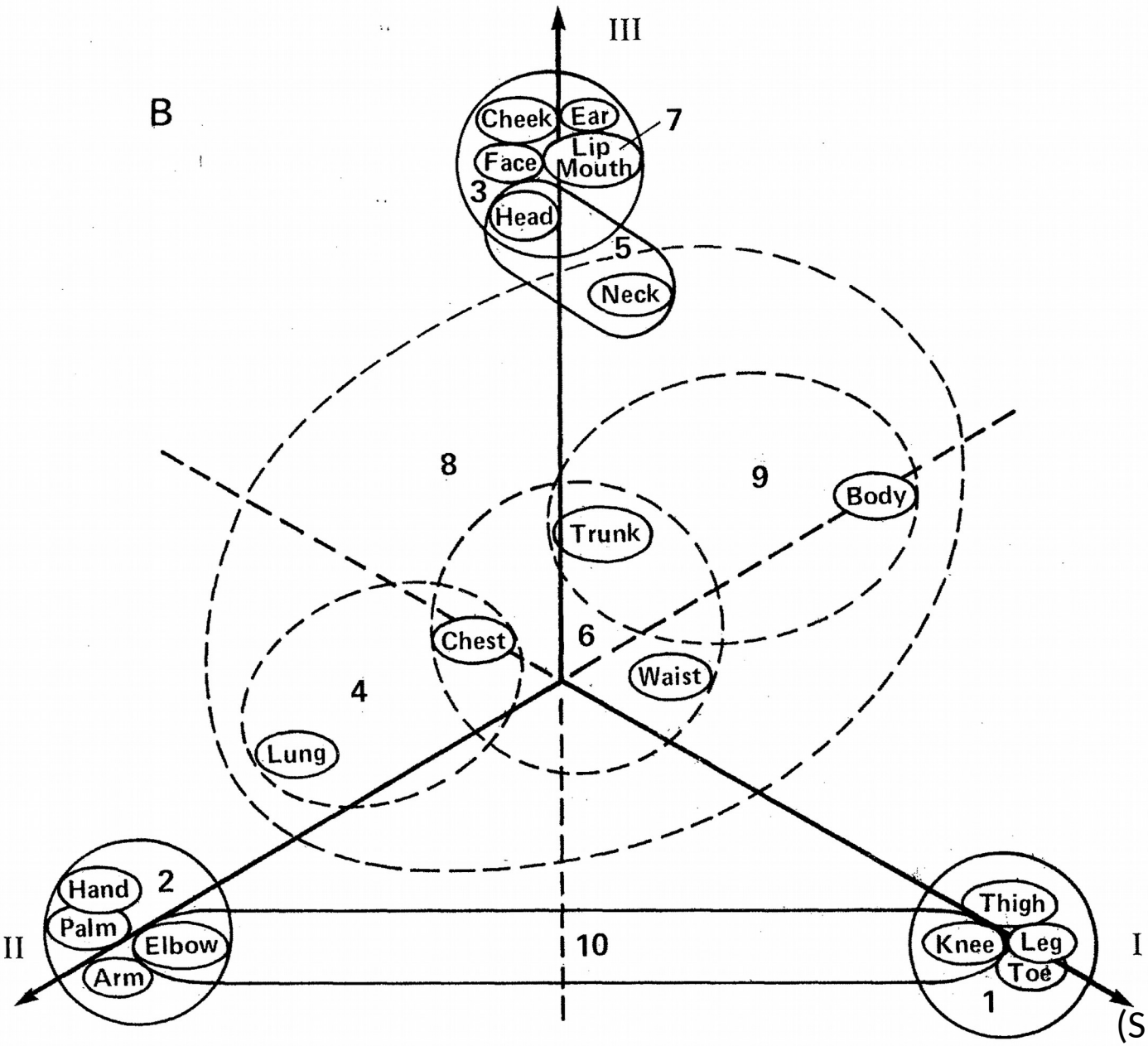
Universal law of Generalization

Cogsci 131



(Shepard 1980)

Property induction

- **One example:**

  Cows have T9 hormones
  Horses have T9 hormones

  How likely are Chimps to have T9 hormones?

  How likely are butterflies to have T9 hormones?

- **Anther example:**

  Birds have T9 hormones
  Horses have T9 hormones

  How likely are Chimps to have T9 hormones?

  How likely are butterflies to have T9 hormones?

General phenomena
(Osherson et al. 1990)

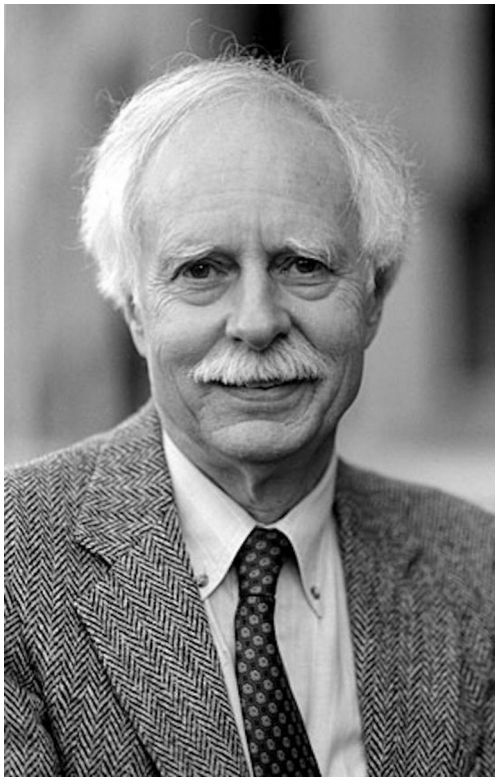
- **Factors that strengthen effects:**
  - Premise typicality
    (robins vs penguins)
  - Premise diversity
    (robins+dolphins vs. robins+bluejays)
  - Conclusion specificity
    (robins+bluejays → birds vs. robins+bluejays → animals)
  - Premise monotonicity
    (hawks+sparrows+bluejays vs hawks+sparrows)

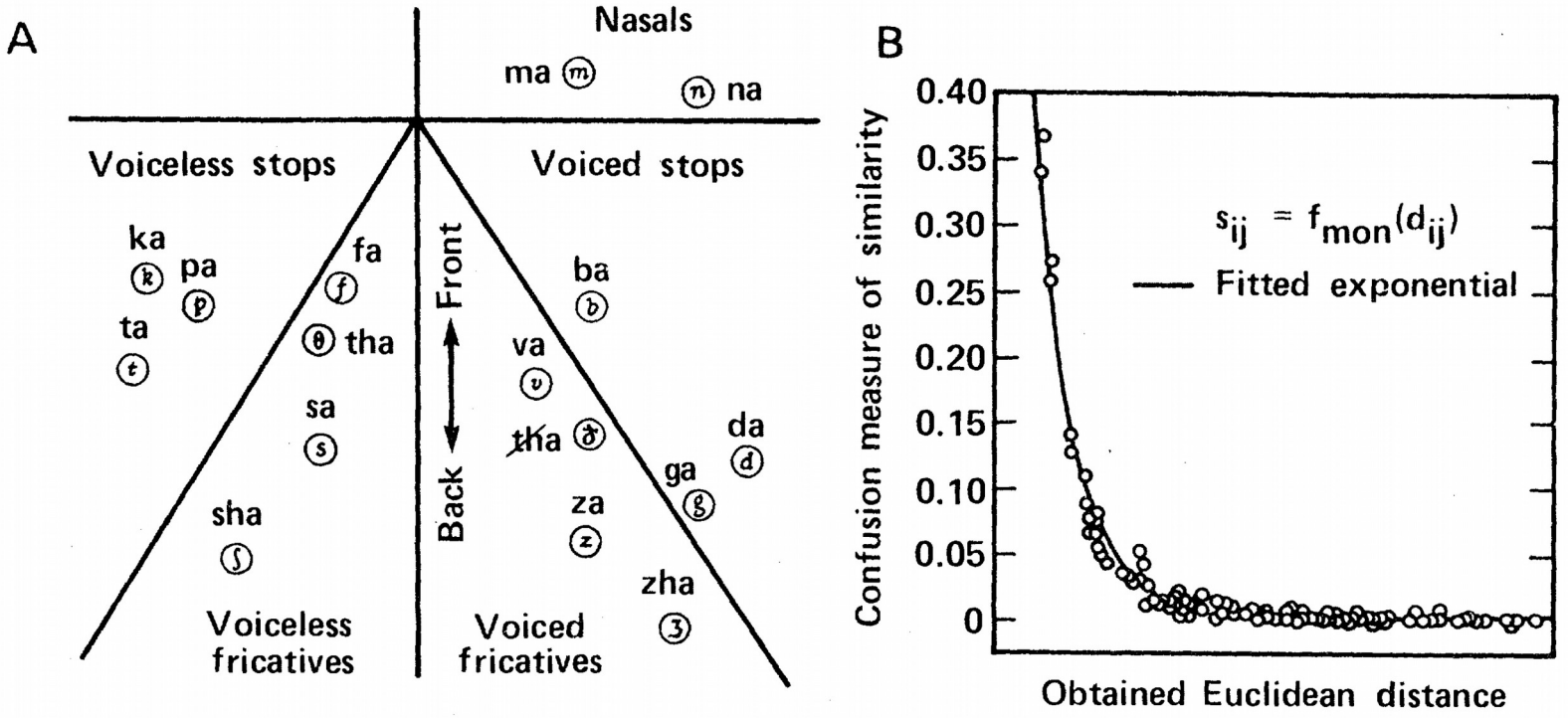
- **Why might these kinds of effects be seen?**

