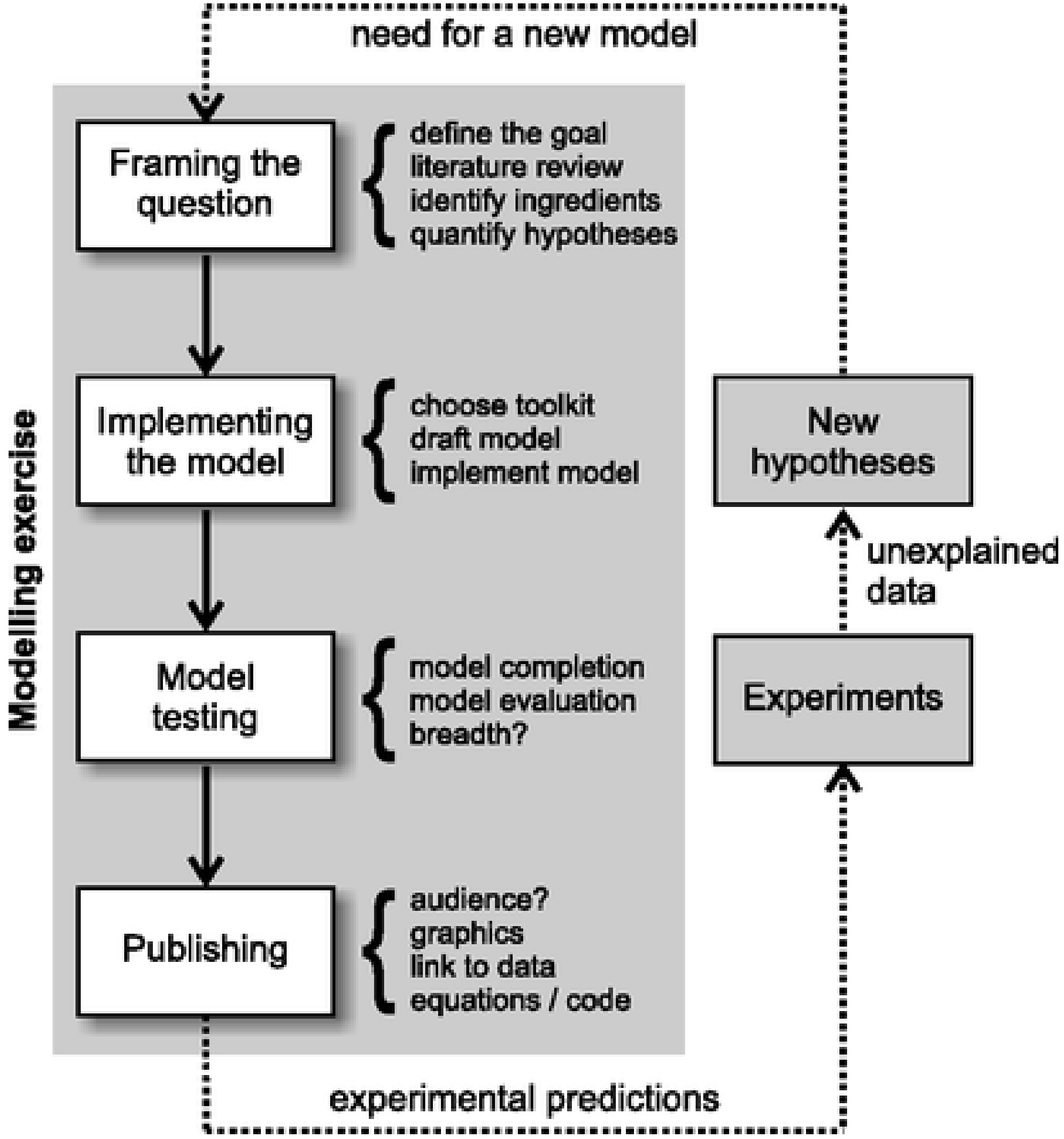
MATLAB for Brain and Cognitive Psychology (Modeling)

Presented by:
Ehsan Rezayat, Ph.D.
Faculty of Psychology and Education, University of Tehran.
Institute for Research in Fundamental Sciences (IPM), School of Cognitive Sciences,
emails: rezayat@ut.ac.ir, rezayat@ipm.ir, erezayat.er@gmail.com

A research question



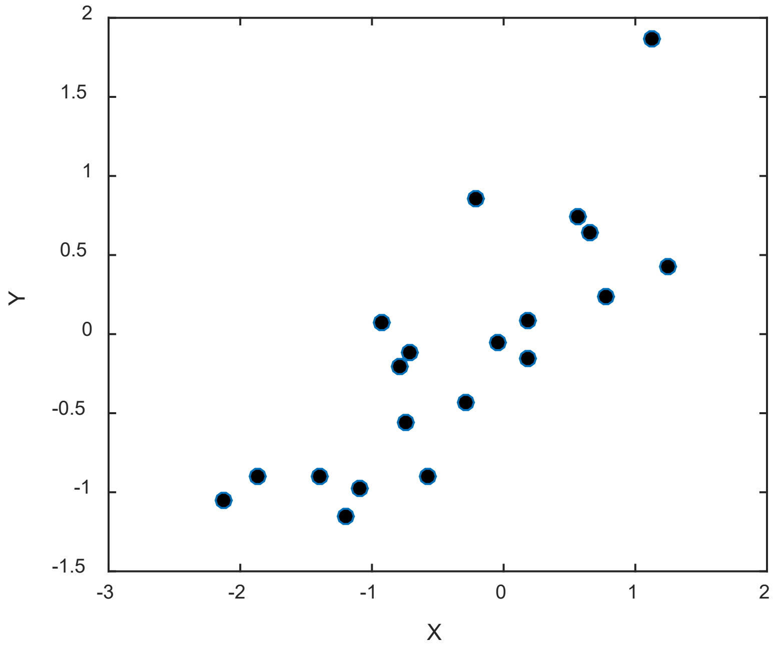
Blohm, G., Kording, K. P., & Schrater, P. R. (2020). A how-to-model guide for Neuroscience. *Eneuro*, *7*(1).

