Lecture 2: Noisy Cognition and Response Biases

Michael Woodford

Columbia University

A Short Course on Cognitive Imprecision
Northwestern University
November 12, 2025

Inference from Noisy Representations

- We have discussed the idea that randomness in observed choices may reflect application of an **optimal** decision criterion, but taking as input a **noisy** representation of the decision problem

Inference from Noisy Representations

- We have discussed the idea that randomness in observed choices may reflect application of an **optimal** decision criterion, but taking as input a **noisy** representation of the decision problem

- Optimal inference from noisy representations will generally imply estimated values that are not even **on average** equal to a **correct** valuation (based on the objective characteristics, rather than their noisy representation)

  — hence cognitive noise of this kind can be a source of systematic **bias** in choices

Inference from Noisy Representations

- We have discussed the idea that randomness in observed choices may reflect application of an **optimal** decision criterion, but taking as input a **noisy** representation of the decision problem

- Optimal inference from noisy representations will generally imply estimated values that are not even **on average** equal to a **correct** valuation (based on the objective characteristics, rather than their noisy representation)

  — hence cognitive noise of this kind can be a source of systematic **bias** in choices

- Biases of this kind are clearly seen in **perceptual** domains, where it is clear what an objectively correct judgment would be

Regression Bias

- In the case of estimates of physical **magnitudes** — judgments about the size of something that may be larger or smaller — it is commonly found that on average "people tend to **underestimate** high values ... and to **overestimate** low ones" (**"conservatism":** Hilbert, 2012; Petzscner *et al.*, 2015, call this **"regression bias"**)

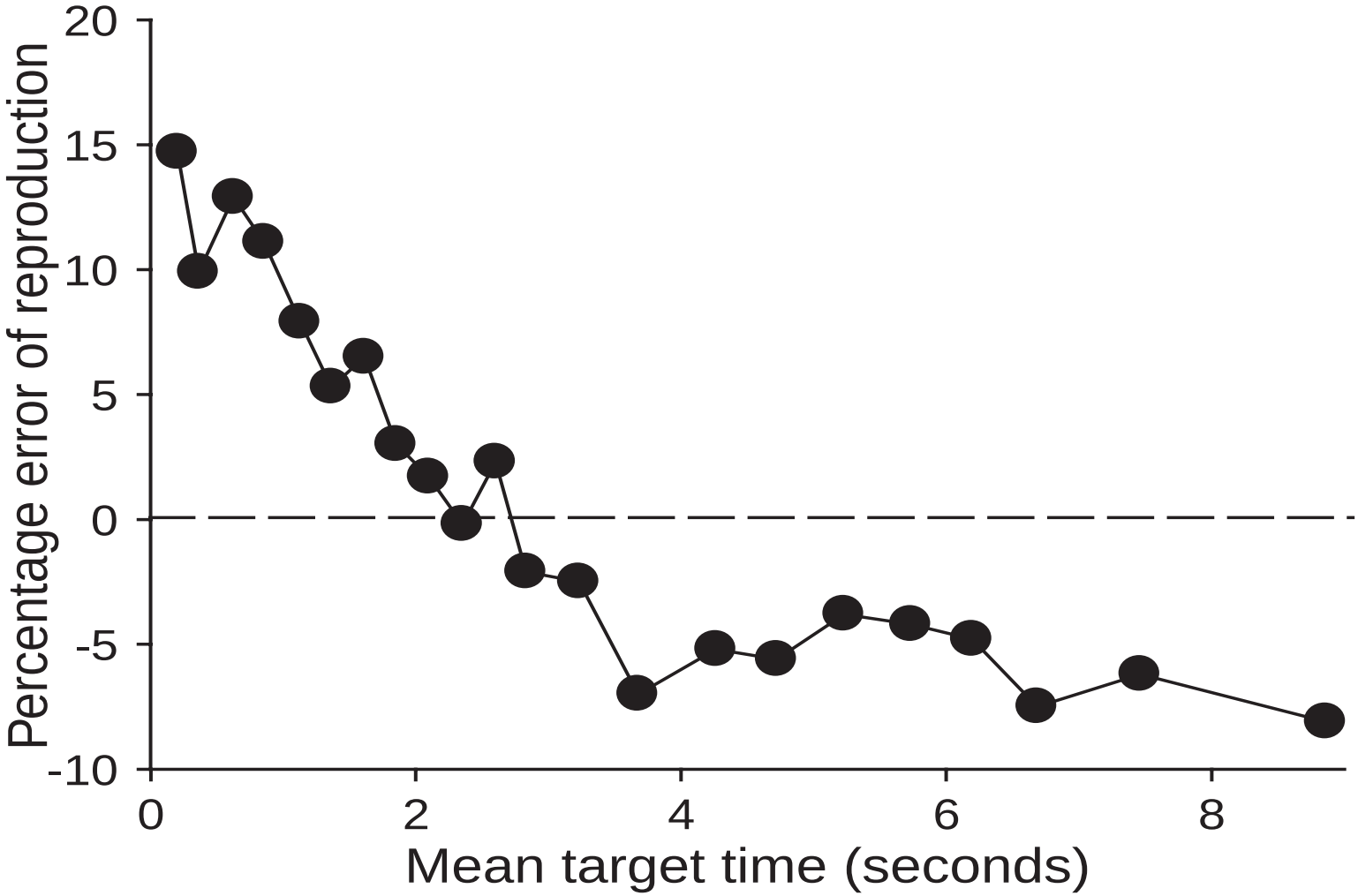
Regression Bias

- In the case of estimates of physical **magnitudes** — judgments about the size of something that may be larger or smaller — it is commonly found that on average "people tend to **underestimate** high values ... and to **overestimate** low ones" (**"conservatism":** Hilbert, 2012; Petzscner *et al.*, 2015, call this **"regression bias"**)

- An early example: **"Vierordt's Law":** short time intervals tend to be over-estimated, while long time intervals tend to under-estimated

  — classic experiment [Vierordt, 1868]: subject hears two successive taps, then must try to reproduce the same time interval themselves

Vierordt's Law



data from Vierordt (1868)
[figure from Lejeune and Wearden (2009)]

Context-Dependent Bias

- However, the distribution of estimates associated with a given stimulus doesn't depend **only** on the objective properties of that individual stimulus

  — it also depends on the **context** in which the stimulus appears

Context-Dependent Bias

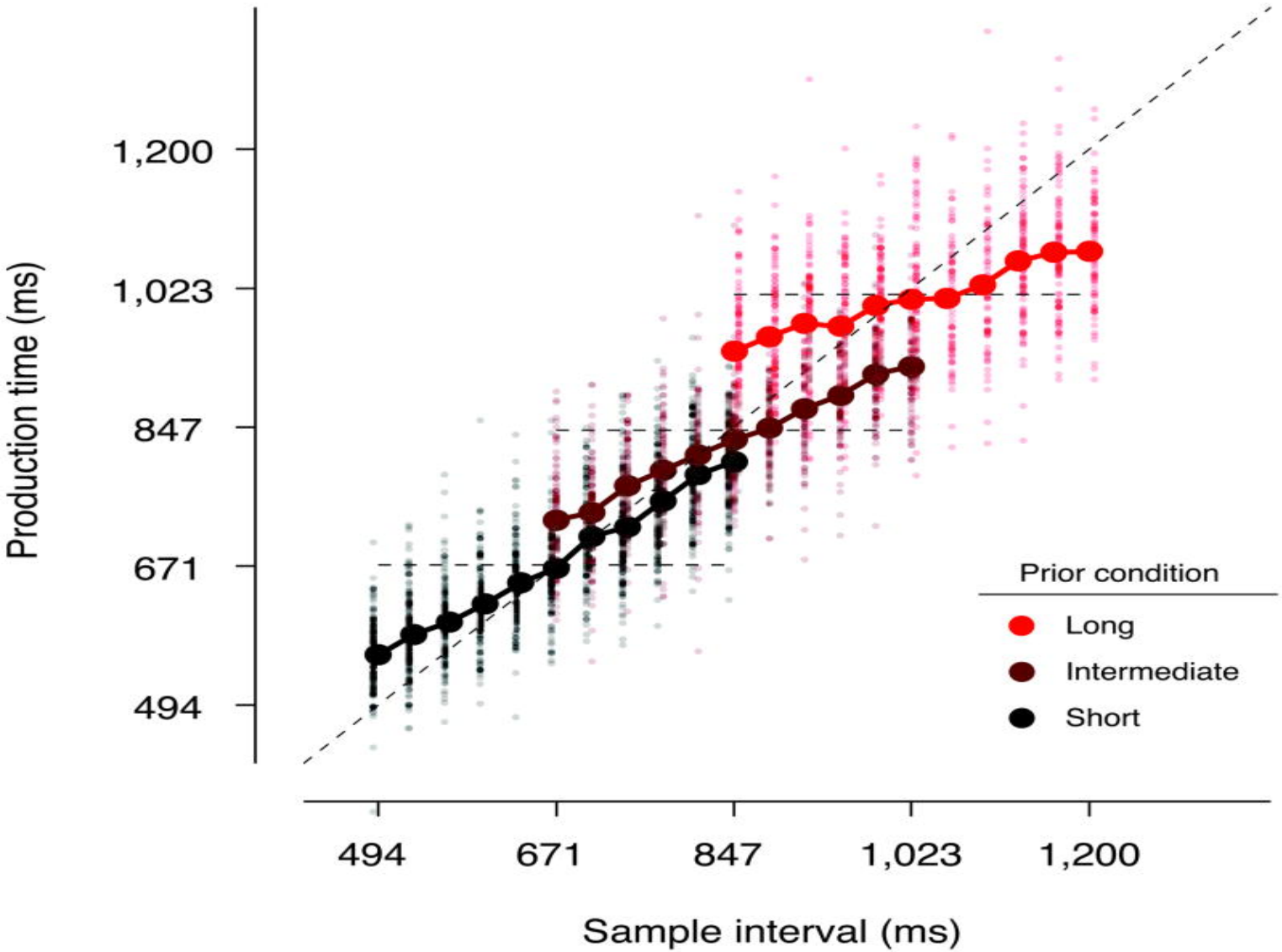
- However, the distribution of estimates associated with a given stimulus doesn't depend **only** on the objective properties of that individual stimulus

  — it also depends on the **context** in which the stimulus appears

- Jazayeri and Shadlen (2010): plot range of production intervals associated with a given presented interval, on days on which the **mean interval previously presented** is different

  — the three different "prior conditions" shown in their figure [prior **mean** is shown by dashed horizontal line for each condition]

Estimated Time Intervals (Jazayeri and Shadlen, 2010)



Central Tendency of Judgment

- Petzschner *et al.* (2015) note that this kind of pattern is observed in a large variety of different cases of **estimation of an extensive magnitude:**

  — estimation of distances, estimation of angles, estimation of sizes of objects, etc.