Similarity and Generalization

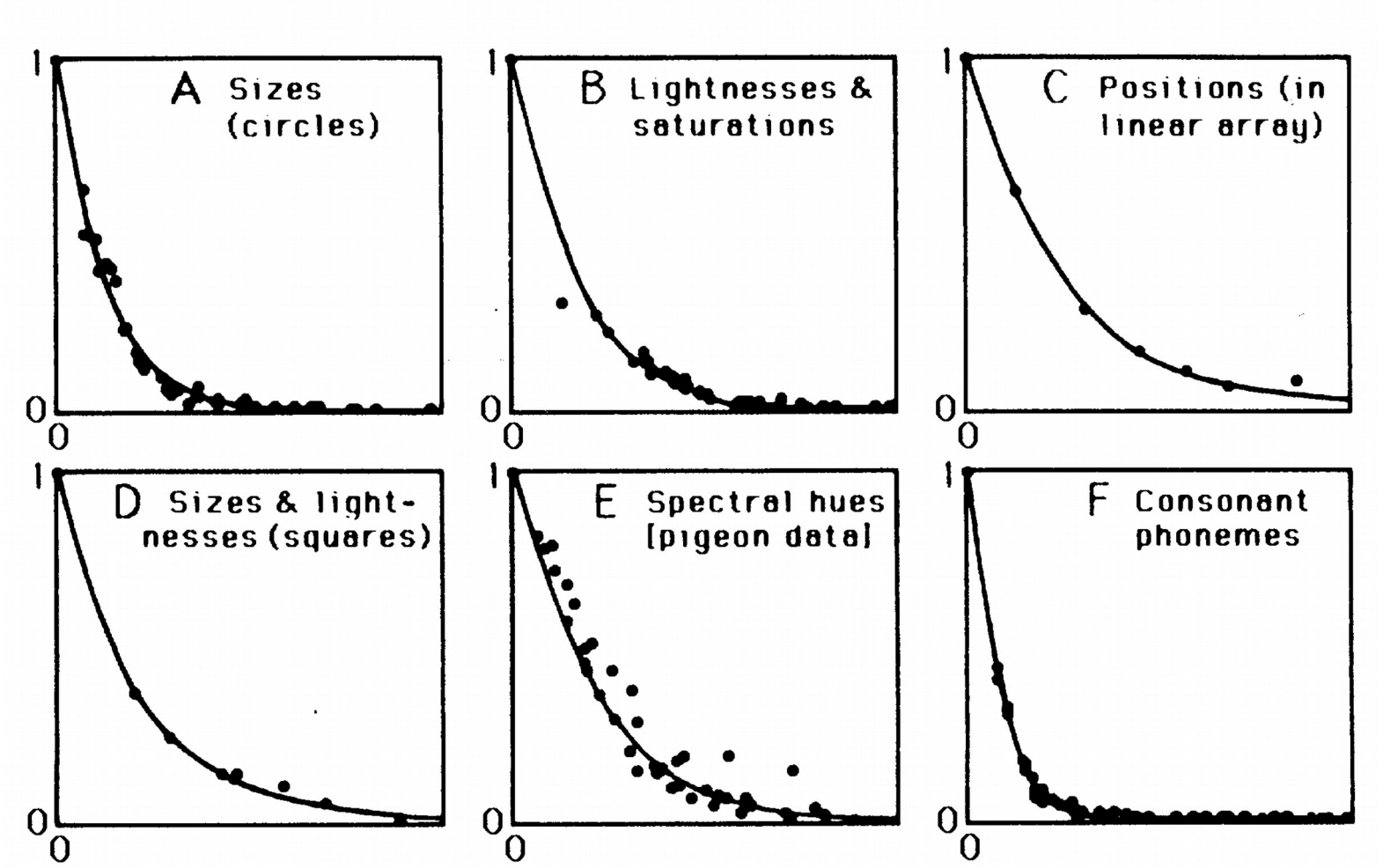
- **Similarity is important in part because it determines generalization on new data.**

- **Shepard's Universal Law:** Generalization drops off exponentially in psychological distance.

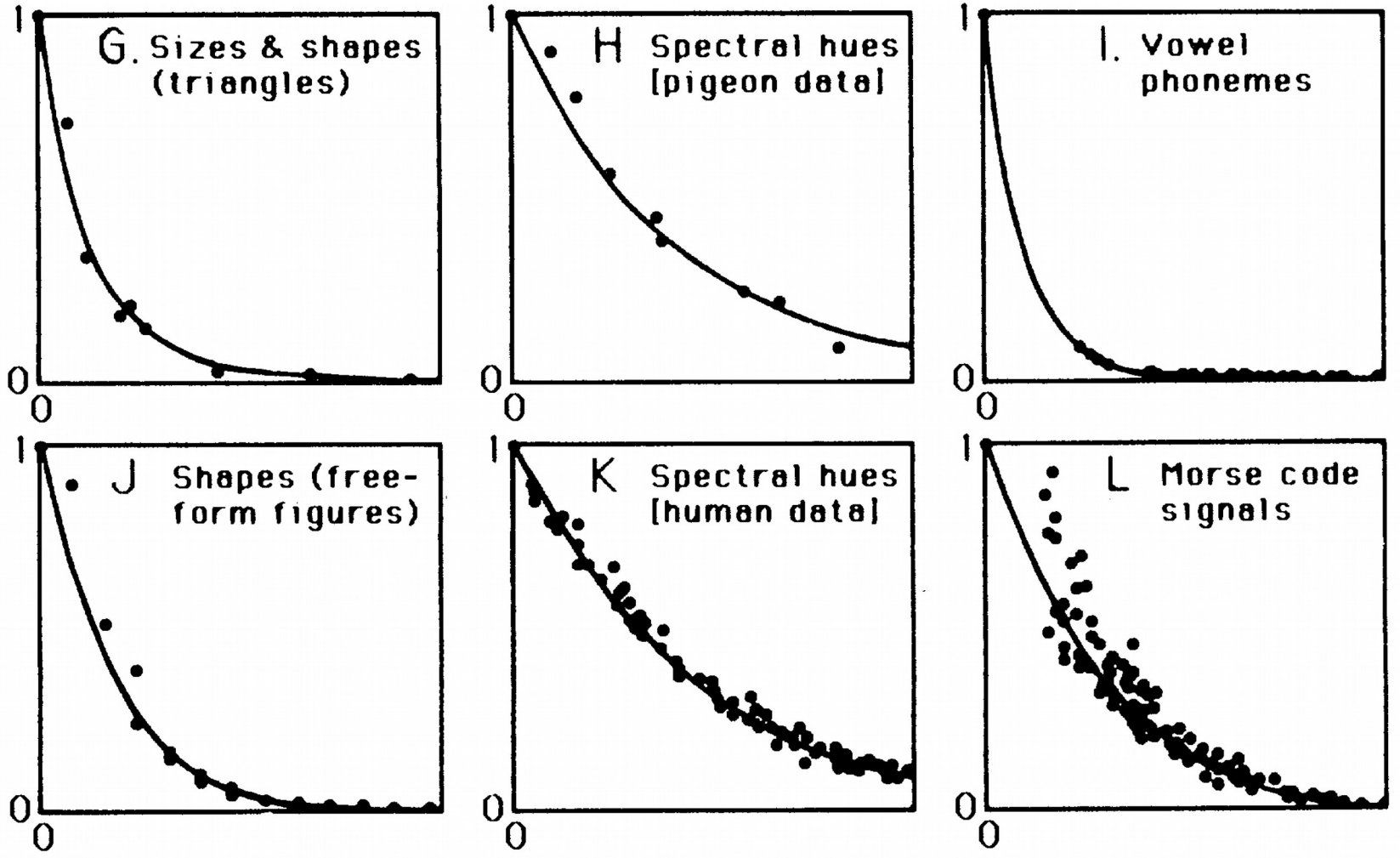




(Shepard 1980)