Model uses

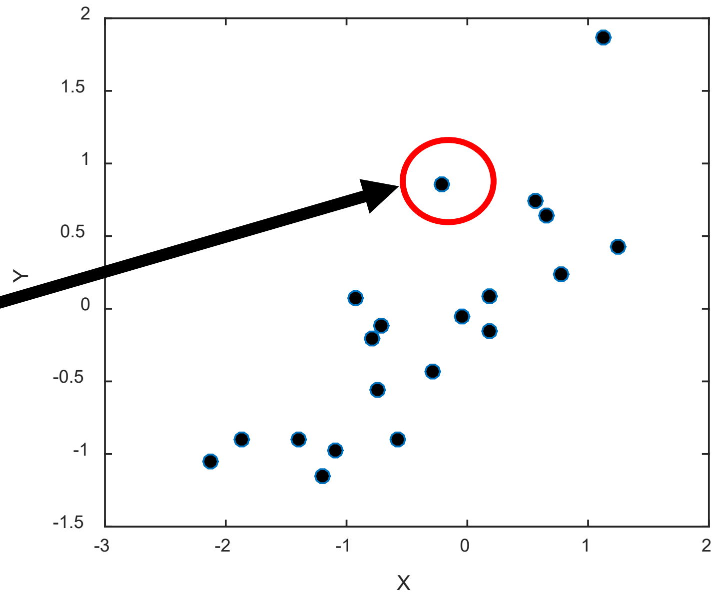
- **Simulation** involves running the model with particular parameter settings to generate ‘fake’ behavioral data.
- **Parameter estimation** involves finding the set of parameter values that best account for real behavioral data for a given model
- **Model comparison** involves trying to compute which of a set of possible models best describes the behavioral data, as a way to understand which mechanisms are more likely to underlie behavior
- **Latent variable inference** involves using the model to compute the values of hidden variables (for example values of different choices) that are not immediately observable in the behavioral data, but which the theory assumes are important for the computations occurring in the brain.

Goal

- fitting a model based on one simple and basic model like the regression line with its two parameters, slope and intercept to the data.



Fitting Models to Data



$$y_i = b_0 + b_1 x_i + e_i$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

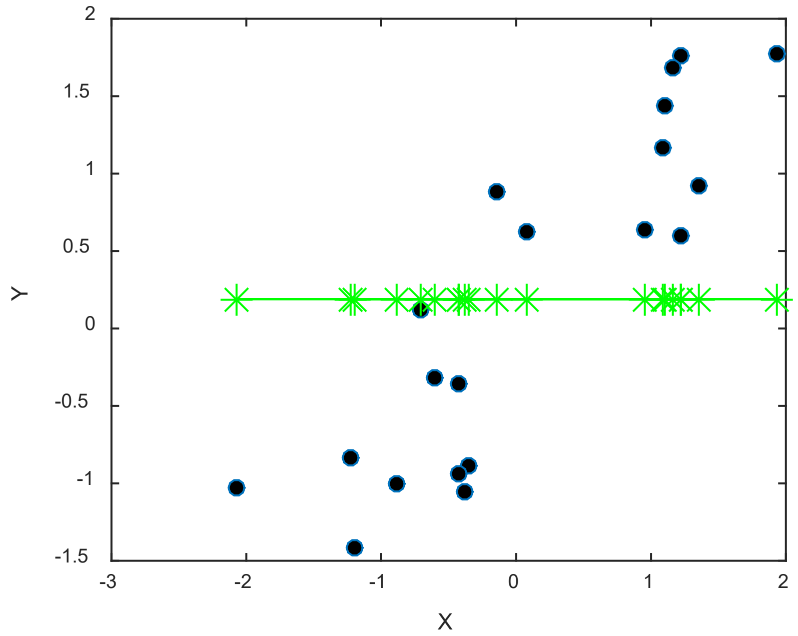
$$Y = Xb + e$$

How exactly the best fitted model is obtained?

- b ???

- Hypothesis 1:
- There is no relationship between Y and X

- $\hat{y} = \frac{y_1 + y_2 + \dots y_n}{n}$

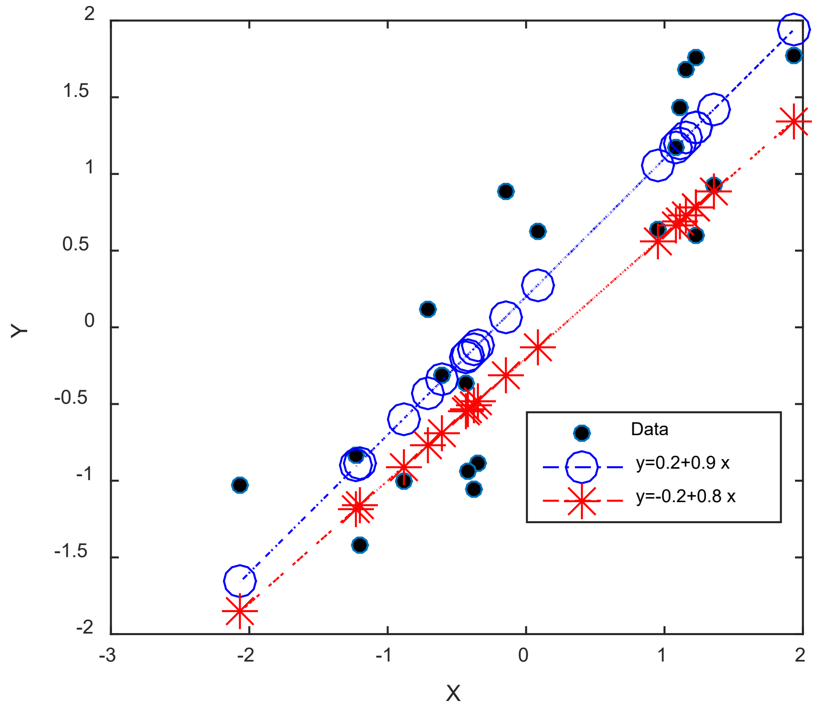


- Hypothesis 2:
- There is linear relationship between Y and X

$$\hat{Y} = Xb$$

- b =?!