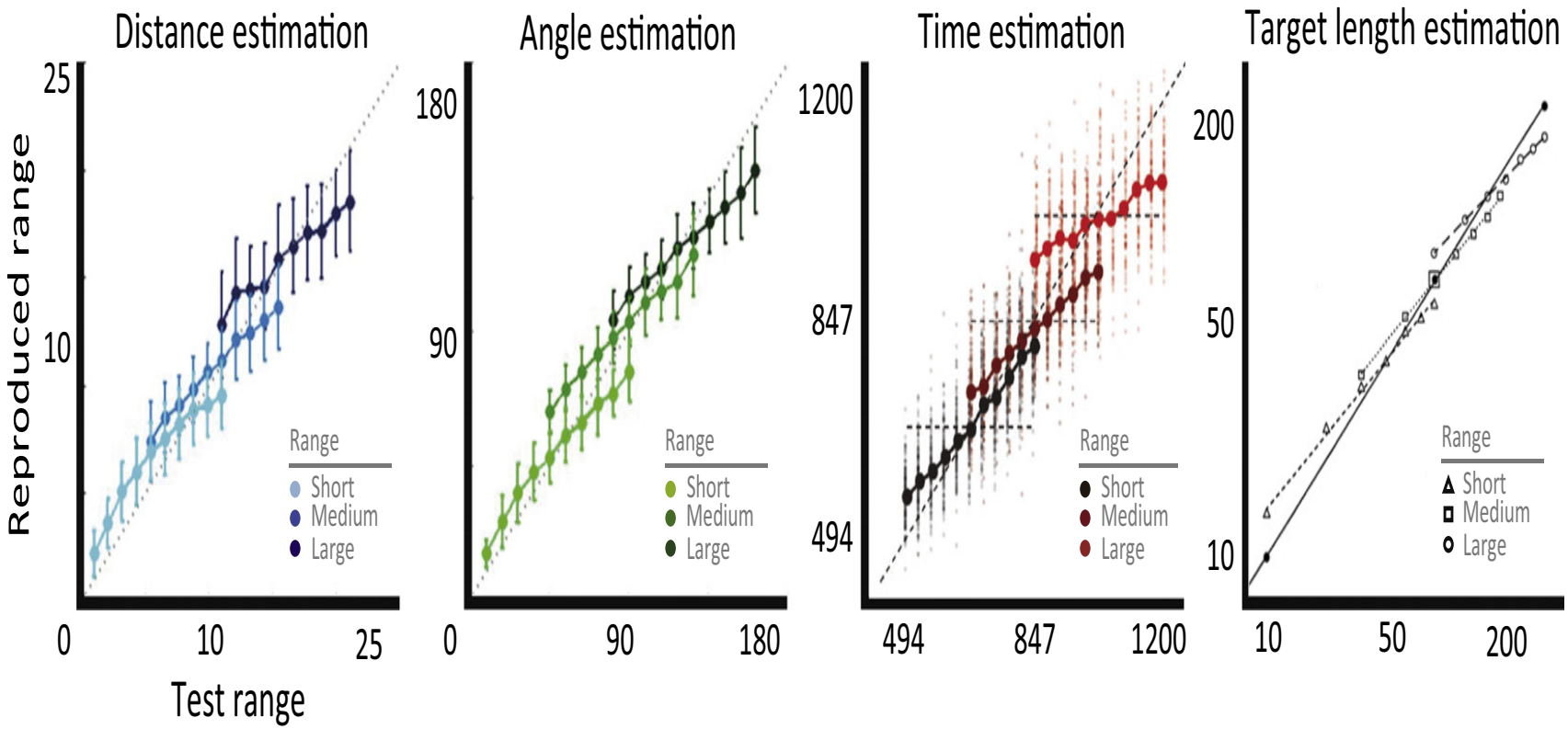
Central Tendency of Judgment

- Petzschner *et al.* (2015) note that this kind of pattern is observed in a large variety of different cases of **estimation of an extensive magnitude:**

  — estimation of distances, estimation of angles, estimation of sizes of objects, etc.

- Hollingworth (1910) calls this the **"central tendency of judgment"**

Central Tendency of Judgment



[Petzschner *et al.* (2015), Figure 1(A)]

A Bayesian Explanation

- The pervasiveness of this kind of bias (and the "central tendency" in particular) can be explained as a consequence of **optimal adaptation** to the presence of **noise** in the internal evidence ("sensory evidence") upon which the subject's estimates are based

A Bayesian Explanation

- The pervasiveness of this kind of bias (and the "central tendency" in particular) can be explained as a consequence of **optimal adaptation** to the presence of **noise** in the internal evidence ("sensory evidence") upon which the subject's estimates are based

- A simple model: suppose that objective magnitude X has a noisy internal representation r, an independent draw [each time a stimulus is presented] from distribution

$$r \sim N(X, \nu^2)$$

where [for now!] the noise variance ν^2 is independent of X

A Bayesian Explanation

- The pervasiveness of this kind of bias (and the "central tendency" in particular) can be explained as a consequence of **optimal adaptation** to the presence of **noise** in the internal evidence ("sensory evidence") upon which the subject's estimates are based

- A simple model: suppose that objective magnitude X has a noisy internal representation r, an independent draw [each time a stimulus is presented] from distribution

$$r \sim N(X, \nu^2)$$

where [for now!] the noise variance ν^2 is independent of X

- What inference about the magnitude X can be drawn from access to the noisy representation r?

A Bayesian Explanation

- Suppose further that in a given environment, the values of X that are encountered are themselves independent draws [each time a stimulus is presented] from a prior distribution that is also Gaussian

$$X \sim N(\mu, \sigma^2)$$

A Bayesian Explanation

- Suppose further that in a given environment, the values of X that are encountered are themselves independent draws [each time a stimulus is presented] from a prior distribution that is also Gaussian

$$X \sim N(\mu, \sigma^2)$$

- Then in this environment, the joint distribution of (X, r) will be **bivariate Gaussian,** and the **posterior** distribution for X given an observation r is of the form

$$X|r \sim N(\hat{\mu}(r), \hat{\sigma}^2)$$

where $$\hat{\mu}(r) \equiv \mu + \beta(r - \mu), \qquad \beta \equiv \frac{\sigma^2}{\sigma^2 + \nu^2} < 1$$

$$\hat{\sigma}^2 \equiv \frac{\sigma^2 \nu^2}{\sigma^2 + \nu^2} = (1 - \beta)\sigma^2$$

A Bayesian Explanation

- What would an **optimal response rule** for subject be?

  — perceptual experiments of this kind typically not incentivized, so subjects' objectives unclear; but if for example response rule is adapted to **minimize MSE,** response will be

$$\hat{X}(r) \;=\; \mathrm{E}[X\,|r] \;=\; \mu + \beta(r-\mu)$$

A Bayesian Explanation

- What would an **optimal response rule** for subject be?

  — perceptual experiments of this kind typically not incentivized, so subjects' objectives unclear; but if for example response rule is adapted to **minimize MSE,** response will be

$$\hat{X}(r) \;=\; \mathrm{E}[X\,|r] \;=\; \mu + \beta(r-\mu)$$

- This predicts **random** responses, conditional on objective stimulus:

$$\mathrm{var}[\hat{X}\,|X] \;=\; \beta^2 \mathrm{var}[r\,|X] \;=\; \beta^2\nu^2$$

  — and **mean** response will be **biased:**

$$\mathrm{E}[\hat{X}\,|X] \;=\; (1-\beta)\mu + \beta\mathrm{E}[r\,|X] \;=\; (1-\beta)\mu + \beta X$$

Cognitive Noise and Estimation Bias

$$\mathrm{E}[\hat{X}\,|X] \;=\; \mu + \beta\,(X-\mu)$$

- Since $\beta < 1$, this implies **regression bias:**

$$\mathrm{E}[\hat{X}\,|X] \;>\; X \quad \text{for all } \textbf{small} \text{ enough } X \quad (X < \mu)$$
$$\mathrm{E}[\hat{X}\,|X] \;<\; X \quad \text{for all } \textbf{large} \text{ enough } X \quad (X > \mu)$$

Cognitive Noise and Estimation Bias

$$\mathrm{E}[\hat{X}\,|X] \;=\; \mu + \beta\,(X - \mu)$$

- Since $\beta < 1$, this implies **regression bias:**

$$\begin{aligned} \mathrm{E}[\hat{X}\,|X] &> X \quad \text{for all } \textbf{small} \text{ enough } X \quad (X < \mu) \\ \mathrm{E}[\hat{X}\,|X] &< X \quad \text{for all } \textbf{large} \text{ enough } X \quad (X > \mu) \end{aligned}$$

- And if the **ranges of stimuli** encountered in different contexts differ, an optimally adapted decision rule for each context should result in a **different mapping** $\mathrm{E}[\hat{X}\,|X]$

  — the "crossover point" should always be where $X = \mu$, the **prior mean** for that context, in accordance with Hollingsworth's **"central tendency of judgment"**

Cognitive Noise and Estimation Bias

- Fixing the statistics of the environment (μ, σ^2), the model also makes predictions about how bias should change if ν^2 is higher in some contexts than others:

$$\text{larger } \nu^2 \quad \Rightarrow \quad \text{smaller } \beta \quad \Rightarrow \quad \text{stronger regression bias}$$

Cognitive Noise and Estimation Bias

- Fixing the statistics of the environment (μ, σ^2), the model also makes predictions about how bias should change if ν^2 is higher in some contexts than others:

$$\text{larger } \nu^2 \quad \Rightarrow \quad \text{smaller } \beta \quad \Rightarrow \quad \text{stronger regression bias}$$

- And indeed regression bias is **stronger** in contexts where sensory evidence is **noisier:**
  - bias in estimated **speed** of a moving visual image greater when image is presented with **lower visual contrast** (Stocker and Simoncelli, 2006)

Cognitive Noise and Estimation Bias

- Another example:
  - when subjects must compare the sizes of two stimuli presented **sequentially,** they **over-estimate** the relative size of the first stimulus when both are relatively small, but **under-estimate** the size of the first stimulus when both are relatively large

    — consistent with Bayesian model of regression bias if information **held longer in working memory** is retrieved with **greater noise** (Ashourian and Loewenstein, 2011)

Cognitive Noise and Estimation Bias

- Another example:
  - when subjects must compare the sizes of two stimuli presented **sequentially,** they **over-estimate** the relative size of the first stimulus when both are relatively small, but **under-estimate** the size of the first stimulus when both are relatively large

    — consistent with Bayesian model of regression bias if information **held longer in working memory** is retrieved with **greater noise** (Ashourian and Loewenstein, 2011)

- Additional kind of **causal manipulation** to increase cognitive noise: central-tendency effect increased by increasing **"cognitive load"** (Allred *et al.*, 2016)

Imprecision in Number Representation

- But do such biases in estimation of **physical magnitudes** on the basis of **sensory evidence** matter for economic decisions?

