by the camera

- Can we derive P(E) from quantities we've already discussed?

  P(E) = P(E|H)P(H) + P(E|~H)P(~H)

- Substituting, we get Bayes' Rule:

$$P(H \mid E) = \frac{P(E \mid H)P(H)}{P(E \mid H)P(H) + P(E \mid \sim H)P(\sim H)}$$

Importance of ‘priors’ or ‘base rates’

- If terrorists and non-terrorists are equally likely, $p(H|E) = .99$
- If only one person in 100 is a terrorist, $p(H|E) = .5$
- If only one person in 10,000 is a terrorist, $p(H|E)$ is ~.01!

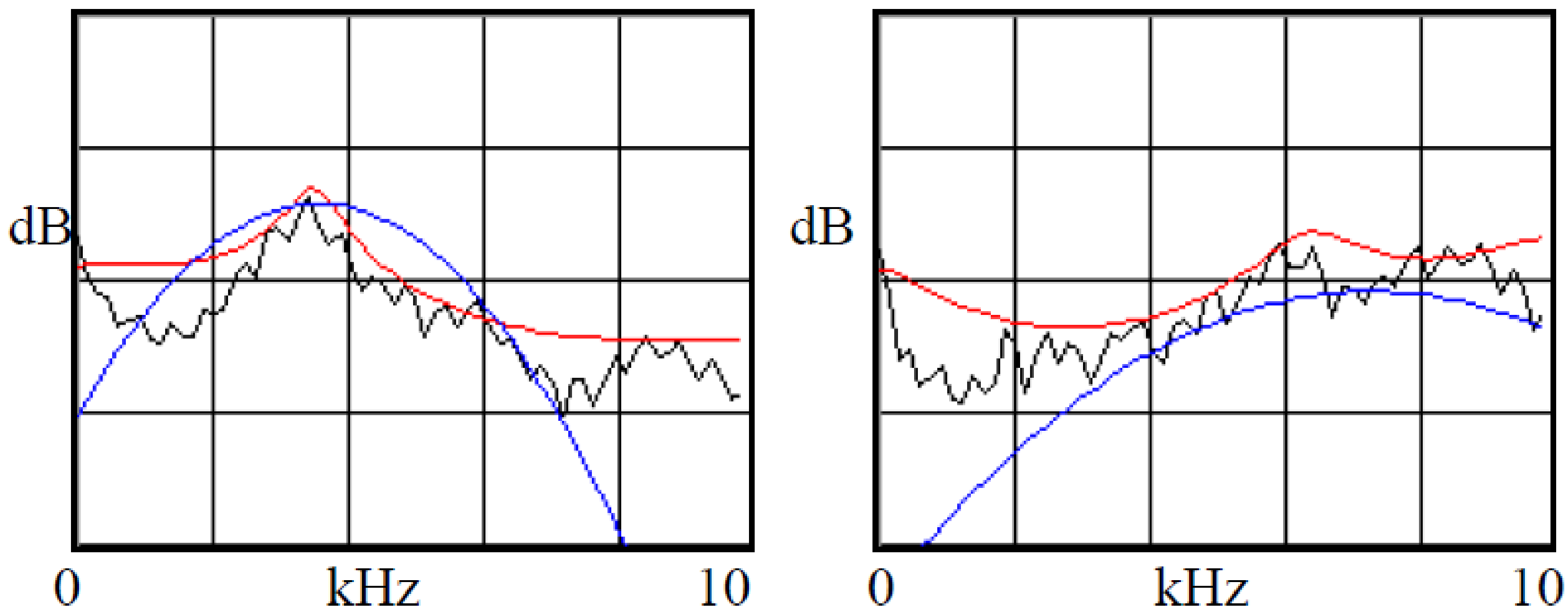
Lessons From this Example

- Priors are crucial if we are to derive a correct inference from the data!
- For items with low 'base rates' (priors), people often overestimate the posterior, p(H|E).
- In explicit inferences, people often neglect base rates (or underweight them)