- Assume b0 = 0.2 b1 =0.9 then b=[0.2 0.9]
- Assume b0 = -0.2 b1 =0.8 then b=[-0.2 0.8]

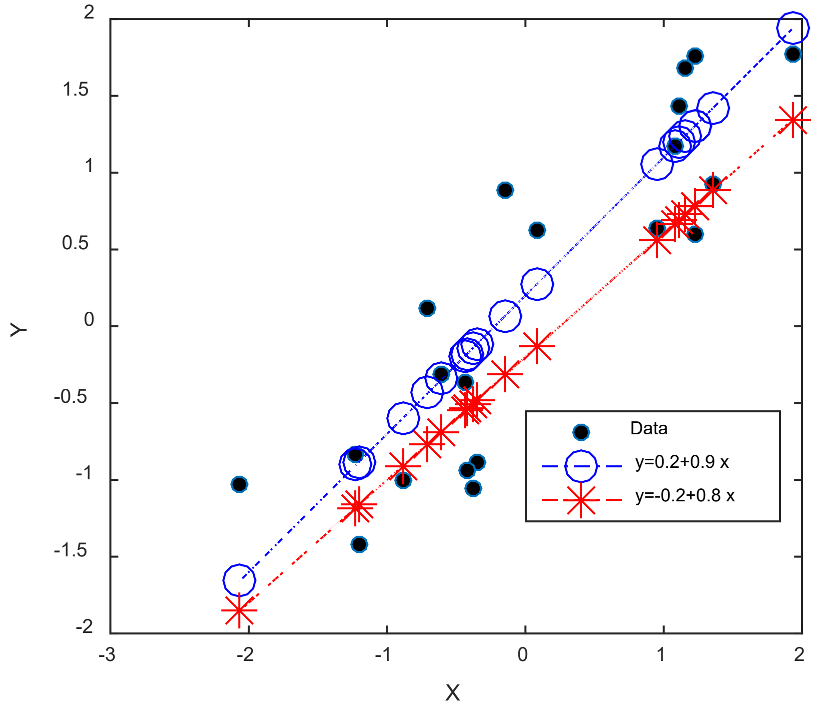


Which parameter is better?

- Calculate the errors

$$RMSD = \sqrt{\frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}{N}}$$

1. Difference between model prediction and observed data
2. Negative error and positive errors are conceptually the same (e.g. we can use the square of error)
3. Make the average and squared root



If b = [0.2 0.9] Error = 0.1163

If b = [-0.2 0.8] Error = 0.1361

One way is checking all parameters pairs Grid search

- Check for

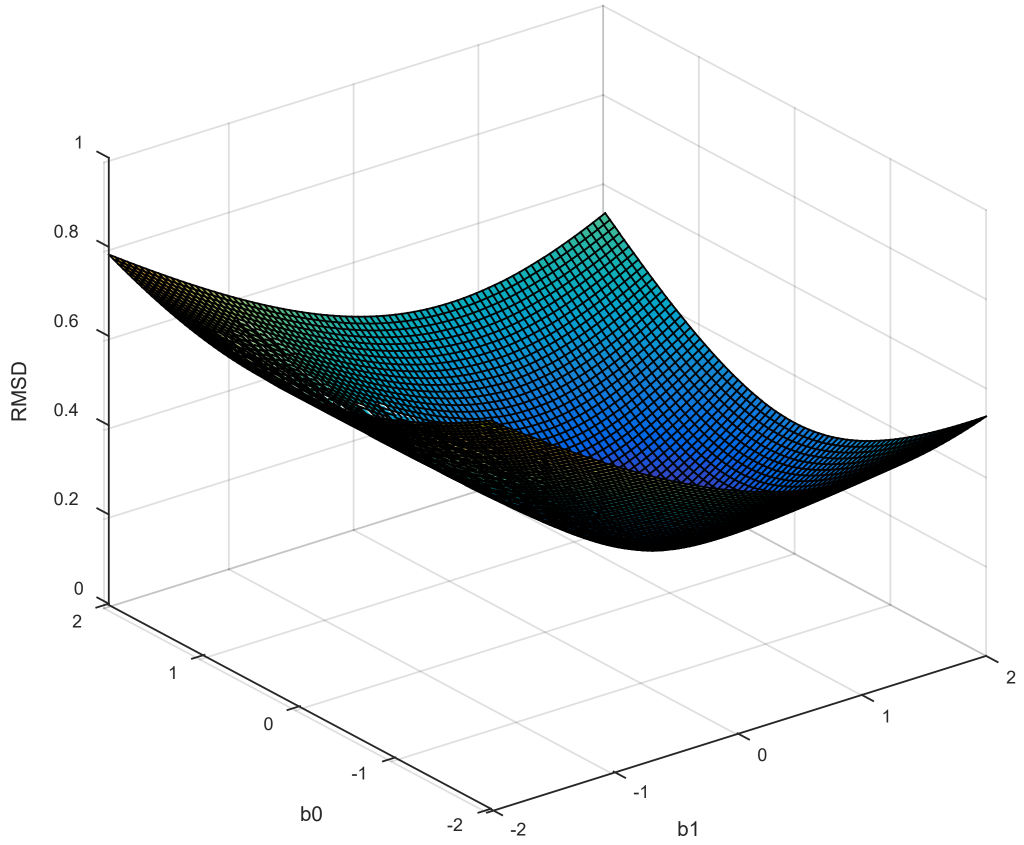
b0 = -2 : 0.05 : 2

b1 = -2 : 0.05 : 2

81 * 81 pairs = 6561 times

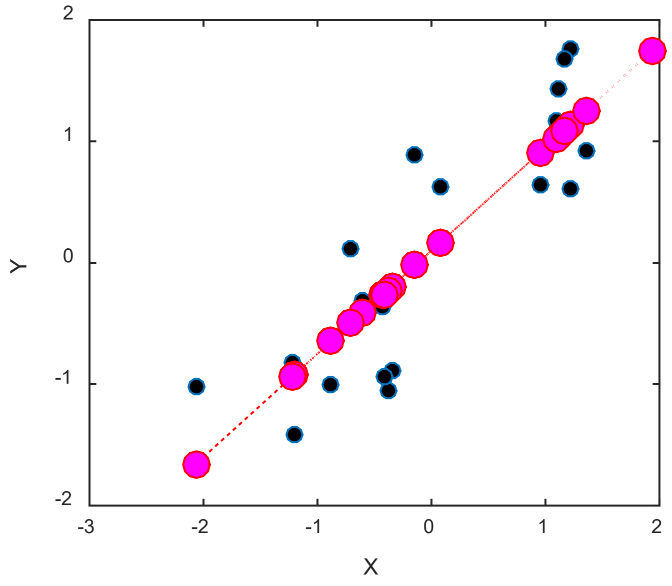
Minimum error pair is 0.1146

$b_0 = 0.1 \;\; b_1 = 0.85$



At minimum error point

$$b_0 = 0.1 \;\; b_1 = 0.85$$



- Grid search is computationally expensive

Least Square Estimator

- Do some calculus

Y = X b

Minimize error = $\sqrt{\frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}{N}}$