- Recall from Lecture 1: there is reason to believe that **numerical information** is also represented **imprecisely** in the brain

Imprecision in Number Representation

- But do such biases in estimation of **physical magnitudes** on the basis of **sensory evidence** matter for economic decisions?
- Recall from Lecture 1: there is reason to believe that **numerical information** is also represented **imprecisely** in the brain
- Much of what we know about imprecision in number representation comes from studies of the accuracy of judgments about the **numerosity** of visual arrays

Imprecision in Number Representation

- For example, the precision of **estimates** of numerosity has been studied experimentally, at least since Jevons (1871)

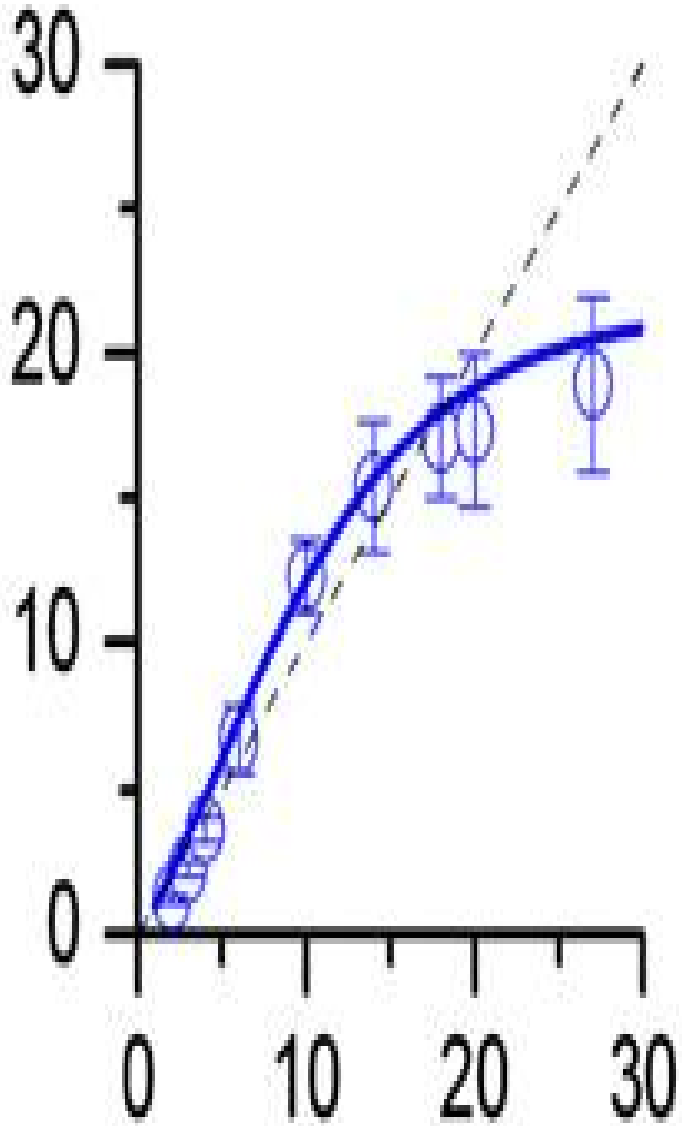
Imprecision in Number Representation

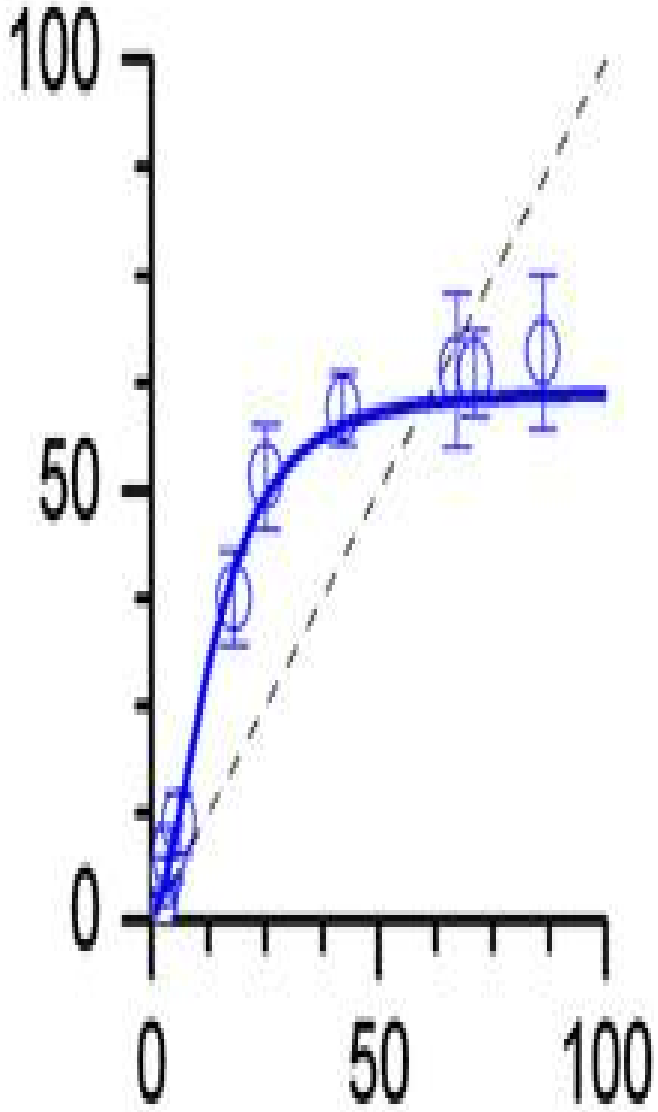
- For example, the precision of **estimates** of numerosity has been studied experimentally, at least since Jevons (1871)

- These estimates also commonly exhibit **regression bias:** over-estimation of small numbers, under-estimation of larger ones

  — as well as the **central tendency of judgment**

Numerosity Estimation [from Anobile *et al.* (2012)]





Imprecision in Number Representation

- The **degree** of regression bias also seems to be higher when internal representation of numerosity is **noisier**

Imprecision in Number Representation

- The **degree** of regression bias also seems to be higher when internal representation of numerosity is **noisier**
- Xiang *et al.* (2021) show this in two ways:
  - sort experimental trials according to degree of **uncertainty** about the number expressed by the subject
  - seek to **increase** cognitive noise by allowing **less time**

Imprecision in Number Representation

- The **degree** of regression bias also seems to be higher when internal representation of numerosity is **noisier**
- Xiang *et al.* (2021) show this in two ways:
  - sort experimental trials according to degree of **uncertainty** about the number expressed by the subject
  - seek to **increase** cognitive noise by allowing **less time**
- Both greater reported uncertainty (for whatever reason) and less time are associated with
  - more **variable** responses
  - stronger **regression bias**

"Behavioral Attenuation"

- Similar biases are also observed in more complex decisions, that seem to require **reasoning**

"Behavioral Attenuation"

- Similar biases are also observed in more complex decisions, that seem to require **reasoning**
- Enke *et al.* (2025) document the ubiquity of a pattern that they call **behavioral attenuation** across a wide range of decision problems:
  - **insufficient responsiveness** of DM's decision to variation in parameters of the problem, relative to an optimal decision
    — as with "regression bias" in perceptual domains
  - moreover, the elasticity of decisions w.r.t. parameter variation is **decreased** when there is greater **uncertainty** about best decision
    — as in the Xiang *et al.* study of numerosity estimation

"Behavioral Attenuation"

- Found both in
  - pure **calculation** tasks (statistical inference, optimization), where there is an objectively correct answer (as with the perceptual tasks above), and in
  - **preferential choice** tasks (effort supply, saving, etc.), where the experimenter doesn't know what subjects' preferences are
- In latter case, can nonetheless observe the signature of this effect when responses are **less elastic** in the case of greater uncertainty

"Recall" Task (Stock Valuation)

- Example of a calculation task ["Recall" task]: subjects must estimate the **value** (in dollars) of a particular (fictitious) stock, on the basis of how many positive or negative news items there have been about the company
  - all news items are simply "positive" or "negative" (only sign matters)
  - the signs of all the news items are displayed on the screen
  - the subject is told the **formula** for converting news into implied dollar value — so task is just a calculation using information on the screen

"Recall" Task: Decision Screen

Reminder: **Stock price (in $) = 100 + number of positive news – number of negative news**

In this round:

Company name: TigerThrive Fitness

Positive news: 

Negative news: 