(Shepard 1987)



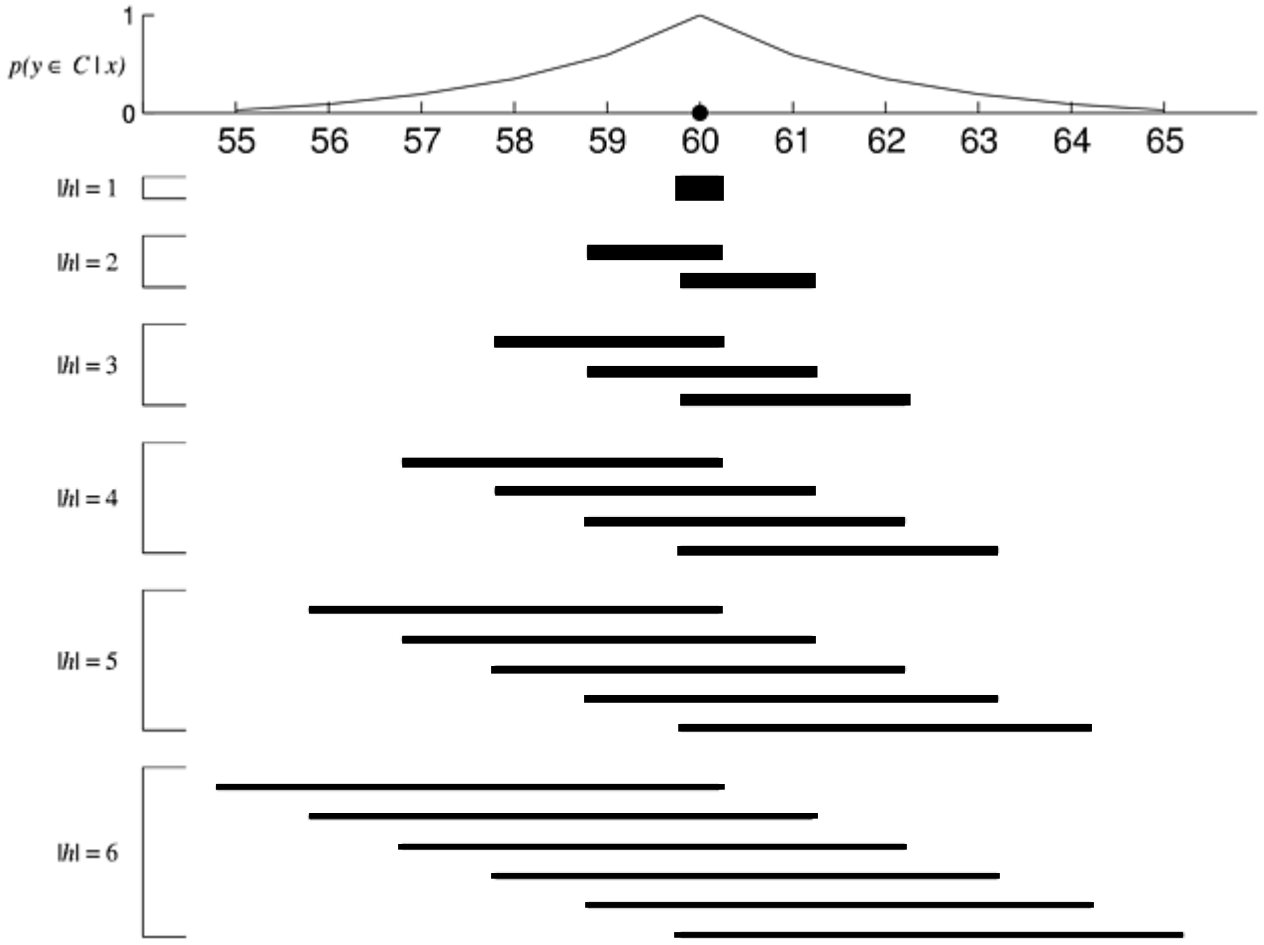
(Shepard 1987)

"How a cognitive psychologist came to seek universal laws", Shepard (2004)

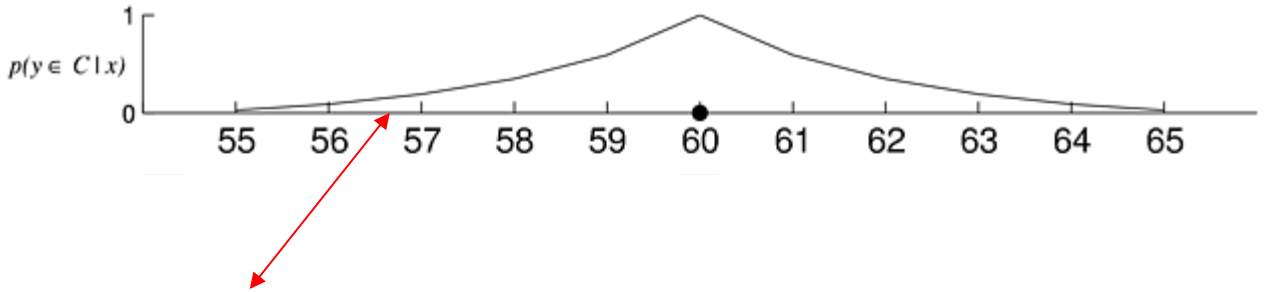
… [one reason] is my unwillingness to be satisfied with any proposed psychological principle whose sole justification is that it fits all the available empirical evidence-whether behavioral or neurophysiological. *I crave, in addition, a reason that that behavioral principle (or that associated neural structure) should have the particular form that it does, rather than some other.* … [I believe] that if, as I fervently hope, psychological principles are not merely arbitrary, some may be shown to have arisen as accommodations to universal features of the world. If this is so, we might aspire to a science of the mind that, like the physical and mathematical sciences, has universal laws.

What level of analysis is this?

Explaining the universal law



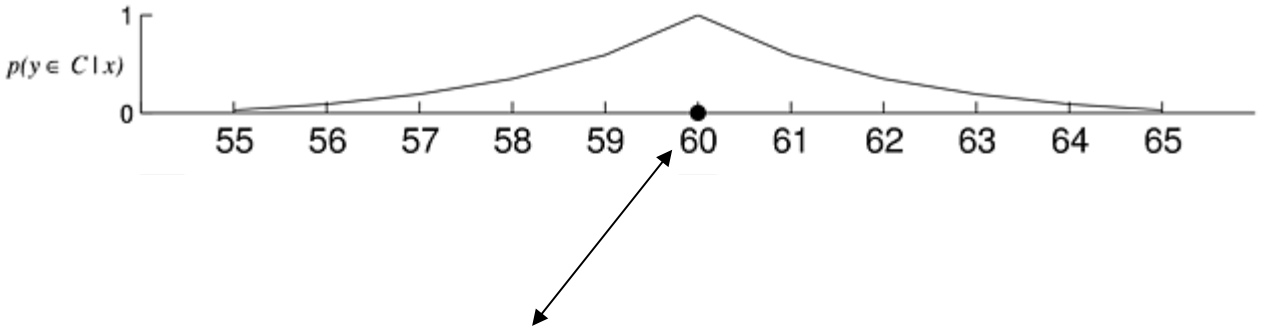
Explaining the universal law



Single dimension (simple case)