Min $\sum_{i=1}^{N}(\hat{y}_i - y_i)^2$

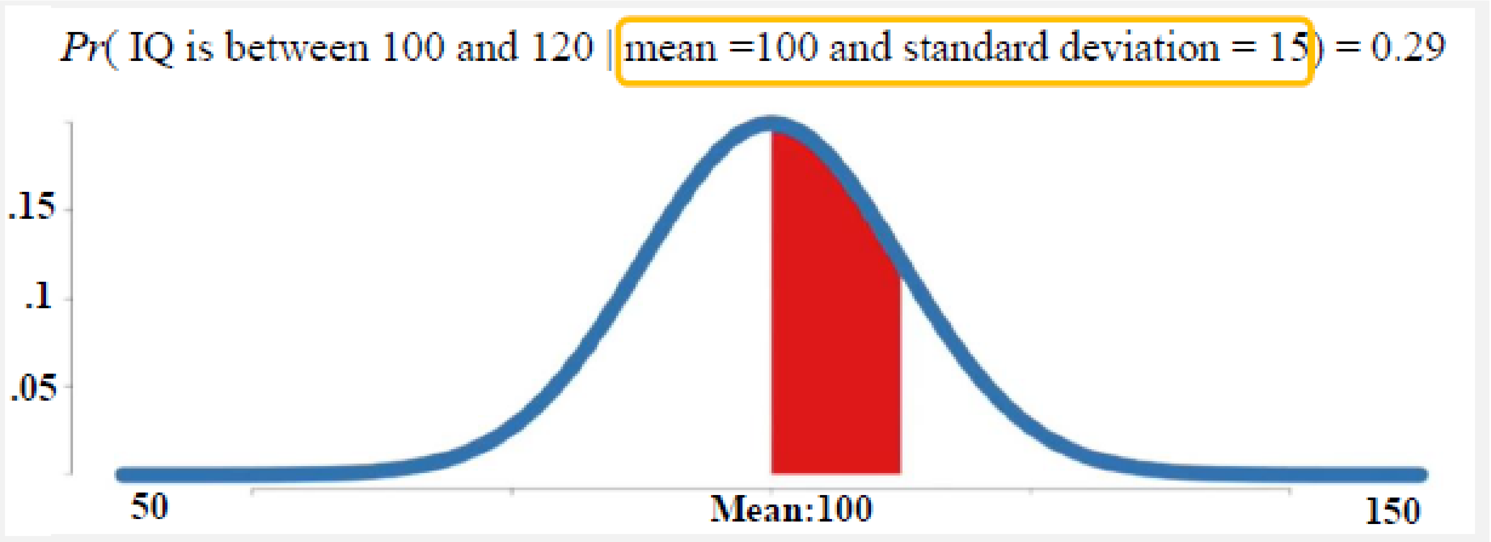
$\sum_{i=1}^{N}(Xb - y_i)^2$ = $(Xb - Y)^T\,(Xb - Y)$

$$MIN\ \overbrace{(Xb - Y)^T(Xb - Y)}^{J(b)}$$

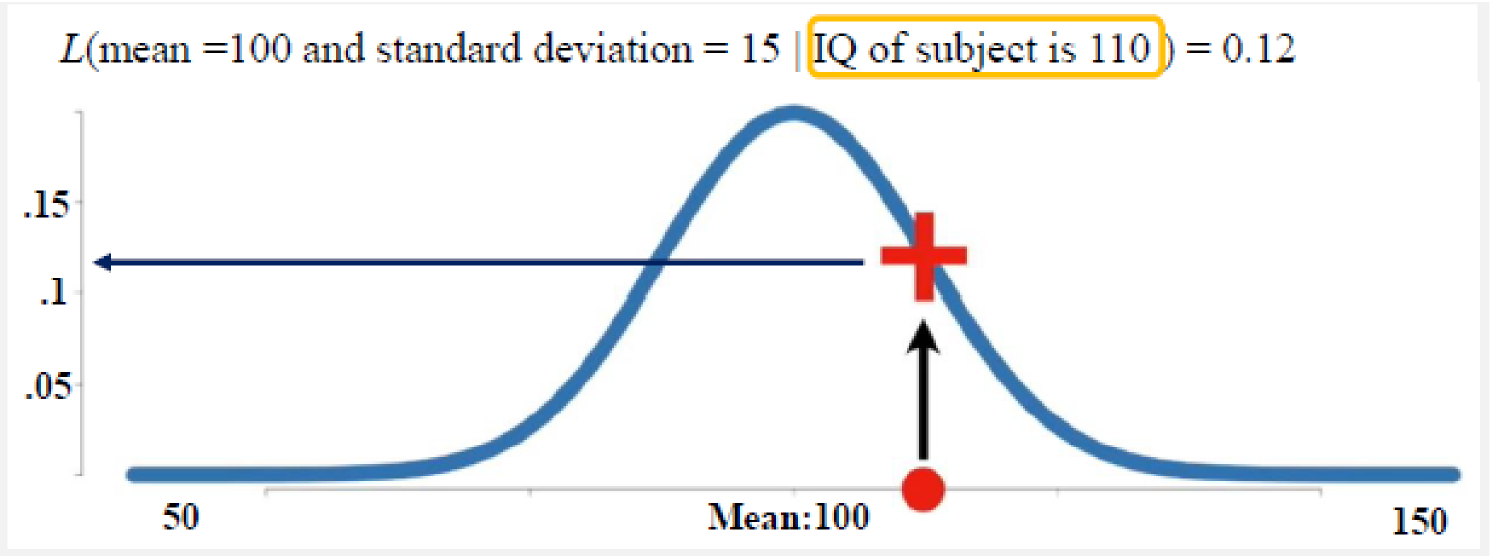
$$\frac{\partial}{\partial b} J(b) = 0$$

What is Maximum likelihood?

Probabilities are the areas under a fixed distribution
***Pr*(data | distribution)**



Likelihood are the y-axis values for fixed data points with distributions that can be moved
***L*(distribution | given data)**



What is likelihood?