What do you think is the stock price of this company? $ 10

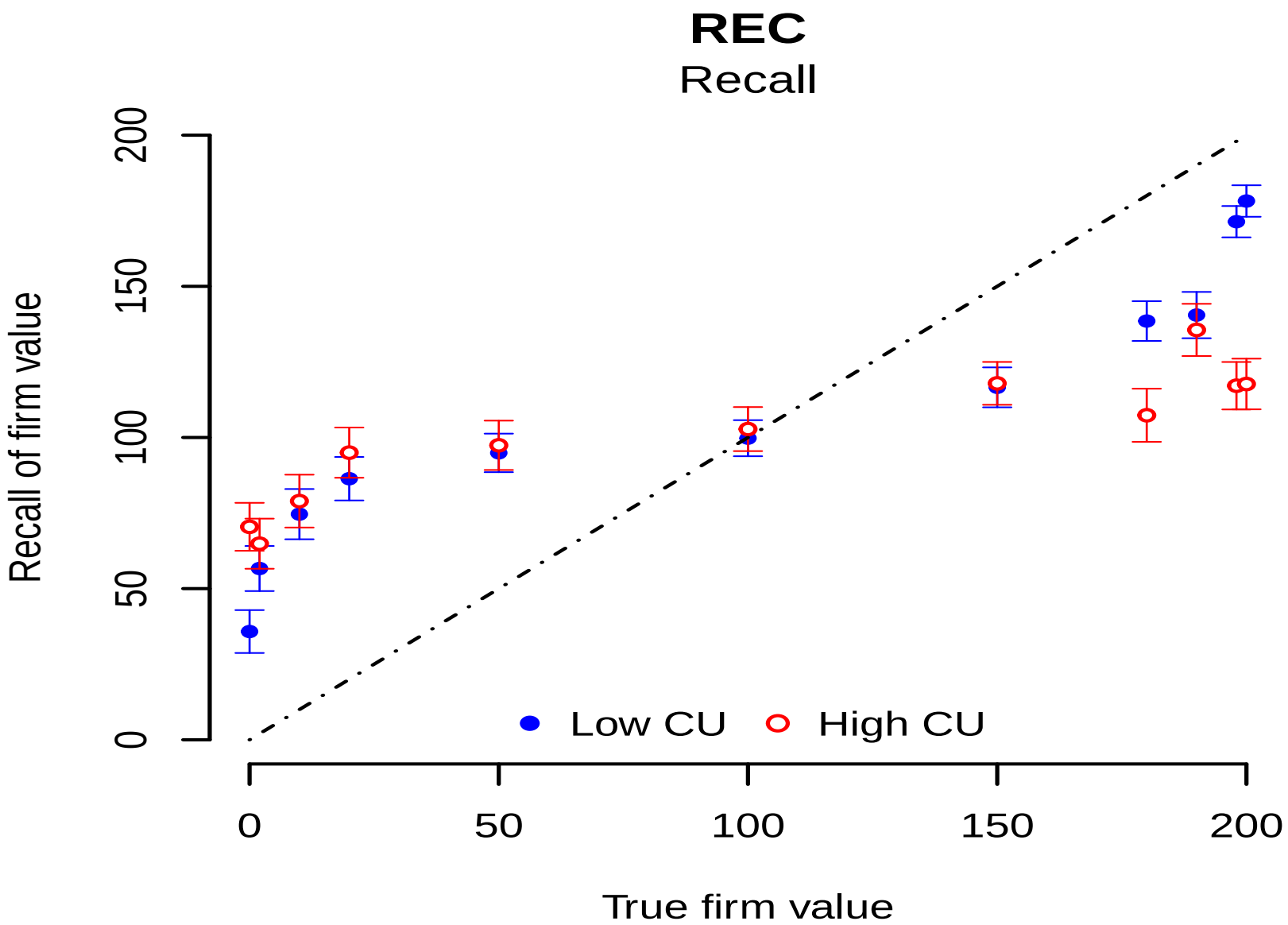
How certain are you that the stock price is actually somewhere between **$9** and **$11**?

Very uncertain — **Completely certain**

0% 25% 50% 75% 100%

79 %

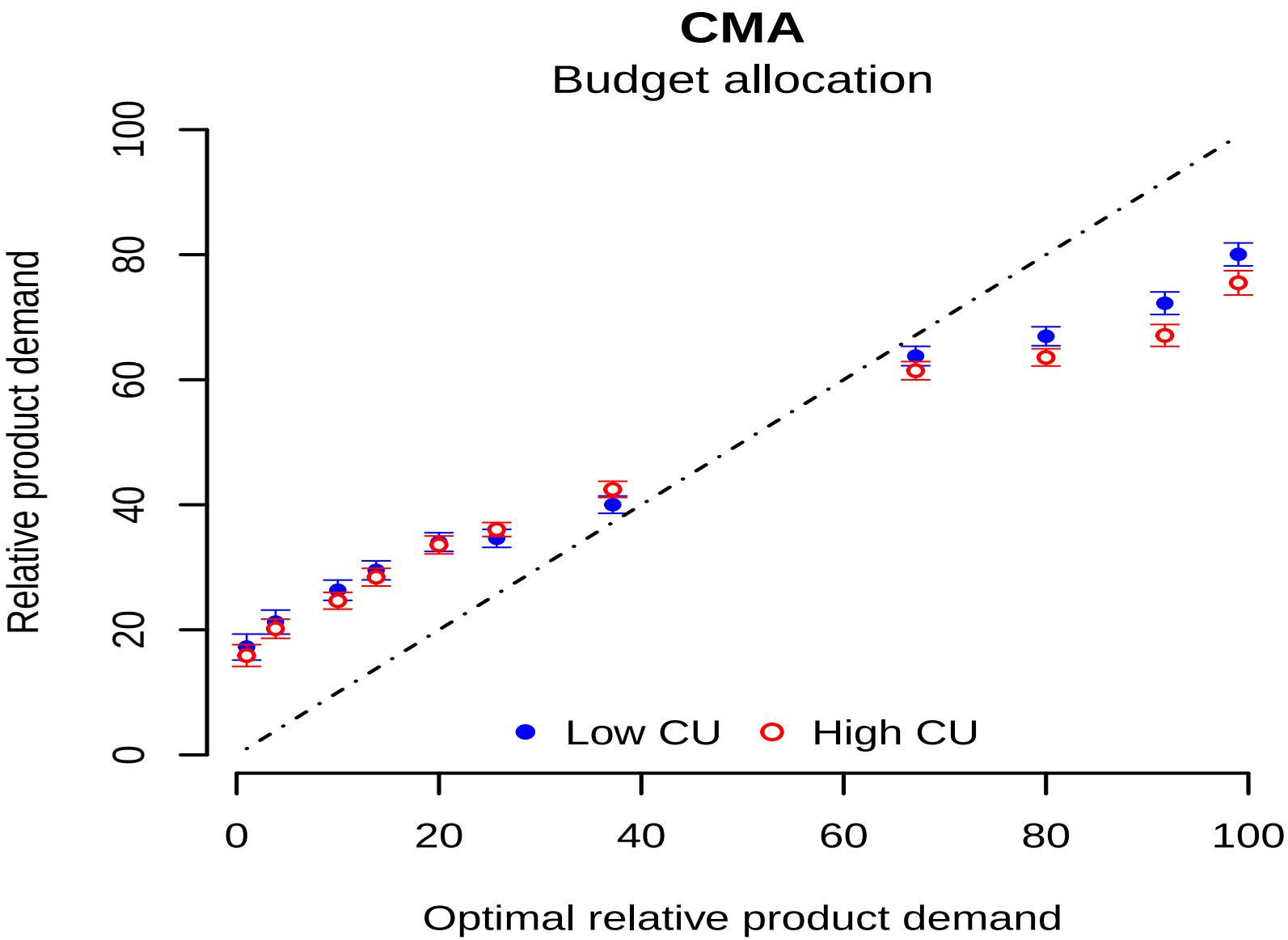
"Recall" Task: Behavioral Attenuation



"Allocation" Task (Consumer Demand)

- Another seemingly "economic" task that is actually a pure computation: subjects must decide how to allocate their budget between purchases of two goods, given their prices
  - subject is **told** the utility function (both a formula and a graph)
  - their monetary reward is proportional to the "utility" obtained — they don't actually consume the items that they choose to "buy"

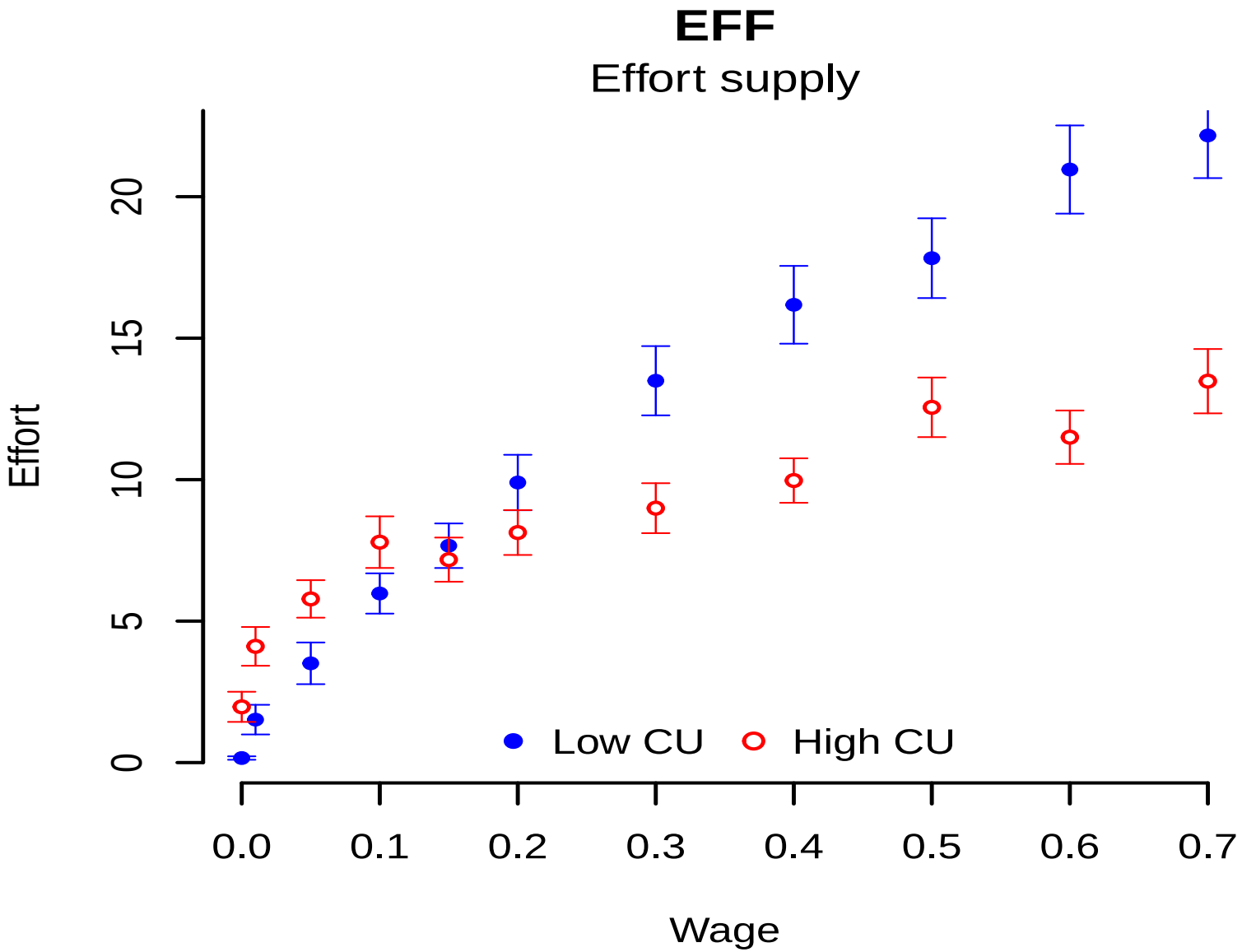
"Allocation" Task: Behavioral Attenuation