We consider only integer stimulus values

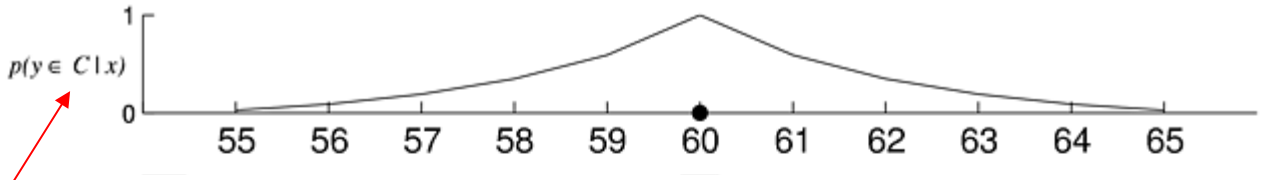
Explaining the universal law



Black dot = stimulus **x**

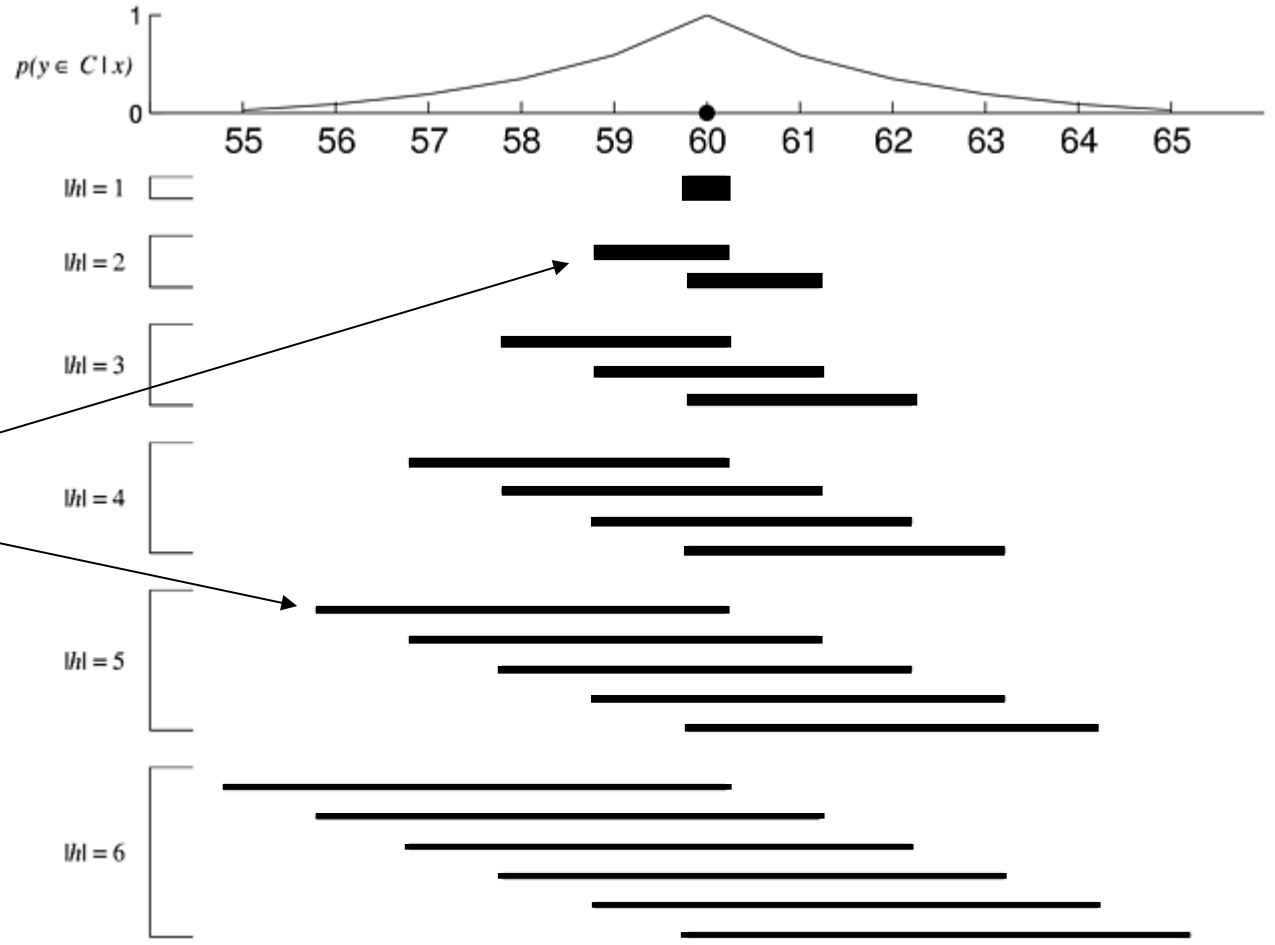
Example: hormone level

Explaining the universal law



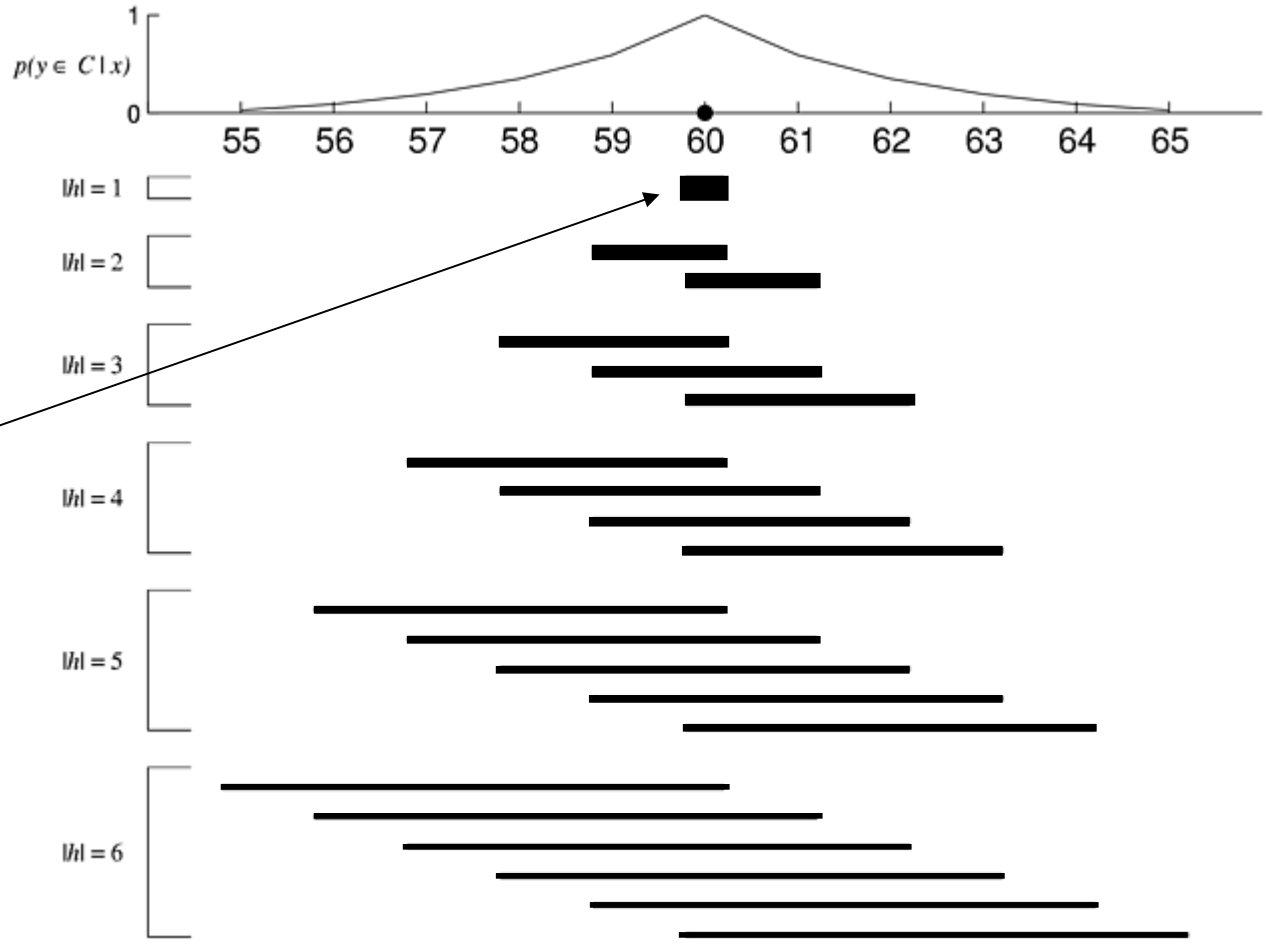
Y-axis:
Probability that y is in the consequential region, **given x is**.

Explaining the universal law



Possible hypotheses (consequential regions) are overlapping intervals of all possible sizes