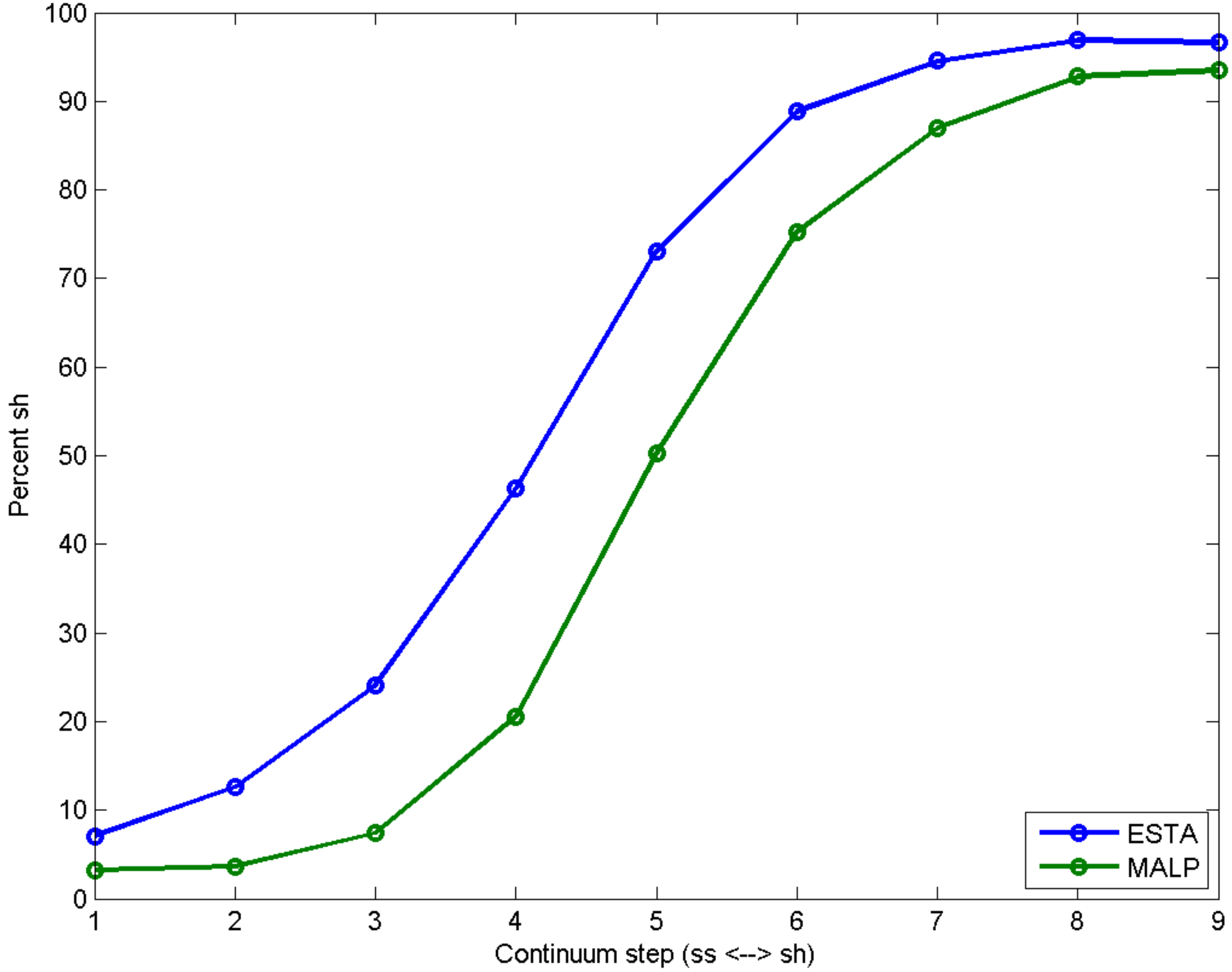
Ganong Experiment

- Spoken words such as ‘establish’ or ‘malpractice’ are presented but the final fricative sound is replaced with an ambiguous sound, somewhere on a continuum from ‘ss’ to ‘sh’.
  - Stimuli are blends of natural ‘ss’ and ‘sh’ sounds.
- The participant must decide: is the last sound ‘ss’ or ‘sh’.