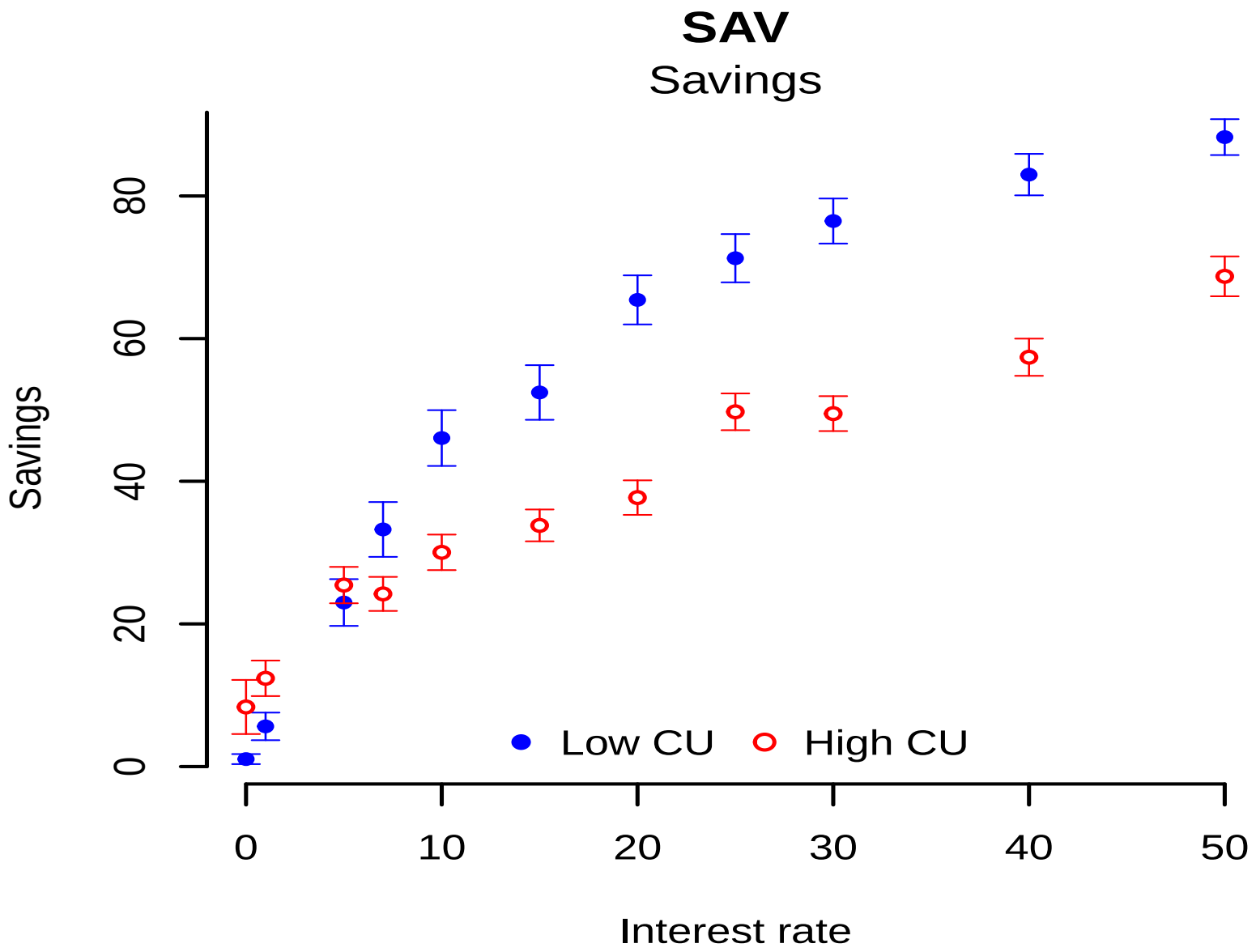
Preferential Tasks

- Other tasks involve choices where the subject's reward will depend on their **own preferences:**
  - "Effort supply" task: subject decides how many tasks to complete, given the wage (piece rate) offered
  - "Saving" task: subject decides how much of an endowment to "save," given the interest factor by which money will be multiplied if payment is taken later

"Effort Supply" Task: Behavioral Attenuation



"Saving" Task: Behavioral Attenuation



Inhomogeneity in Cognitive Precision

- Enke *et al.* show that variations in subjective uncertainty co-vary with the degree of behavioral attenuation, not just **across subjects** or for other reasons orthogonal to the nature of the decision problem (e.g., variation across trials in the degree of distraction or fatigue)

  — but also as functions of the **parameter values** defining the particular decision problem (for a given type of problem)

  — different ranges of parameter values result in different degrees of **subjective uncertainty,** and this is associated with correspondingly different **elasticities of behavioral response** to parameter variation when the parameters are in different ranges

Inhomogeneity in Cognitive Precision

- Enke *et al.* further propose a general regularity about how cognitive uncertainty (and hence behavioral elasticity) varies with parameter values: CU increases the greater the distance from any **"simple points"**

  — special parameter values for which the decision simplifies

Inhomogeneity in Cognitive Precision

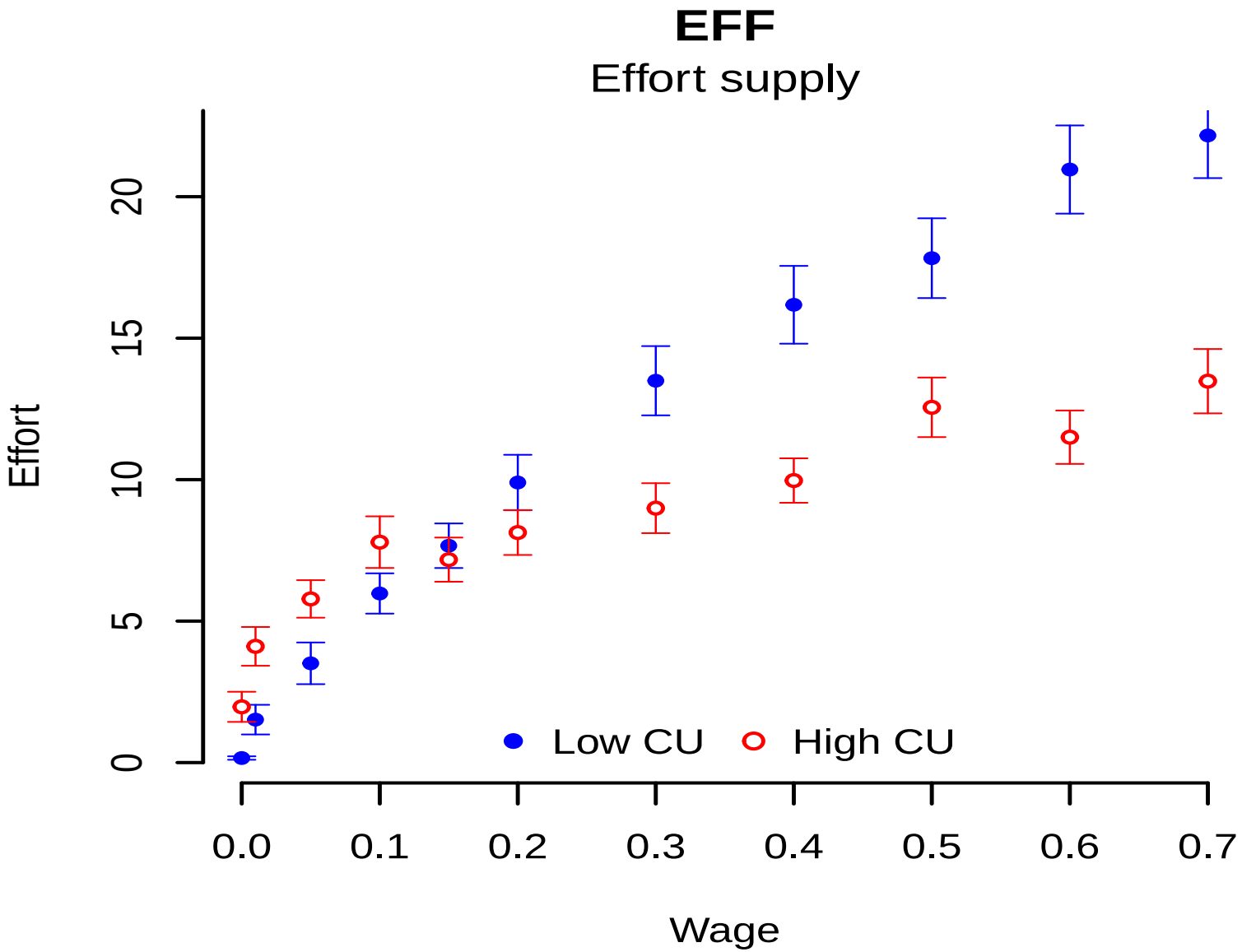
- Enke *et al.* further propose a general regularity about how cognitive uncertainty (and hence behavioral elasticity) varies with parameter values: CU increases the greater the distance from any **"simple points"**

  — special parameter values for which the decision simplifies

- In many cases, the "simple point" will be a **zero** value (e.g., a zero wage in "effort supply" task)

  — in this case, the prediction of declining behavioral elasticity as a parameter increases away from zero coincides with the familiar idea of **"diminishing marginal sensitivity"** (Kahneman and Tversky, 1979)

"Effort Supply" Task: Diminishing Sensitivity



Inhomogeneity in Cognitive Precision

- Nonlinear response distortion of this kind is what is predicted by a Bayesian model of optimal decision making in the presence of cognitive noise, if the degree of noise is inhomogeneous [unlike what was assumed in linear-Gaussian example above]

Inhomogeneity in Cognitive Precision

- Nonlinear response distortion of this kind is what is predicted by a Bayesian model of optimal decision making in the presence of cognitive noise, if the degree of noise is inhomogeneous [unlike what was assumed in linear-Gaussian example above]

- And it's familiar in the study of imprecise **perceptual** judgments that the size of the difference in a **physical** stimulus magnitude required for two stimuli to be **discriminated** with a given accuracy is **not constant** over the stimulus space