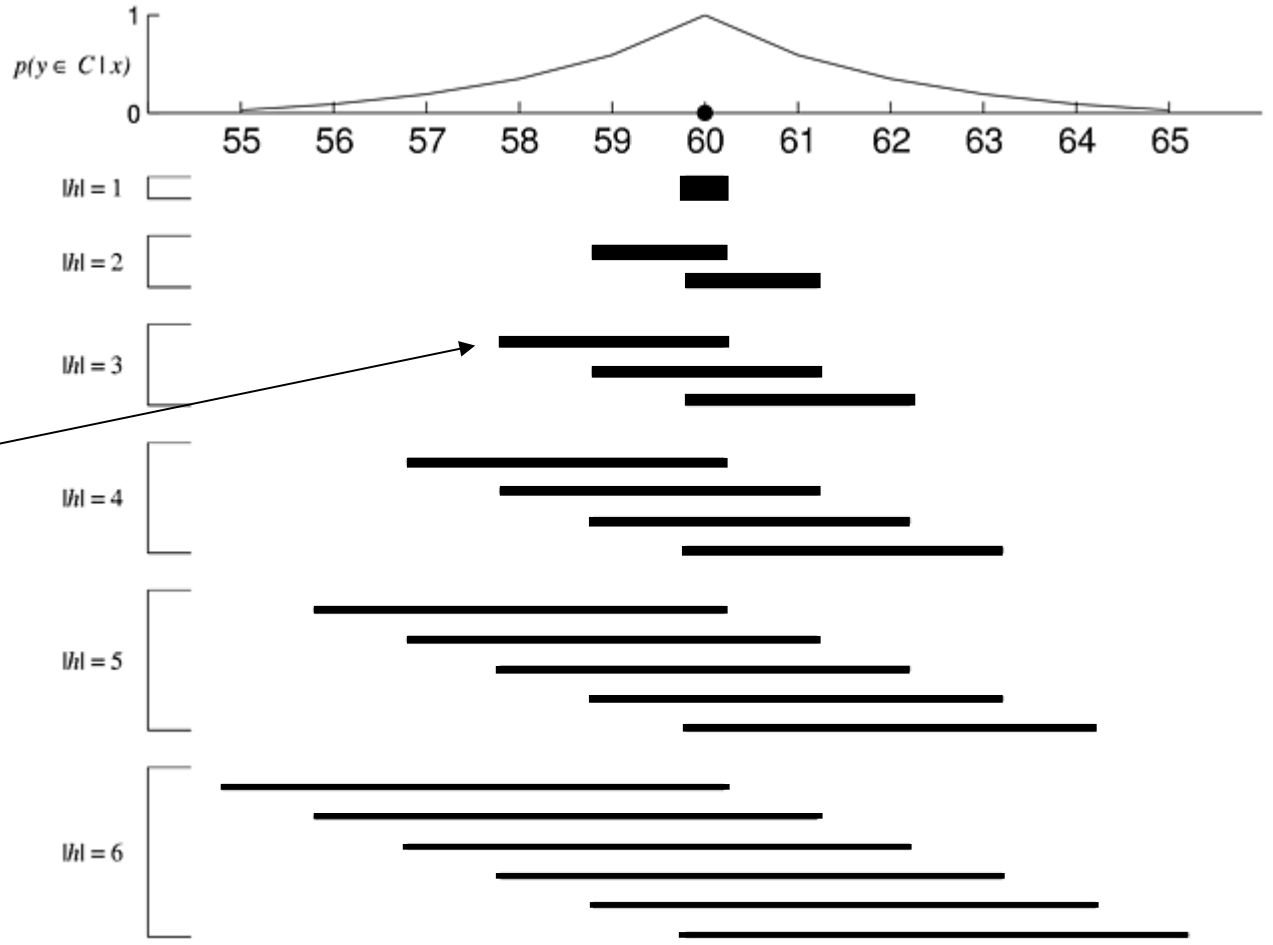
Explaining the universal law



Example 1

Interval includes only the original value of 60

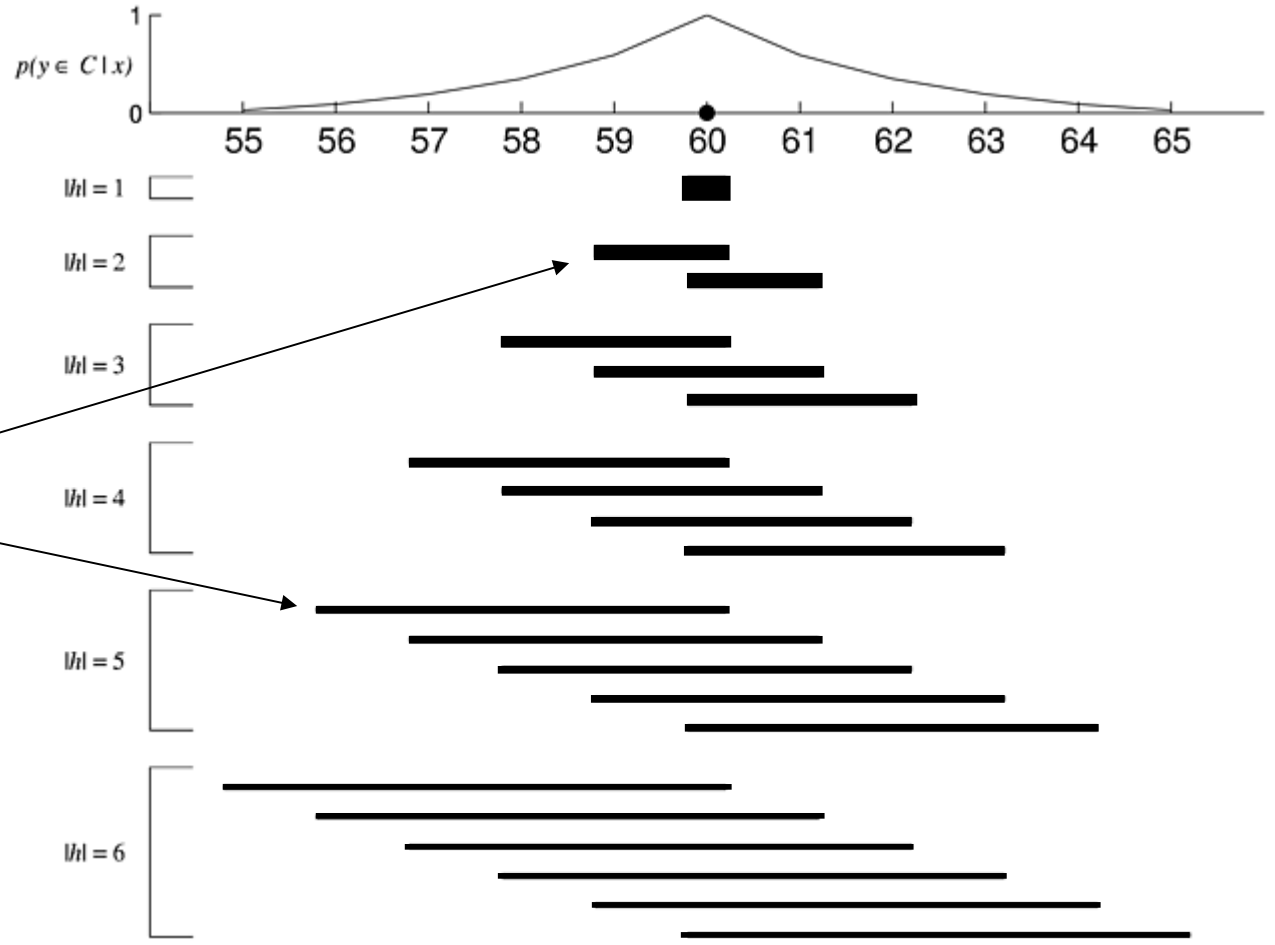
Explaining the universal law



Example 2

Interval includes 58-60 (size of 3)

Explaining the universal law



Smaller intervals more probable than large intervals **(bar thickness**)