For a single data point y , the model M , and a vector of parameter values θ , we will therefore refer to the probability or probability density for an observed data point given the model and parameter values as **f(y |θ , M)**

where f is the probability density function

So, we can obtain a joint probability or probability density for the data in the vector y under the assumption that the observations in *y* are **independent**

$$f(\mathbf{y}) = \prod^{k} f(y_k|\boldsymbol{\theta})$$

- Rather than keeping the model and the parameter values **fixed** and looking at what happens to the probability function or probability density across different possible data points,

Pr(data | distribution)

- we instead keep the data and the model **fixed** and observe changes in ***likelihood*** values as the parameter values change.

***L*(distribution | given data)**

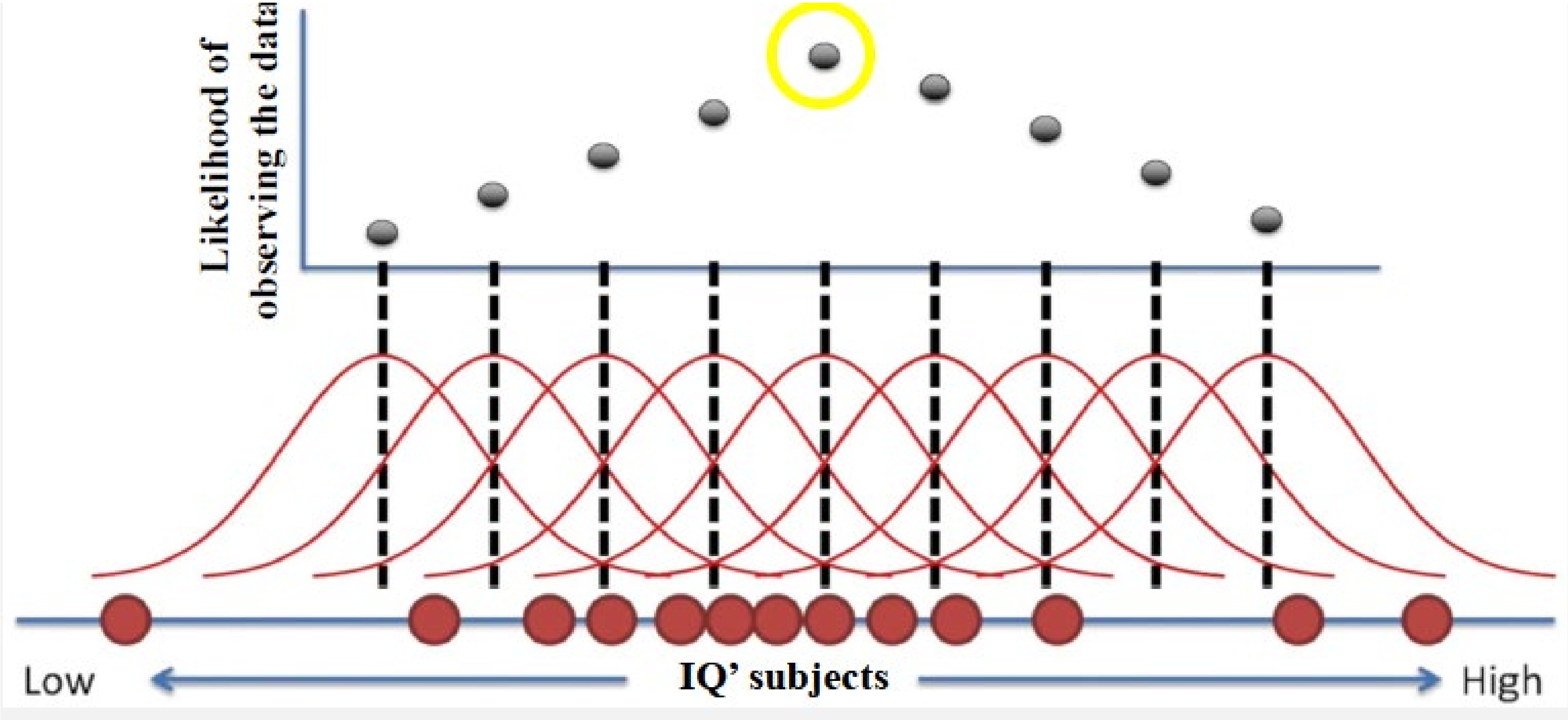
$$P(y|\theta)P(\theta) = P(\theta|y)P(y).$$

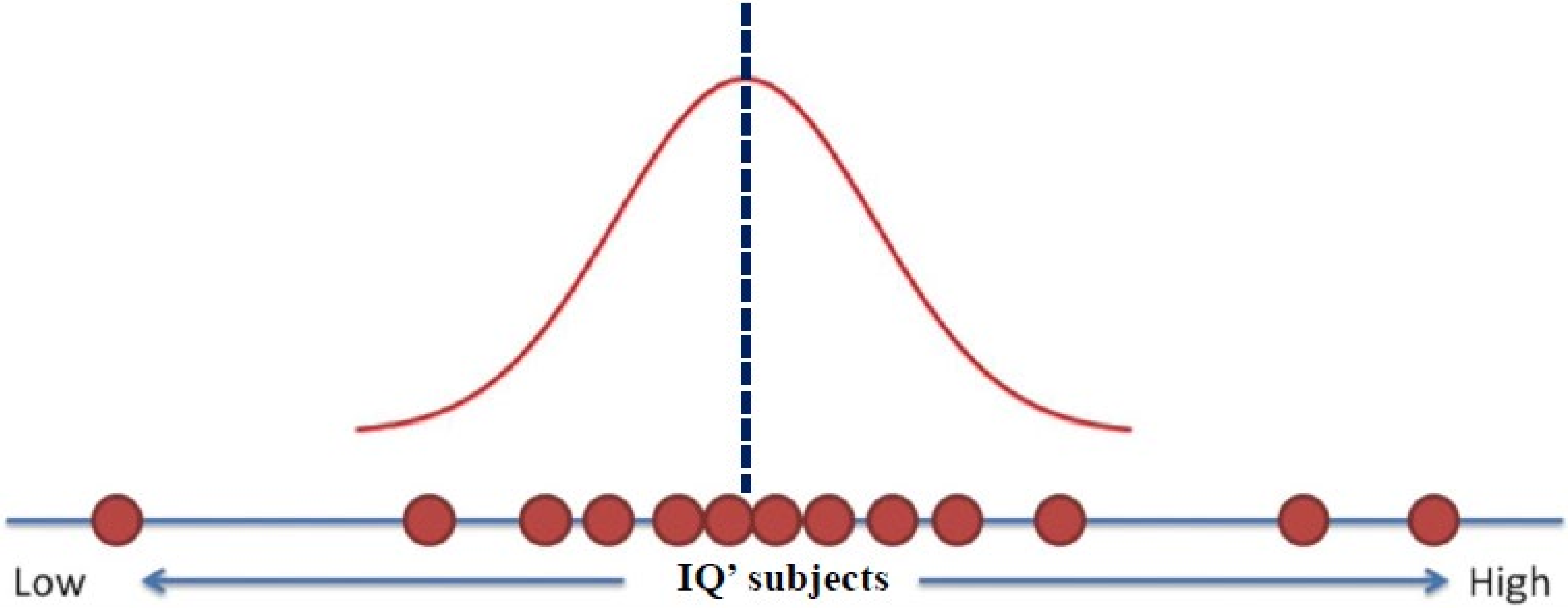
$$L(\theta|\mathbf{y}) = \prod^{k} L(\theta|y_k)$$