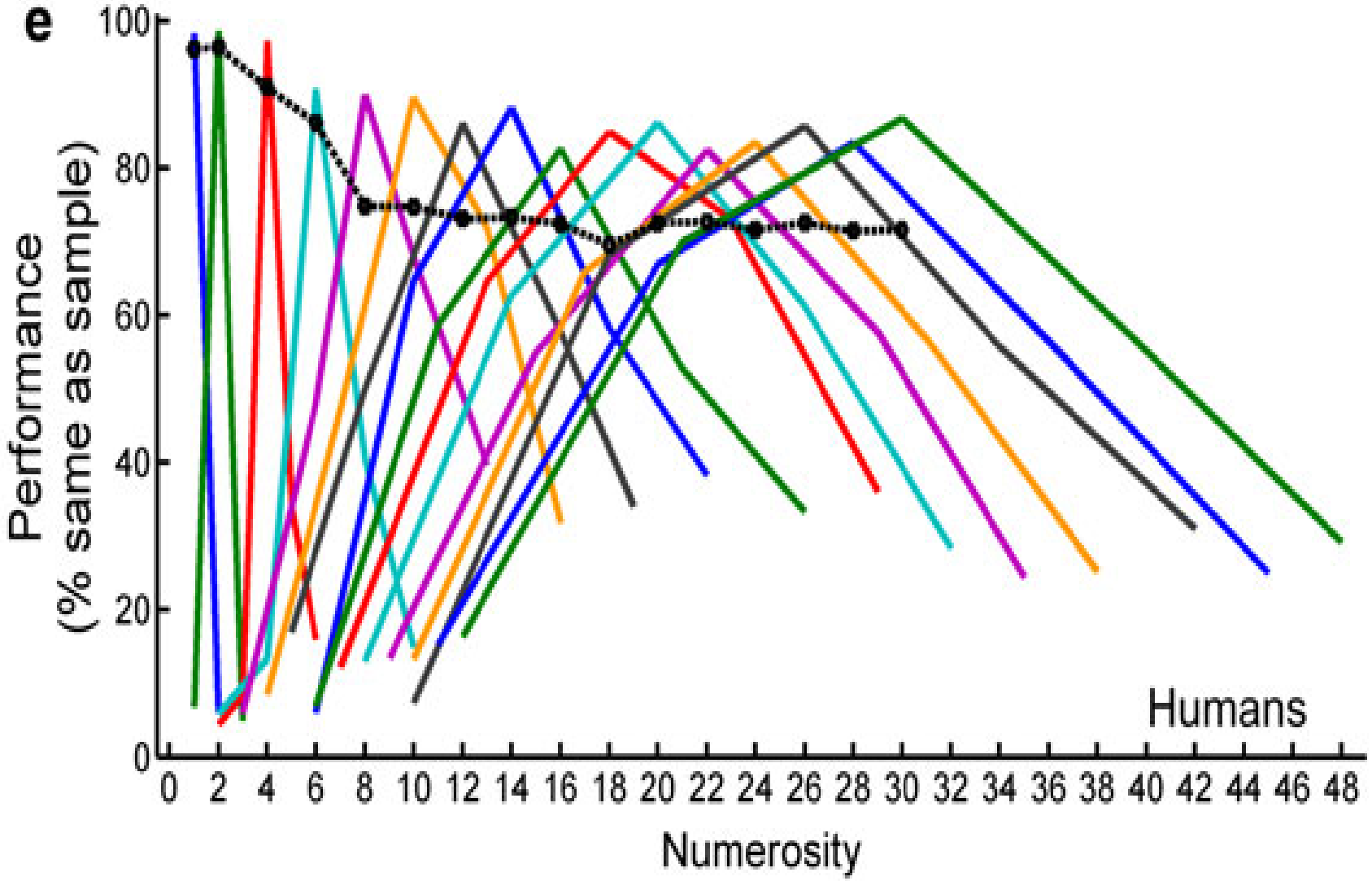
  — famously, **"Weber's Law"** asserts that the amount by which a second magnitude (length, weight, ...) must be greater for a given degree of discriminability **grows in proportion** to the first magnitude

Imprecise Number Representation

- Imprecise judgments of **numerosity** are often argued to be like this

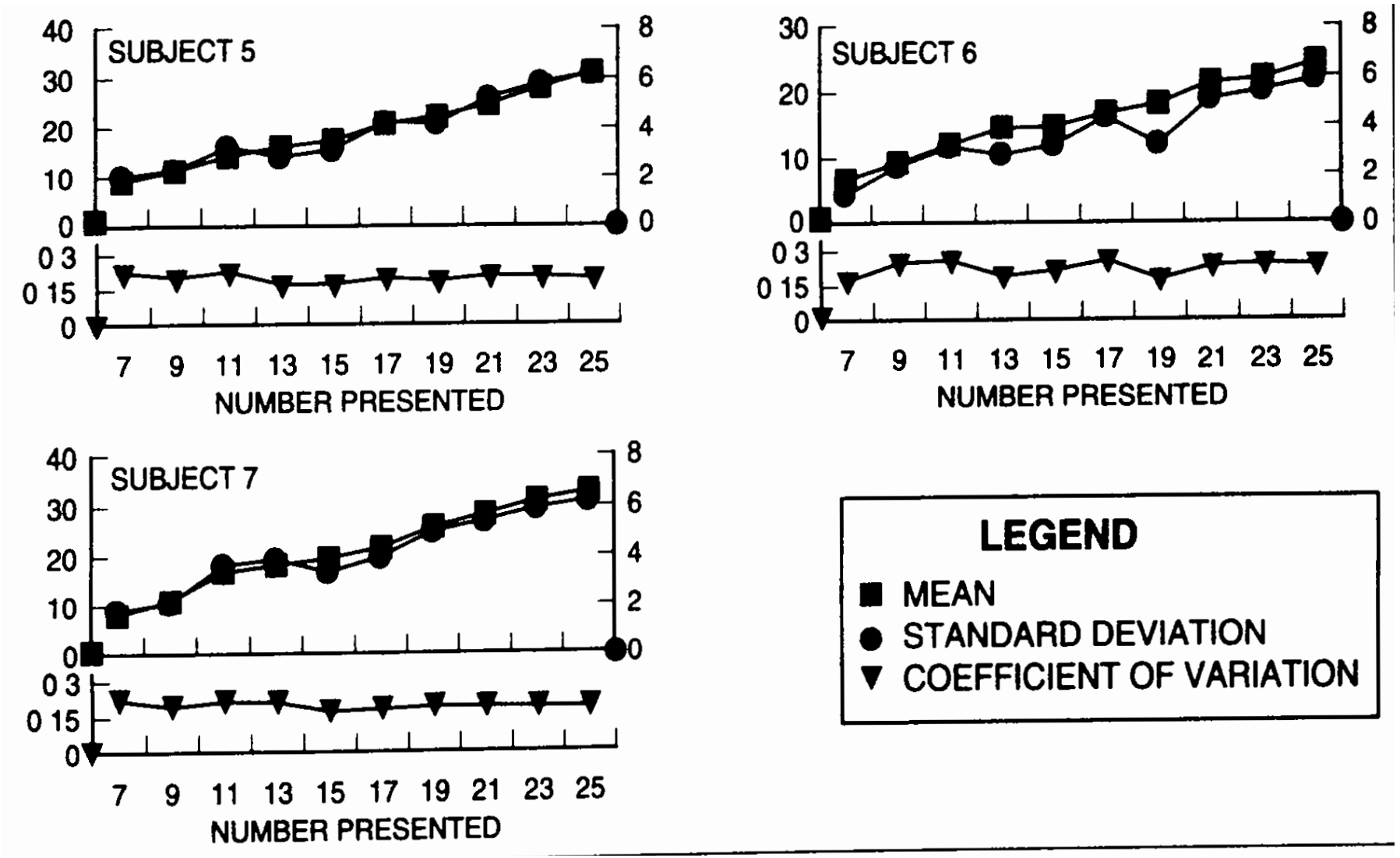
  — leading authors like Dehaene (2011) to argue that numbers are subjectively represented on a **logarithmic** "mental number line"

Numerosity Discrimination



from Nieder (2013)

Numerosity Estimation



from Whalen *et al.* (1999)

Imprecise Number Representation

- This kind of variability in the precision of number representation can be captured by the hypothesis that the size of a number $X > 0$ is a represented by a random quantity r_x drawn from a distribution

$$r_x \sim N(\log X, \nu^2)$$

— then the degree of **overlap** between the distributions of internal representations of two numbers X_1, X_2 will be a decreasing function of $|\log(X_2/X_1)|$

Imprecise Number Representation

- This kind of variability in the precision of number representation can be captured by the hypothesis that the size of a number $X > 0$ is a represented by a random quantity r_x drawn from a distribution

$$r_x \sim N(\log X, \nu^2)$$

— then the degree of **overlap** between the distributions of internal representations of two numbers X_1, X_2 will be a decreasing function of $|\log(X_2/X_1)|$

- But this kind of inhomogeneity in the precision of number representation also has implications for the kind of **bias** should be observed in numerosity **estimates,** if the response rule is optimally adapted to the nature of the noisy representation

Imprecise Number Representation

- Suppose that the **prior** distribution over numerosities to which the response rule is adapted is **log-normal:**

$$\log X \sim N(\mu, \omega^2)$$

Imprecise Number Representation

- Suppose that the **prior** distribution over numerosities to which the response rule is adapted is **log-normal:**

$$\log X \sim N(\mu, \omega^2)$$

- Then the **posterior** distribution for X, conditional on a given representation r_x, will also be **log-normal,**

$$\log X \,|r_x \sim N(\hat{\mu}(r_x), \hat{\omega}^2)$$

and the posterior mean (minimum-MSE estimate of X) $\hat{X}$ will be given by

$$\hat{X} = \exp(\hat{\mu}(r_x) + \frac{1}{2}\hat{\omega}^2)$$

Imprecise Number Representation

- This provides a model of the distribution of estimates $\hat{X}$ associated with a given true numerosity X, that should be **log-normal** with conditional moments

$$m(X) \equiv \mathrm{E}[\hat{X}\,|X] = AX^{\beta}, \qquad \mathrm{var}[\hat{X}\,|X] = B\mathrm{E}[\hat{X}\,|X]^2$$

where $A, B > 0$ are constants, and

$$\beta = \frac{\sigma^2}{\sigma^2 + \nu^2} < 1$$

[see slides at end for details]