Natural ‘sh’ and ‘ss’



Data from experiment performed last year: P('sh') for 'establi..' and 'malpracti…' contexts



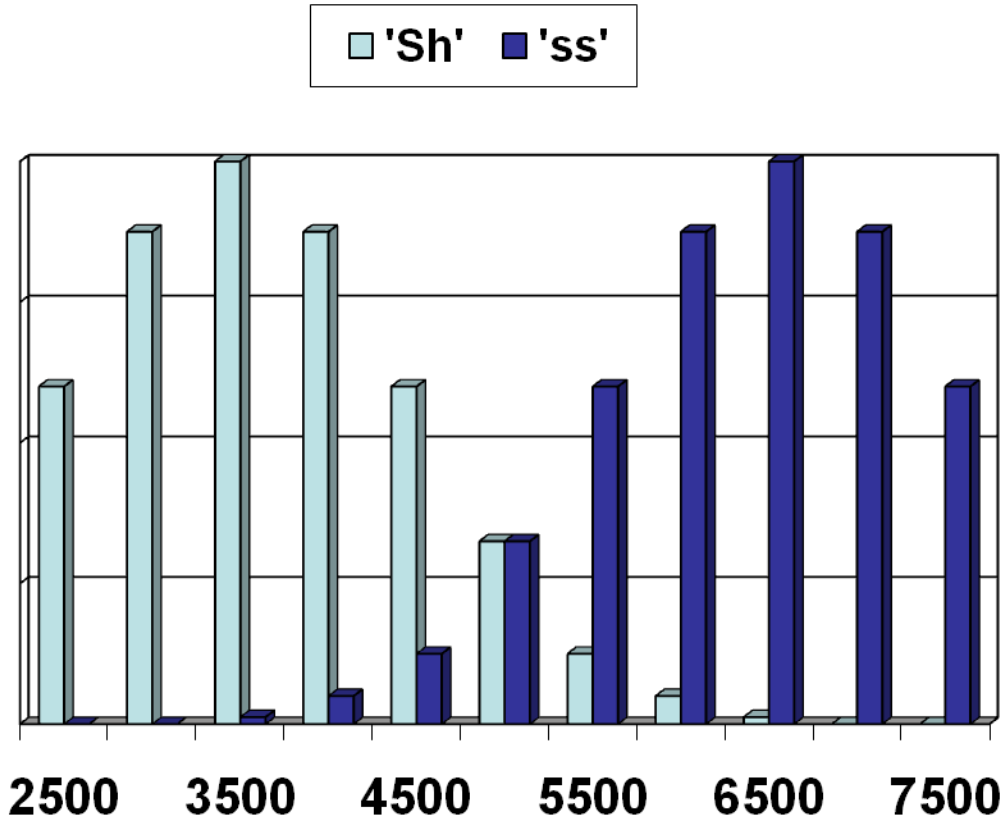
Can we understand these data from a Bayesian Perspective?

- We can think of the different stimuli used as differing in their probability, given that the stimulus was ‘sh’ or ‘ss’.
- We can think of ‘establi…’ increases the prior probability of ‘sh’ relative to ‘ss’
- Similarly, we can think of ‘malpracti…’ as increasing the prior probability of ‘ss’ relative to ‘sh’
- The effect is substantial, but not overwhelming:

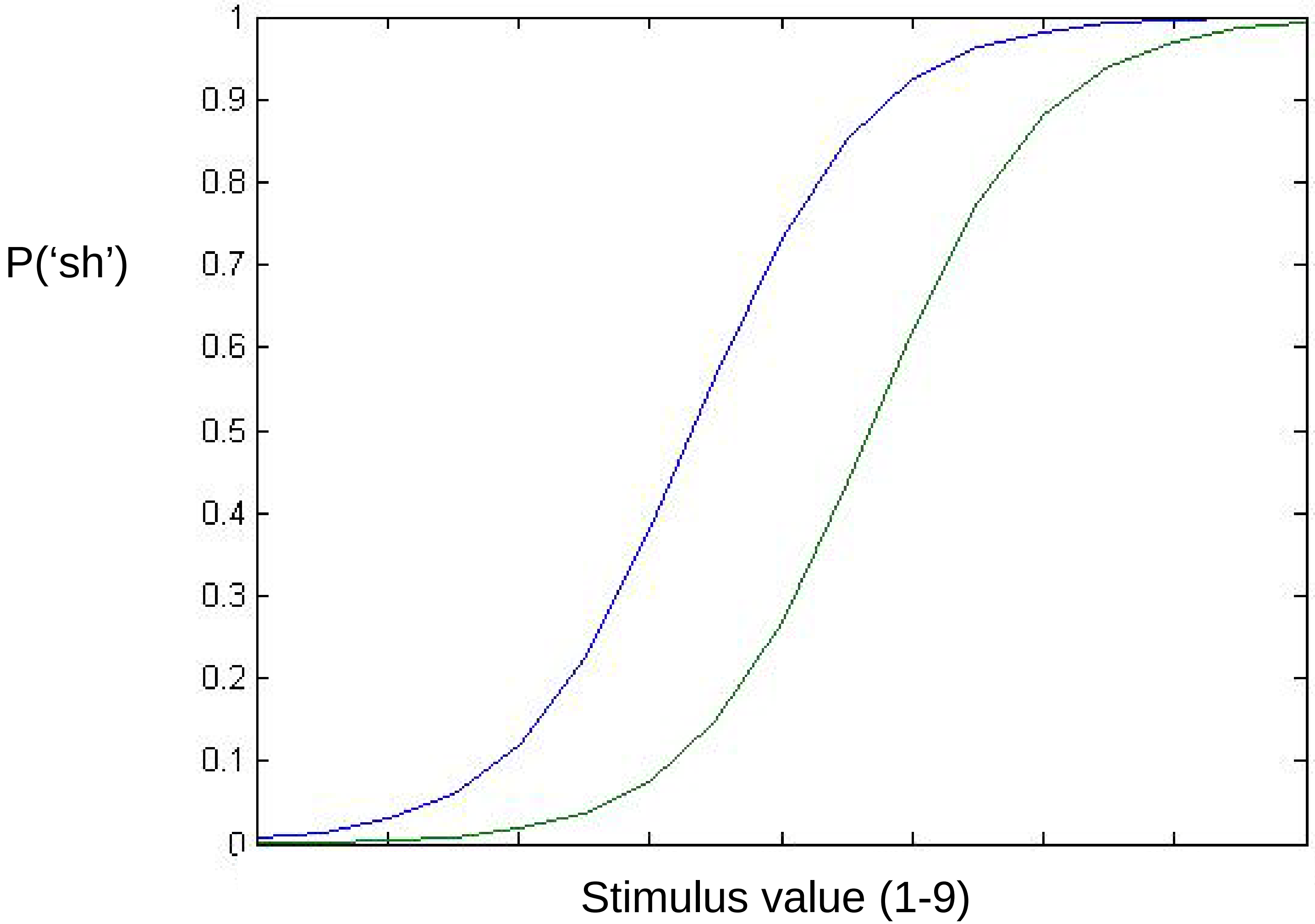
  p(sh|establi…)/(p(sh|establi…)+p(s|establi…)) = .63
  p(sh|establi…)/p(s|establi…) = .63/.37 ≈ 1.7

- Is the effect too big, or is it too small?