$$L(\theta|\mathbf{y}) = \prod^{k} L(\theta|y_k)$$

$$\ln L(\theta|\mathbf{y}) = \sum_{k=1}^{K} \ln L(\theta|y_k),$$

$$\ln \prod_{k=1}^{K} f(k) = \sum_{k=1}^{K} \ln f(k).$$


Likelihood of
observing the data
Low
IQ' subjects
High


Low
IQ' subjects
High

Psychometric function

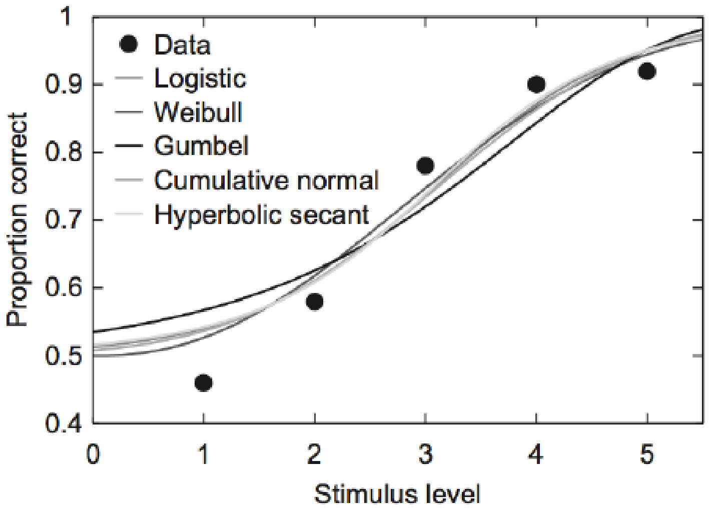
- Cumulative Normal Distribution

$$F_N(x;\alpha,\beta) = \frac{\beta}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp\left(-\frac{\beta^2(x-\alpha)^2}{2}\right)$$