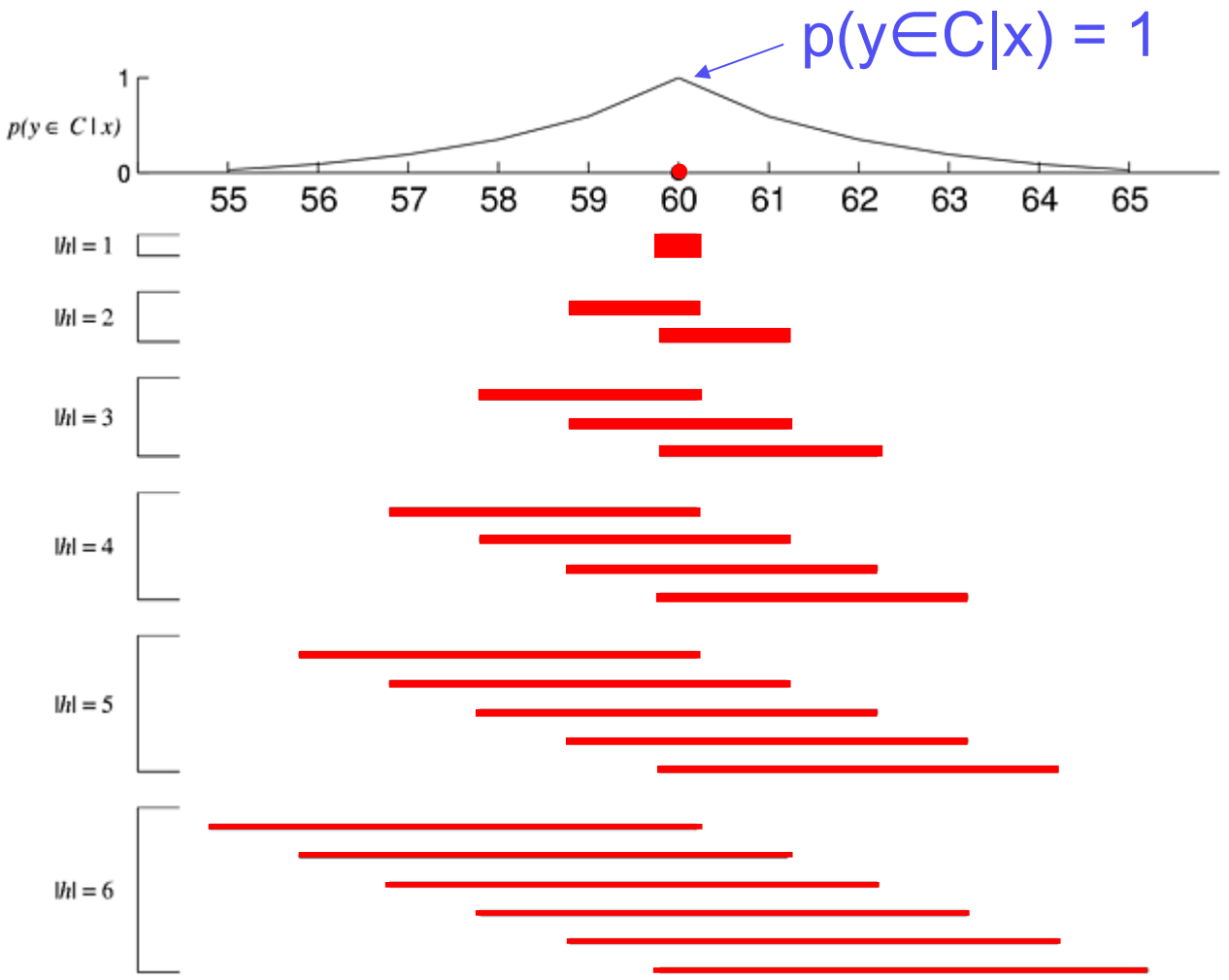
All probabilities sum to 1

Explaining the universal law

Case 1:
Stimulus **y** also has a value of 60

y is inside all possible intervals that include x

Probability that y is in region is 1.



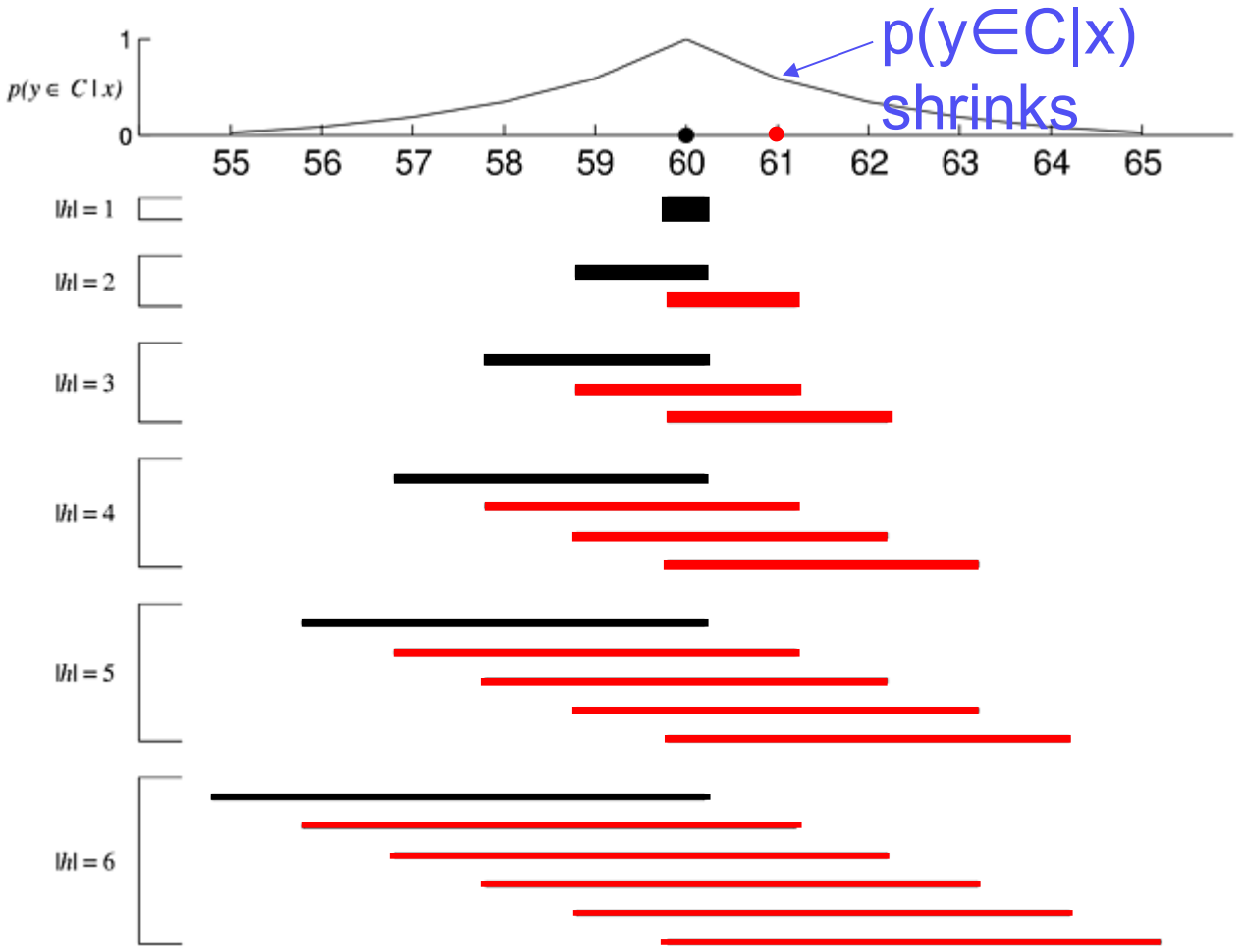
Explaining the universal law

Case 2:
Stimulus **y** has a value of 61

y is inside 16 out of 21 bars

y is inside 6 less bars than before

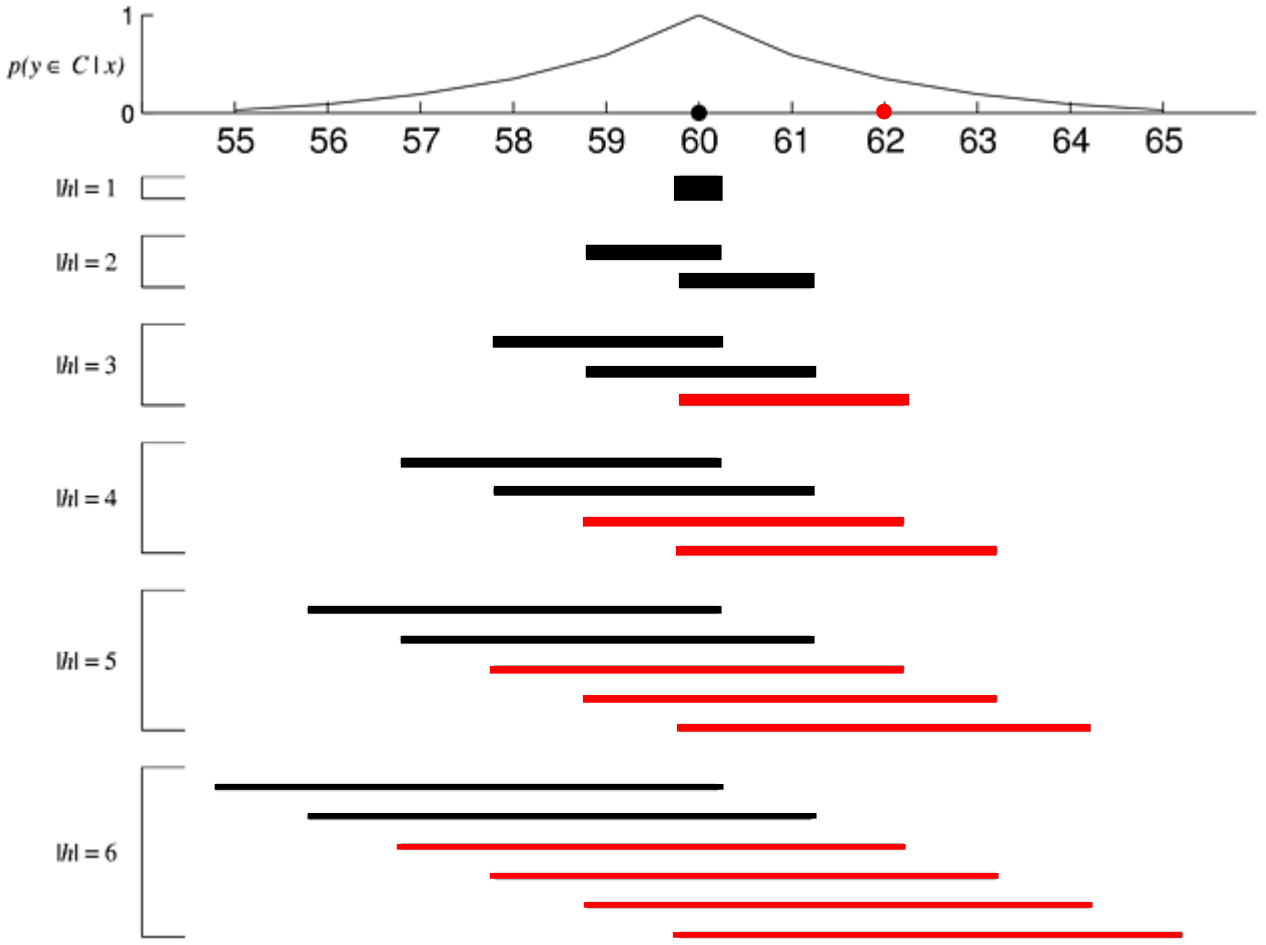
Total probability of being in the consequential region shrinks