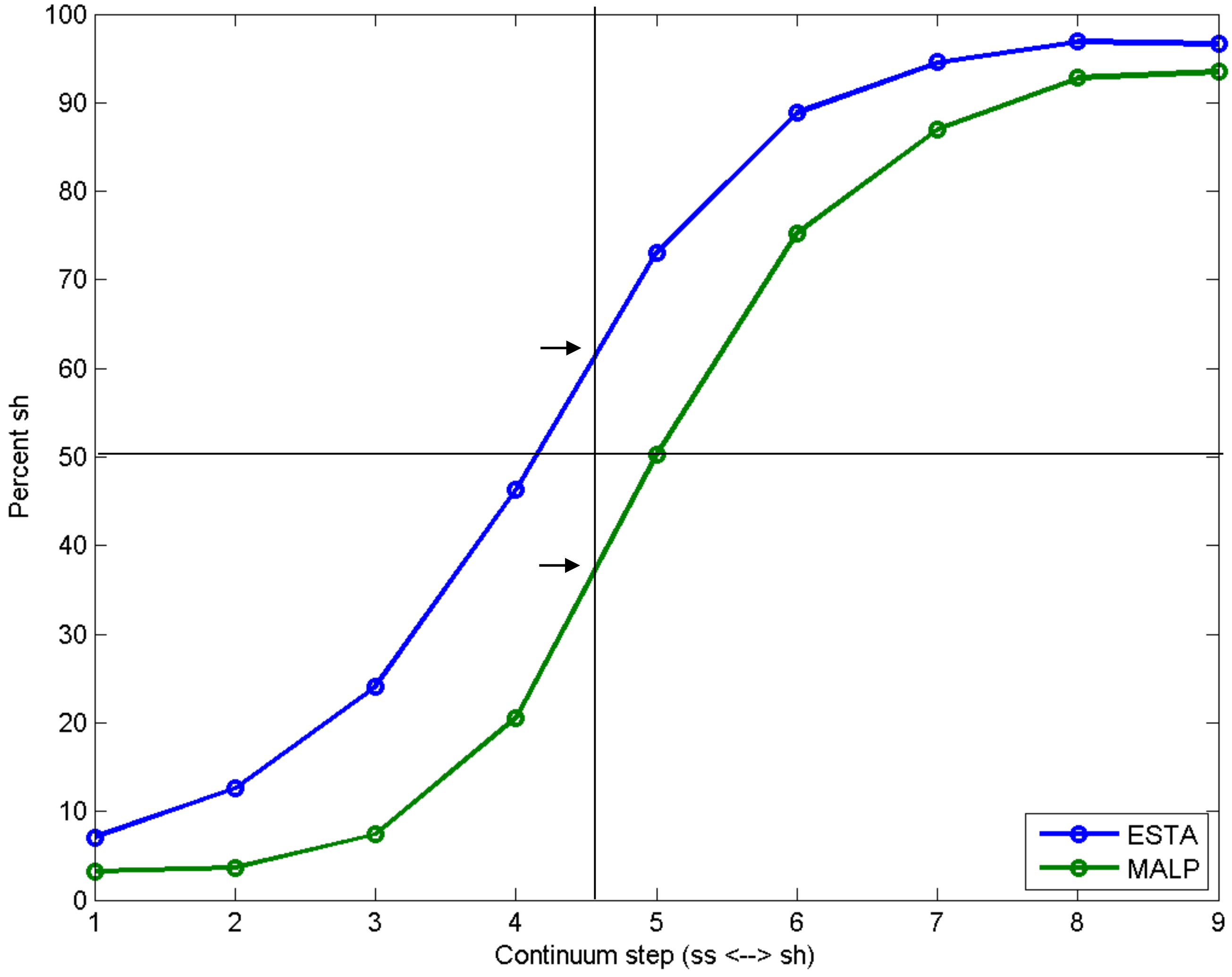
P(stim|sh) & P(stim|ss) [Hypothetical!]



Expected form of the data from Ganong experiment



Data from experiment performed last year: P('sh') for 'establi..' and 'malpracti…' contexts



How might we combine information from different sources?

- For example, visible and heard speech (The *McGurk* effect)
- Auditory cue: starting frequency of syllable
- Visible cue: degree of mouth opening
- Treat A and V as *conditionally independent*
  - *P(Ea&Ev|H) = P(Ea|H)P(Ev|H) for H = ‘ba’ or ‘da’*
- Then we get this normative model:

$$P('da' \mid Ea \,\&\, Ev) = \frac{P(Ea \mid 'da')P(Ev \mid 'da')P('da')}{P(Ea \mid 'da')P(Ev \mid 'da')P('da') + P(Ea \mid 'ba')P(Ev \mid 'ba')P('ba')}$$