- Logistic $$F_L(x;\alpha,\beta) = \frac{1}{1+\exp(-\beta(x-\alpha))}$$

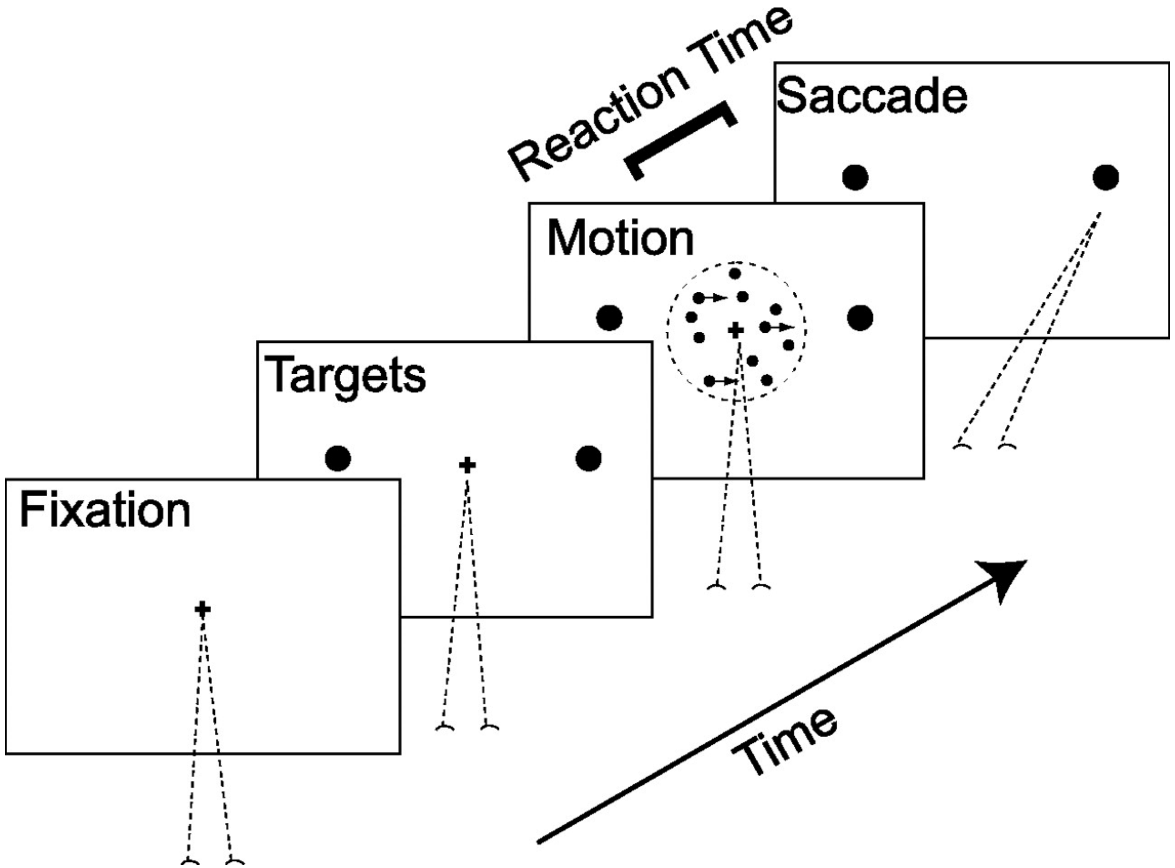
- Weibull

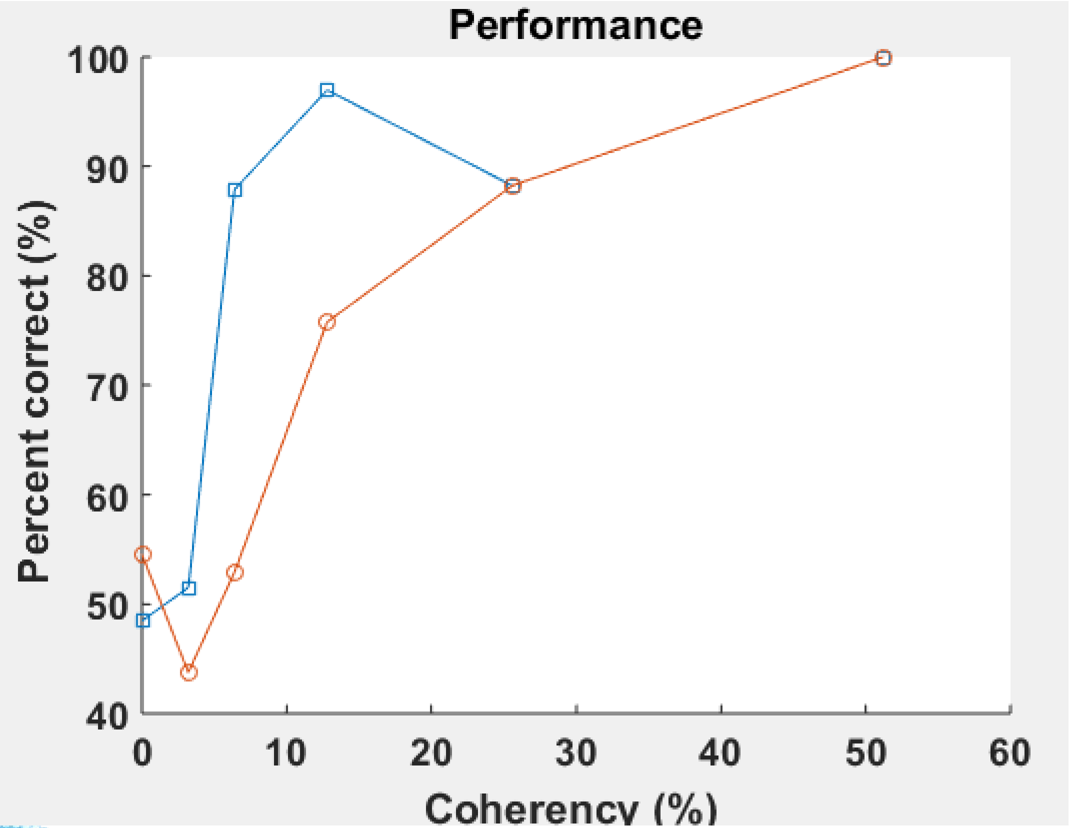
$$F_W(x;\alpha,\beta) = 1-\exp\left(-\left(\frac{x}{\alpha}\right)^{\beta}\right)$$