Imprecise Number Representation

- This provides a model of the distribution of estimates $\hat{X}$ associated with a given true numerosity X, that should be **log-normal** with conditional moments

$$m(X) \equiv \mathrm{E}[\hat{X}\,|X] = AX^{\beta}, \qquad \mathrm{var}[\hat{X}\,|X] = B\mathrm{E}[\hat{X}\,|X]^2$$

where $A, B > 0$ are constants, and

$$\beta = \frac{\sigma^2}{\sigma^2 + \nu^2} < 1$$

[see slides at end for details]

- The model implies that $\partial m(X)/\partial X$ will be **decreasing** the farther one gets from the "simple point" $X = 0$, as is indeed observed in experimental data for numerosity estimation

Implications for Behavioral Biases

- This model of imprecise assessment of the size of a numerical quantity provides an explanation for **small-stakes risk aversion** (Khaw *et al.*, 2021)

Implications for Behavioral Biases

- This model of imprecise assessment of the size of a numerical quantity provides an explanation for **small-stakes risk aversion** (Khaw *et al.*, 2021)
  - in Lecture 1, we proposed a model in which we would obtain an "indifference point" indicating risk-aversion if and only if the noisy representations of monetary payoffs X and C implied that

$$\mathrm{E}[X \,|\boldsymbol{r}] \;>\; 2 \cdot \mathrm{E}[C \,|\boldsymbol{r}],$$

    with probability **less than 1/2,** even for values of X/C modestly **greater than 2**

Implications for Behavioral Biases

- This model of imprecise assessment of the size of a numerical quantity provides an explanation for **small-stakes risk aversion** (Khaw *et al.*, 2021)
  - in Lecture 1, we proposed a model in which we would obtain an "indifference point" indicating risk-aversion if and only if the noisy representations of monetary payoffs X and C implied that
$$\mathrm{E}[X \,|\boldsymbol{r}] \;>\; 2 \cdot \mathrm{E}[C \,|\boldsymbol{r}],$$
with probability **less than 1/2,** even for values of X/C modestly **greater than 2**
  - we now have a model of noisy coding where that will be true: suppose that each payoff $Q_i = X, C$ has a noisy representation r_i, a conditionally independent draw from
$$r_i \;\sim\; N(\log Q_i,\, \nu^2)$$

Implications for Behavioral Biases

- Then if there is also a (common) independent log-normal prior for each of these quantities, we will have

$$\begin{aligned} \mathrm{E}[Q_i \,|\boldsymbol{r}] \;=\; \mathrm{E}[Q_i \,|\boldsymbol{r}_i] \;&=\; \exp(\alpha + \beta r_i) \\ &=\; A Q_i^{\beta} \xi_i \end{aligned}$$

  where ξ_i is an independent log-normal multiplicative error term

Implications for Behavioral Biases

- Then if there is also a (common) independent log-normal prior for each of these quantities, we will have

$$\begin{aligned} \mathrm{E}[Q_i \,|\boldsymbol{r}] \;=\; \mathrm{E}[Q_i \,|\boldsymbol{r}_i] \;&=\; \exp(\alpha + \beta r_i) \\ &=\; A Q_i^{\beta} \xi_i \end{aligned}$$

  where ξ_i is an independent log-normal multiplicative error term

- So DM chooses the risky gamble more often than not (conditional on X, C) if and only if

$$\frac{1}{2} u(X) \;>\; u(C)$$

  where we define $u(Q) = \mathrm{E}[\hat{Q} \,|Q] = A Q^{\beta}$

Implications for Behavioral Biases

$$\frac{1}{2}u(X) > u(C)$$

- The modal choice is determined **as if** the outcome of the gamble is considered in isolation from the DM's other sources of wealth ["narrow bracketing"], and payoffs are valued in accordance with a **strictly concave** vonN-M utility function [or PT "value function"]

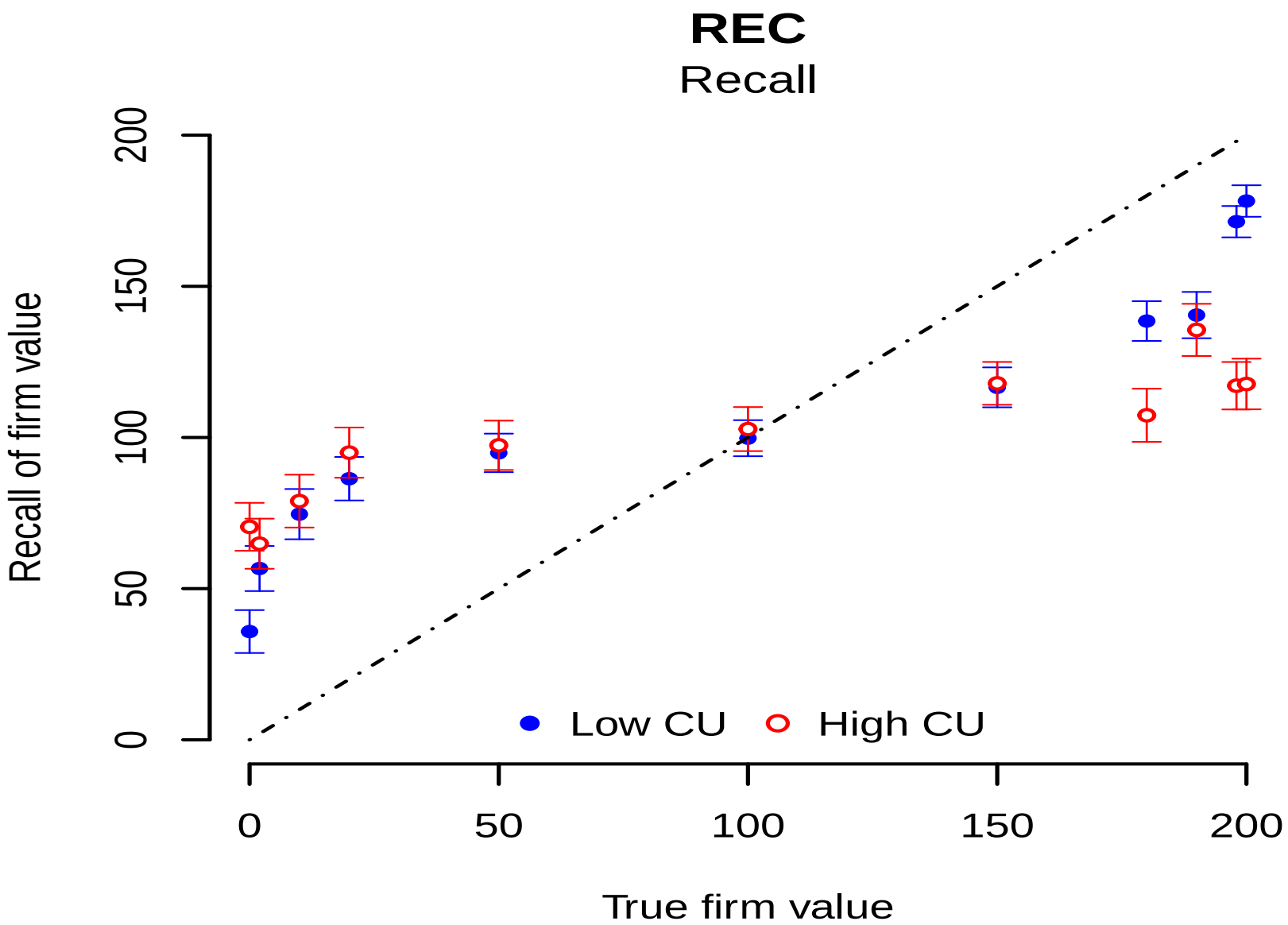
  — even though the decision rule is actually optimized to maximize the expected **total financial wealth** of the DM [no narrow bracketing, no actual risk aversion for gambles of this size]

Another Kind of Inhomogeneity in Precision

- There can also be **multiple "simple points",** as in the "recall" task of Enke *et al.*:
  - both the case of **all good news** and the case of **all bad news** are cases that can be **easily distinguished** from nearby cases

    — leading to **repulsion** from both points

"Recall" Task: Repulsion from Both Ends



Another Kind of Inhomogeneity in Precision

- There can also be **multiple "simple points",** as in the "recall" task of Enke *et al.*:
  - both the case of **all good news** and the case of **all bad news** are cases that can be **easily distinguished** from nearby cases

    — leading to **repulsion** from both points

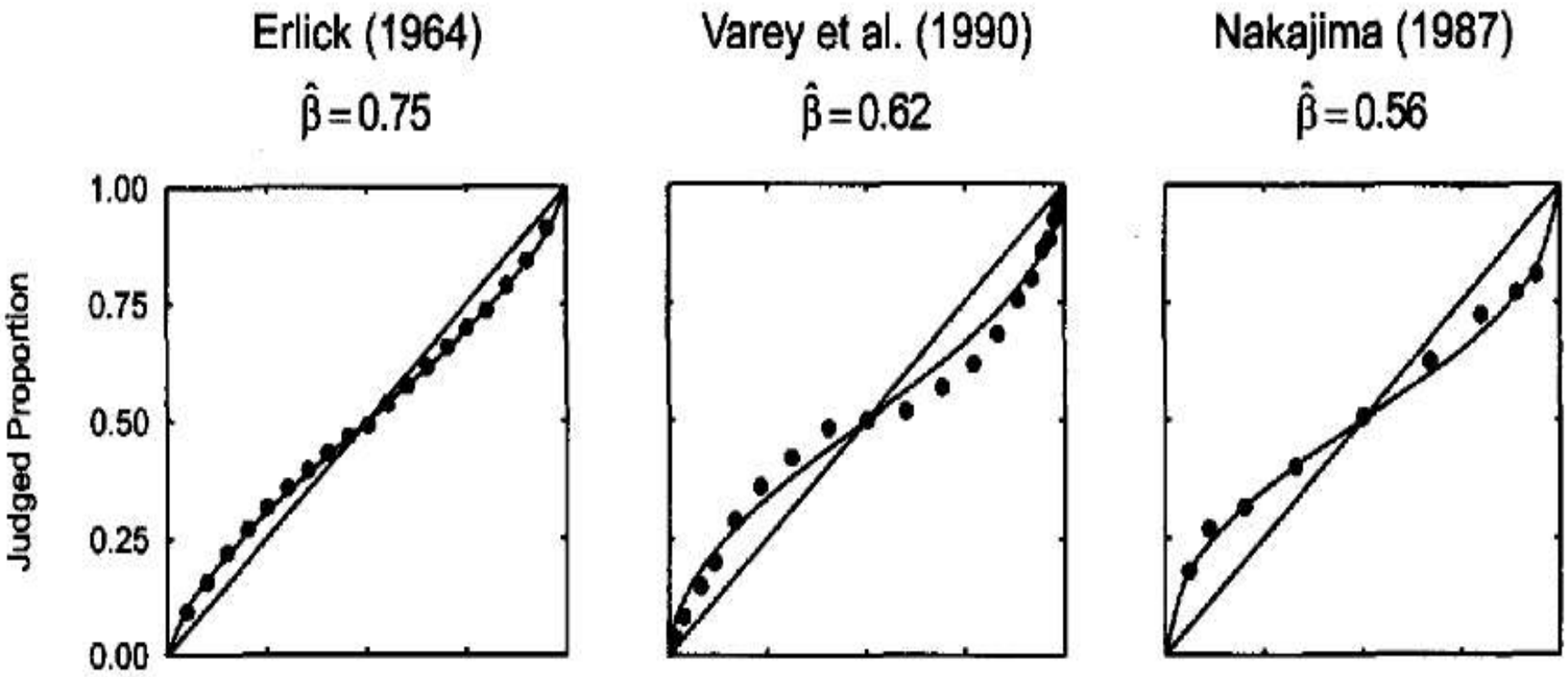
- And the special character of these points seems more clearly to relate to the precision with which the **evidence is represented,** rather than **simplicity of solution** of the arithmetic calculation