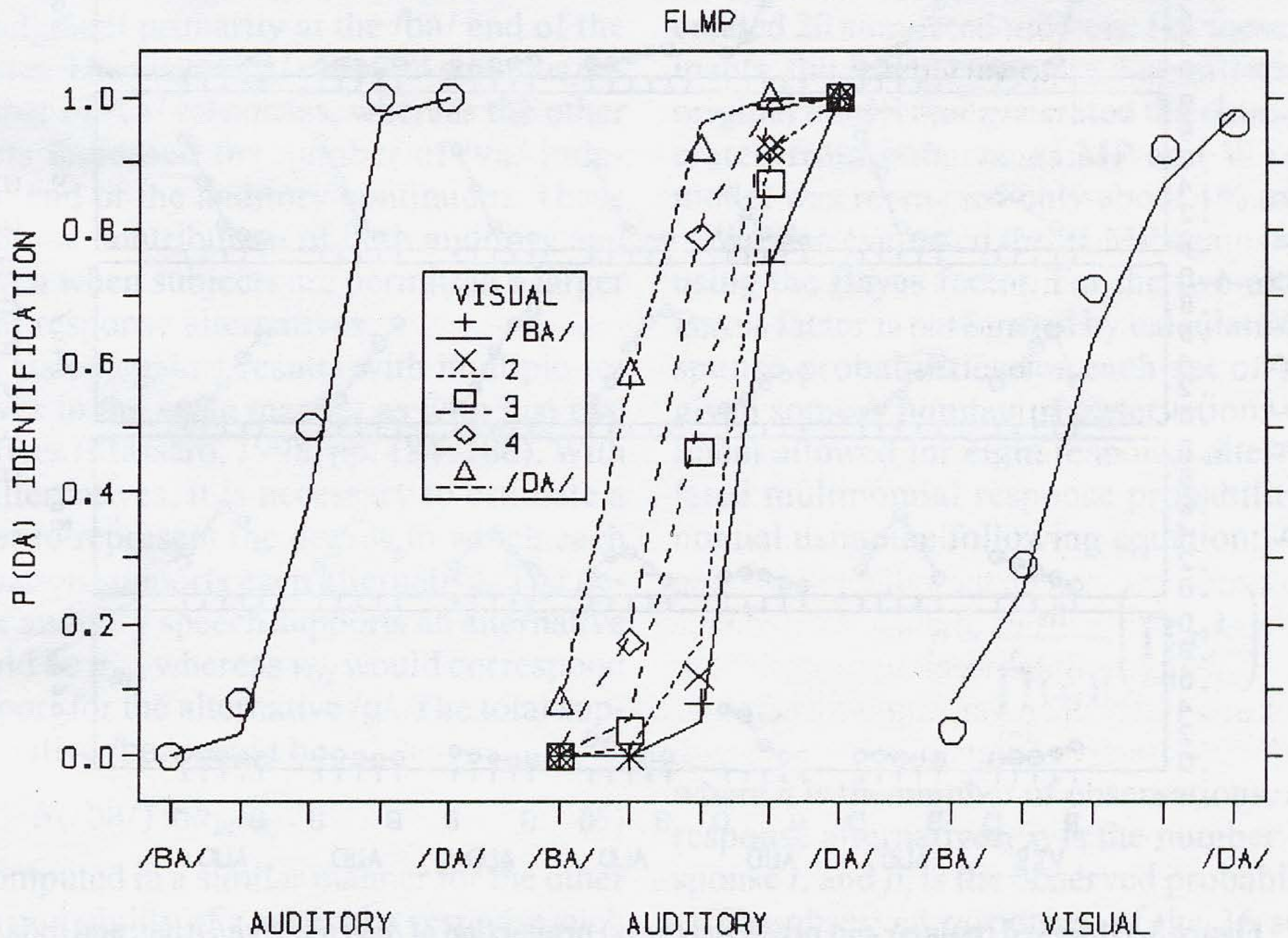


Figure 2. The points give the observed proportion of /da/ identifications for a typical observer in the auditory-alone (left panel), the factorial auditory–visual (center panel), and the visual-alone (right panel) conditions as a function of the five levels of the synthetic auditory and visual speech varying between /ba/ and /da/. The lines give the predictions of the FLMP.

Summary

- People make unconscious inferences all the time.
- The responses they make often approximately match an estimate of the posterior, assuming that context adjusts their priors towards alternatives that are consistent with the given context.
- People also often combine different sources of information in a way that is consistent with Bayes rule, and the additional assumption that the different sources of evidence are treated as conditionally independent.
- Next time we will consider how neurons in the brain could carry out such computations.