Explaining the universal law

Case 3:
Stimulus **y** has a value of 62

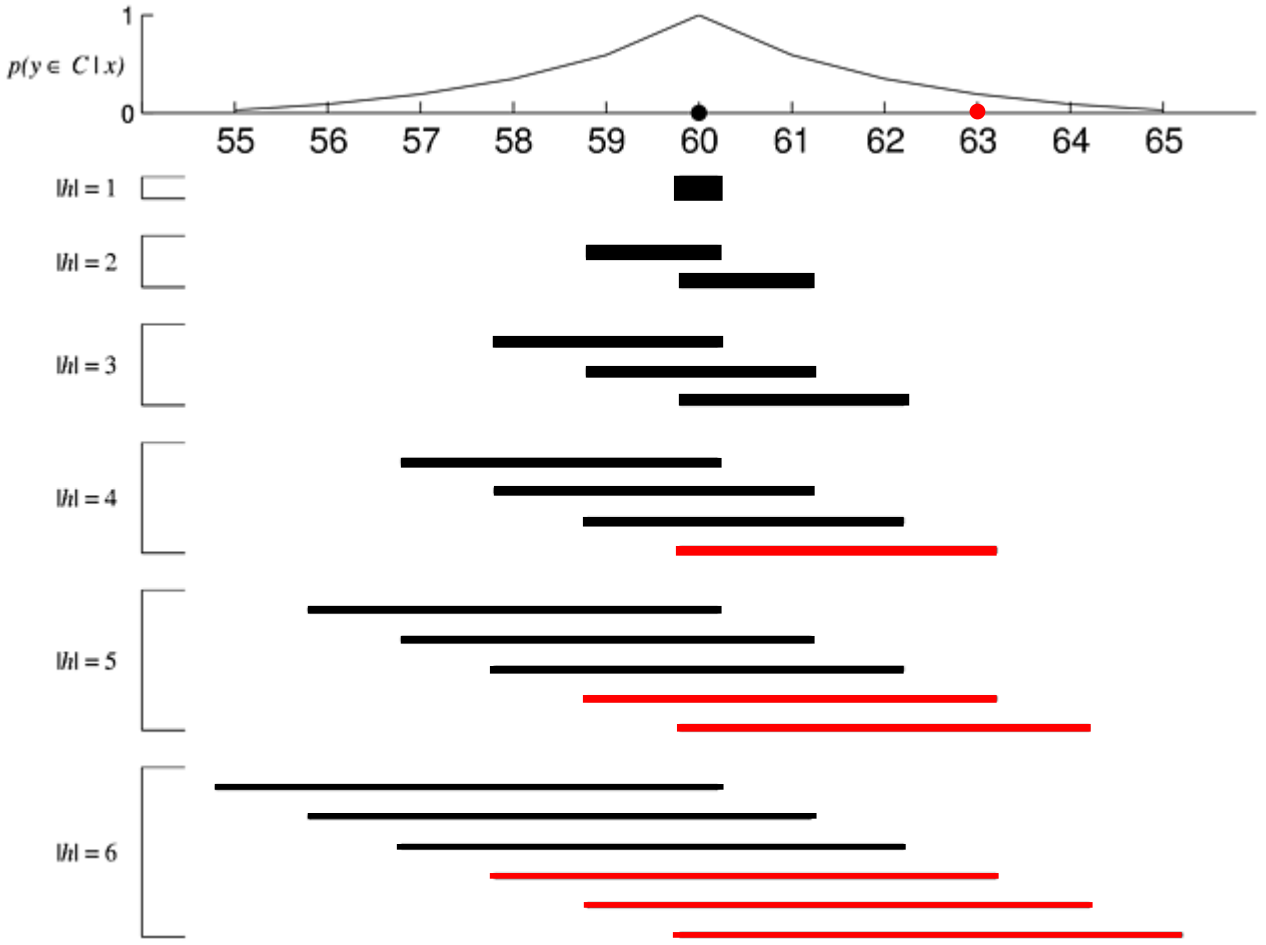
y is inside 5 less bars than before



Explaining the universal law

Case 4:
Stimulus **y** has a value of 63

y is inside 4 less bars than before

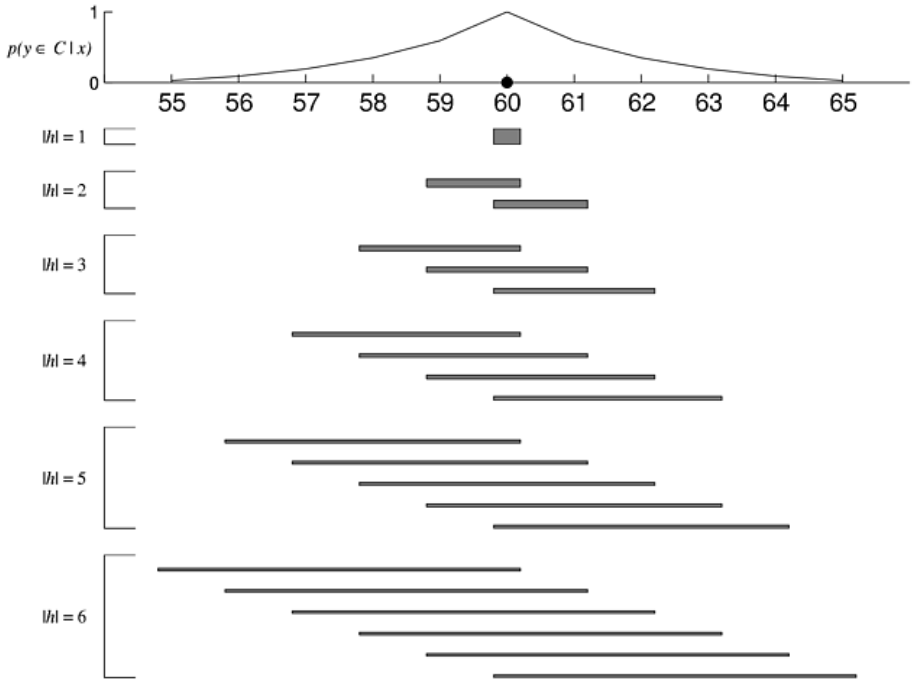


Explaining the universal law

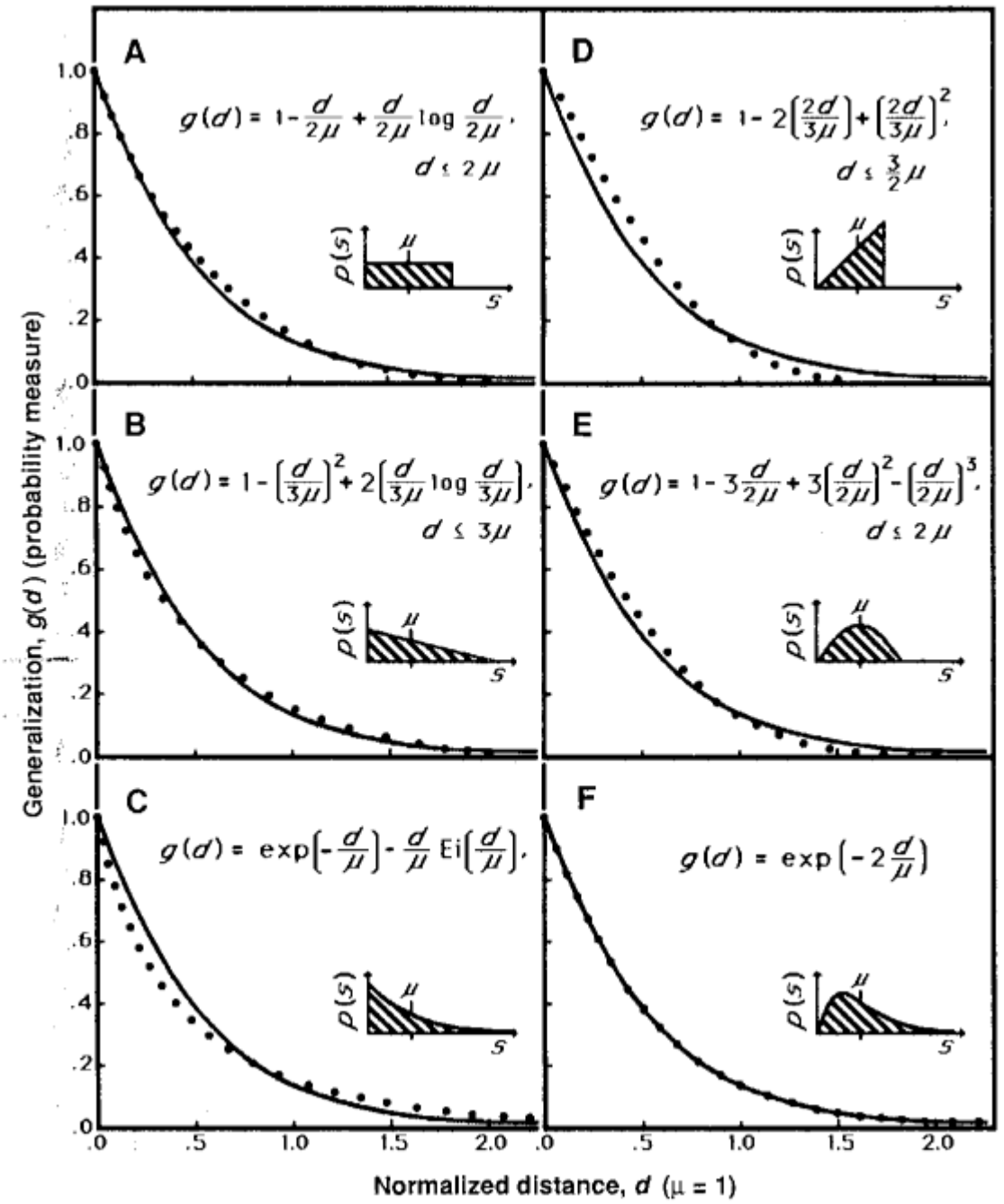
Summary

As y value gets further away from x value, the number of intervals containing it decreases,

BUT, at a decreasing rate **(lost 6,5,4 bars)**.

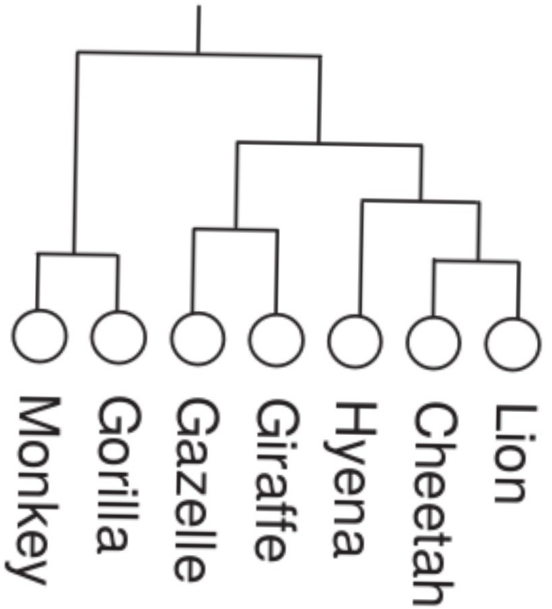


Question: What function decreases at a decreasing rate? One answer: $f(x) = e^{-x}$

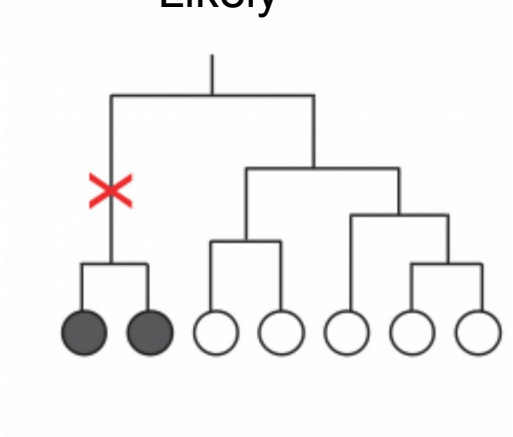

A
$g(d) = 1 - \frac{d}{2\mu} + \frac{d}{2\mu} \log \frac{d}{2\mu}$,
$d \leq 2\mu$
B
$g(d) = 1 - \left(\frac{d}{3\mu}\right)^2 + 2\left(\frac{d}{3\mu} \log \frac{d}{3\mu}\right)$,
$d \leq 3\mu$
C
$g(d) = \exp\left(-\frac{d}{\mu}\right) - \frac{d}{\mu} \mathrm{Ei}\left(\frac{d}{\mu}\right)$,