Another Kind of Inhomogeneity in Precision

- In fact, a similar phenomenon is common in purely **perceptual** estimation tasks

  — when subjects have to estimate the **relative proportion** of two extensive magnitudes: fraction of dots in an array that are of a particular color, fraction of letters in a sequence that have been *A*s, etc. [see Hollands and Dyre (2000)]

Bias in Judged Proportions



[figure from Hollands and Dyre (2000)]

horizontal axis = true proportion

note similar "inverse-S shaped" curve in each case

Another Kind of Inhomogeneity in Precision

- In fact, a similar phenomenon is common in purely **perceptual** estimation tasks

  — when subjects have to estimate the **relative proportion** of two extensive magnitudes: fraction of dots in an array that are of a particular color, fraction of letters in a sequence that have been *A*s, etc. [see Hollands and Dyre (2000)]

- This suggests that the pattern in the "recall" task is mainly about imprecision in recognition of **which fraction of the news** is good rather than bad, rather than imprecision in the **calculations** required by the formula

Another Kind of Inhomogeneity in Precision

- The "inverse-S shaped" pattern of biases can also be explained as **optimal adaptation to noise** in the internal representation of the relative proportions

  — under the hypothesis that the representation of proportions is **more precise at both extremes** [see Khaw *et al.* (2025) for details]

Another Kind of Inhomogeneity in Precision

- The "inverse-S shaped" pattern of biases can also be explained as **optimal adaptation to noise** in the internal representation of the relative proportions

  — under the hypothesis that the representation of proportions is **more precise at both extremes** [see Khaw *et al.* (2025) for details]

- This provides a possible explanation for the shape of the "probability weighting function" of prospect theory

[to discuss further in Lecture 3]

References

Allred, S.R., L.E. Crawford, S. Duffy, and J. Smith, "Working Memory and Spatial Judgments: Cognitive Load Increases the Central Tendency Bias," *Psychonomic Bull. and Review* 23: 1825-31 (2016).

Anobile, G., G.M. Cicchini, and D.C. Burr, "Linear Mapping of Numbers Onto Space Requires Attention," *Cognition* 122: 454-459 (2012).

Ashourian, P., and Y. Loewenstein, "Bayesian Inference Underlies the Contraction Bias in Delayed Comparison Tasks," *PloS One* 6, e19551 (2011).

Dehaene (2011): see references for Lecture 1

References (p.2)

Enke, B., T. Graeber, R. Oprea, and J. Yang, "Behavioral Attenuation," working paper, Harvard University, September 2025.

Hilbert, M., "Toward a Synthesis of Cognitive Biases: How Noisy Information Processing Can Bias Human Decision Making," *Psychological Bulletin* 138: 211-237 (2012).

Hollands, J.G., and B.P. Dyre, "Bias in Proportion Judgments: The Cyclical Power Model," *Psychological Review* 107: 500-524 (2000).

Hollingworth, H.L., "The Central Tendency of Judgment," *Journal of Philosophy* 7: 461-469 (1910).

References (p.3)

Jazayeri, M., and M.N. Shadlen, "Temporal Context Calibrates Interval Timing," *Nature Neuroscience* 13: 1020-1026 (2010).

Jevons, W.S., "The Power of Numerical Discrimination," *Nature* 3: 281-282 (1871).

Kahneman, D., and A. Tversky, "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica* 47: 263-291 (1979).

Khaw *et al.* (2021): see references for Lecture 1

References (p.4)

Khaw, M.W., Z. Li, and M. Woodford, "Cognitive Imprecision and Stake-Dependent Risk Attitudes," NBER Working Paper no. 30417, revised August 2025.

Lejeune, H., and J.H. Wearden, "Vierordt's *The Experimental Study of the Time Sense* (1868) and its Legacy," *European Journal of Cognitive Psychology* 21: 941-960 (2009).

Nieder, A., "Coding of Abstract Quantity by 'Number Neurons' in the Primate Brain," *J. Comparative Physiology A* 199: 1-16 (2013).

Petzschner, F.H., S. Glasauer, and K.E. Stephan, "A Bayesian Perspective on Magnitude Estimation," *Trends in Cognitive Sciences* 19: 285-293 (2015).

References (p.5)

Stocker, A.A., and E.P. Simoncelli, "Noise Characteristics and Prior Expectations in Human Visual Speed Perception," *Nature Neuroscience* 9: 578-585 (2006).

Vierordt, K., *Der Zeitsinn nach Versuchen* [The Sense of Time From Experiments], Tubingen, Germany: Laupp, 1868.

Whalen, J., C.R. Gallistel, and R. Gelman, "Nonverbal Counting in Humans: The Psychophysics of Number Representation," *Psychological Science* 10: 130-137 (1999).

Xiang, Y., T. Graeber, B. Enke, and S. Gershman, "Confidence and Central Tendency in Perceptual Judgments," *Attention, Perception & Psychophysics* 83: 3024-3034 (2021).

Logarithmic Encoding Model: Details

- Suppose that for each of two stimuli with magnitudes X_i $(i = 1, 2)$,

$$r_i \sim N(\log X_i, \nu^2)$$

- And suppose also that 2 is judged to be greater than 1 if and only if $r_2 > r_1$

- Then

$$\text{Prob}[\text{“2 is greater”}] = Prob[r_2 - r_1 > 0] = \Phi\left(\frac{\log X_2/X_1}{\sqrt{2}\nu}\right)$$

is an increasing function of the **ratio** X_2/X_1, independent of absolute magnitudes **[“Weber’s Law”]**

Logarithmic Encoding Model: Details

- Now consider instead an **estimation** task. Suppose

$$r \sim N(\log X, \nu^2)$$

and estimation rule is optimized for a prior distribution

$$\log X \sim N(\mu, \sigma^2)$$

- Then same calculations as on slide 10 imply that, conditional on subjective representation r, the **posterior distribution** for $x \equiv \log X$ is given by $N(\hat{\mu}, \hat{\sigma}^2)$, with

$$\hat{\mu} \equiv \mu + \beta(r - \mu), \quad \hat{\sigma}^2 \equiv \frac{1}{\sigma^{-2} + \nu^{-2}}$$

Properties of a Log-Normal Posterior

- If the posterior distribution of $\log X$ is normal, then the posterior distribution of X is **log-normal.** This is a two-parameter family of probability distributions (parameterized by μ and σ, the mean and standard deviation of $\log X$).

- Properties of a log-normally distributed random variable: if $\log X \sim N(\mu, \sigma^2)$,

$$\mathrm{E}[X] = e^{\mu+(1/2)\sigma^2}$$

$$\mathrm{Var}[X] = \left[e^{\sigma^2} - 1\right] e^{2\mu+\sigma^2}$$

Logarithmic Encoding Model: Details

- If we assume that a subject's estimate $\hat{X}$ of the magnitude X, based on the subjective representation r, is chosen so as to minimize the **mean squared error,**

$$\mathrm{E}[(\hat{X} - X)^2],$$

then $\hat{X}(r) = \mathrm{E}[X|r]$, so that

$$\log \hat{X}(r) = \hat{\mu} + (1/2)\hat{\sigma}^2 = \mu + \beta(r - \mu) + \hat{\sigma}^2/2$$

- Then conditional on the value of X,

$$\log \hat{X} \ \sim \ N(\mu + \beta(\log X - \mu) + \hat{\sigma}^2/2, \ \beta^2\nu^2)$$

Logarithmic Encoding Model: Details

- The properties of a log-normal distribution then imply:

$$\log \mathrm{E}[\hat{X}|X] = \mu + \beta(\log X - \mu) + \hat{\sigma}^2/2 + \beta^2 v^2/2$$

$$\frac{\mathrm{s.d.}[\hat{X}|X]}{\mathrm{E}[\hat{X}|X]} = [e^{\beta^2 v^2} - 1]^{1/2}$$

- Thus the model predicts:
  - $\log \mathrm{E}[\hat{X}]$ a linear function of $\log X$ [linear log-log plot] with slope $0 < \beta < 1$; so $\mathrm{E}[\hat{X}]$ an increasing, **strictly concave** function of X
  - cross-over point X^* between over/under-estimation satisfies $X^* = k \cdot \mathrm{E}[X]$, where $k > 0$ is independent of μ (but depends on σ) [central tendency of judgment]
  - $\mathrm{s.d.}[\hat{X}]$ grows in proportion to $\mathrm{E}[\hat{X}]$ as X increases

Logarithmic Encoding Model: Details

- The model also implies that, in the case of adaptation to a given distribution of presented stimuli [fixing μ and σ], a manipulation that increases the imprecision of the internal representation [i.e., increases ν], such as an increase in time pressure, should
  - increase **subjective uncertainty** about the right estimate, since larger ν implies a larger $\hat{\sigma}$, and hence a larger standard deviation of the posterior,

$$\text{s.d.}[X\,|r] \;=\; [\exp(\hat{\sigma}^2) - 1]^{1/2} \hat{X}(r),$$

    associated with any given response $\hat{X}(r)$; and
  - **reduce the elasticity** β of the mean response $\mathrm{E}[\hat{X}\,|X]$ with respect to increases in X

as observed by Xiang *et al.* (2021).