D
$g(d) = 1 - 2\left(\frac{2d}{3\mu}\right) + \left(\frac{2d}{3\mu}\right)^2$,
$d \leq \frac{3}{2}\mu$
E
$g(d) = 1 - 3\frac{d}{2\mu} + 3\left(\frac{d}{2\mu}\right)^2 - \left(\frac{d}{2\mu}\right)^3$,
$d \leq 2\mu$
F
$g(d) = \exp\left(-2\frac{d}{\mu}\right)$
$\rho(s)$
μ
s
1.0
.8
.6
.4
.2
.0
.5
1.0
1.5
2.0
Generalization, $g(d)$ (probability measure)
Normalized distance, d ($\mu = 1$)

Property Induction

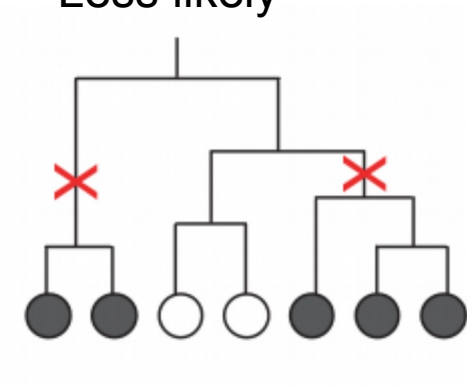
(Kemp & Tenenbaum 2009)



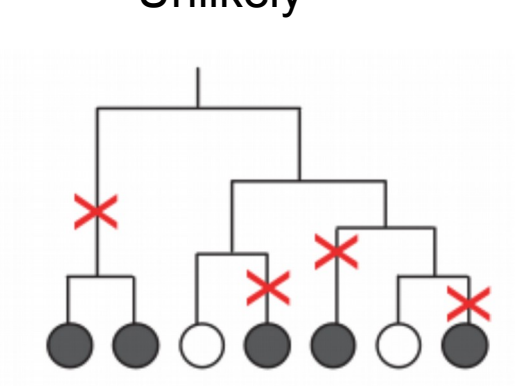
Likely



Less likely

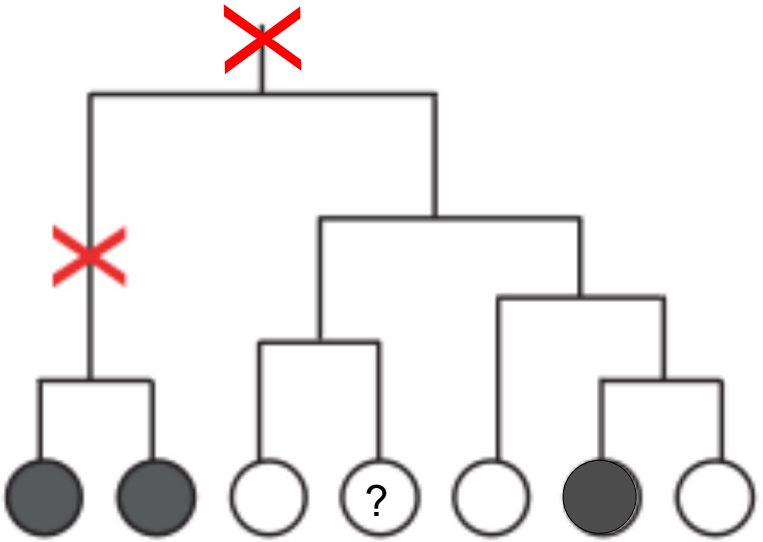


Unlikely



Property Induction
(Kemp & Tenenbaum 2009)

- **Broad observations give evidence for high Xes, because many Xes are unlikely**
- **But high Xes mean that many animals will have it.**



Property induction

- **Critically, you need a structure!**
- **And the story is more complex:** what structure is relevant may depend on what the property is!