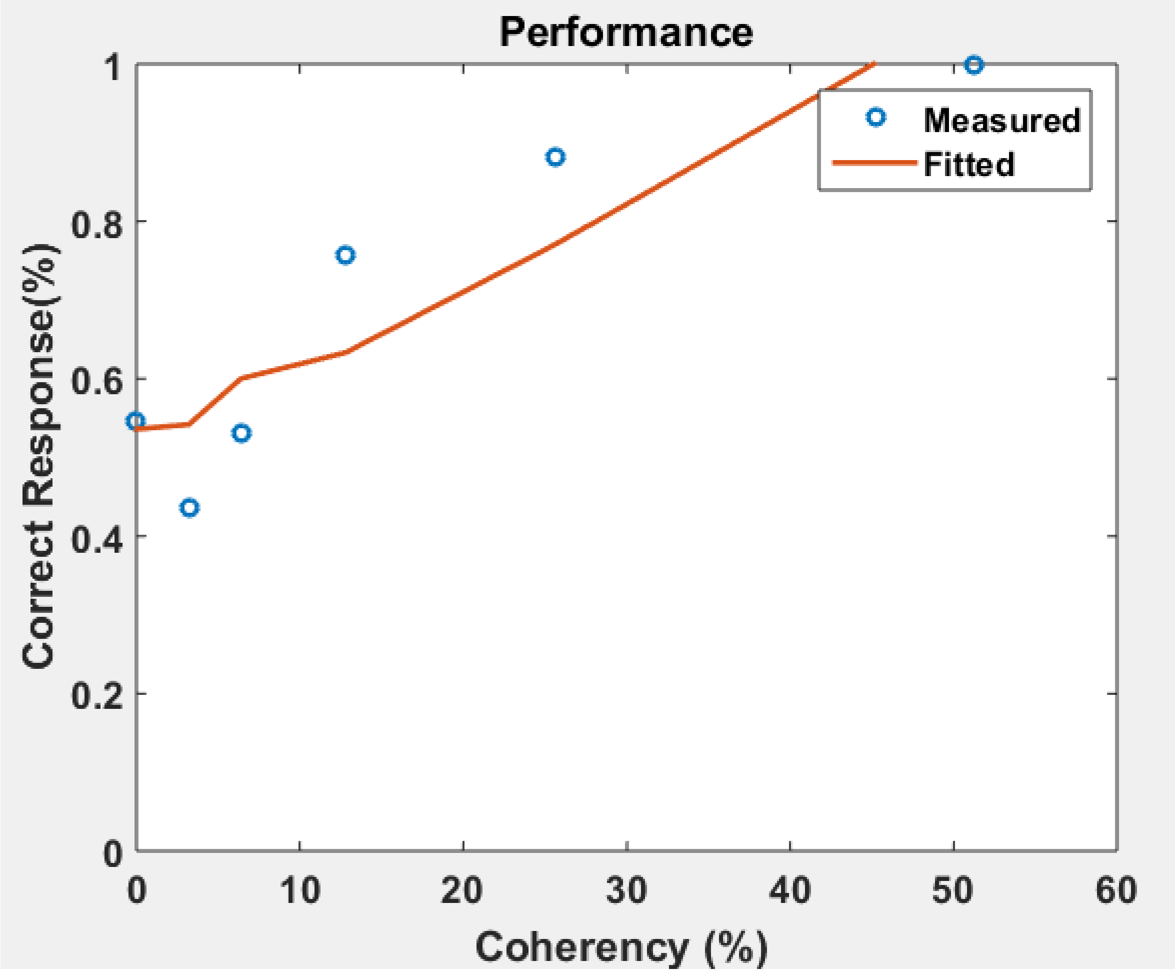


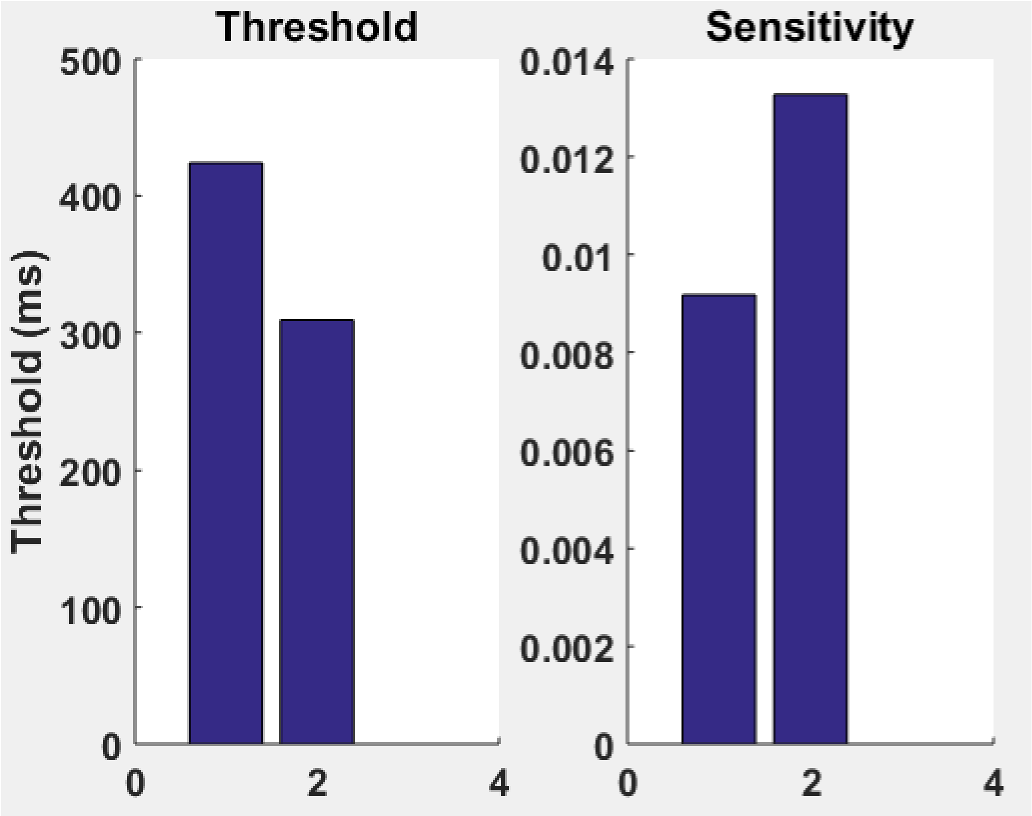
Kindom & Prins (2010)

RDM task

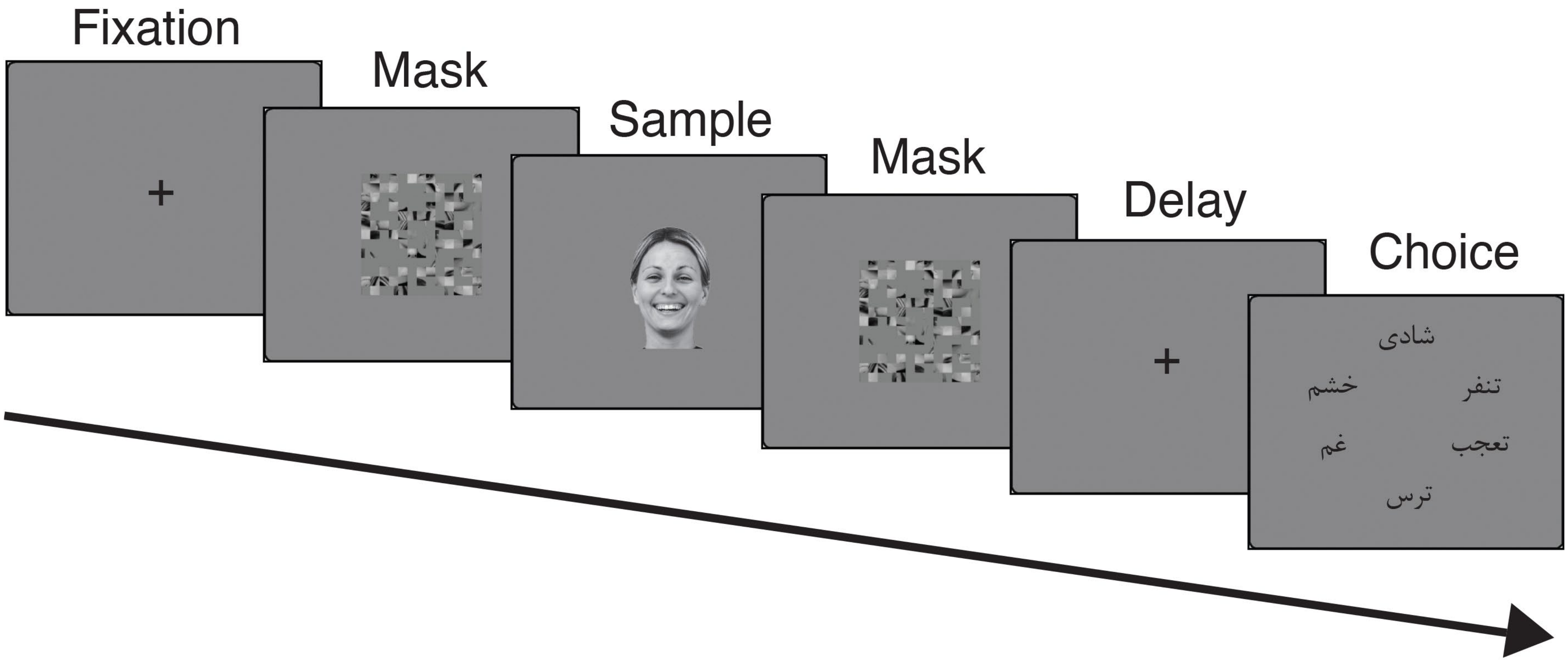


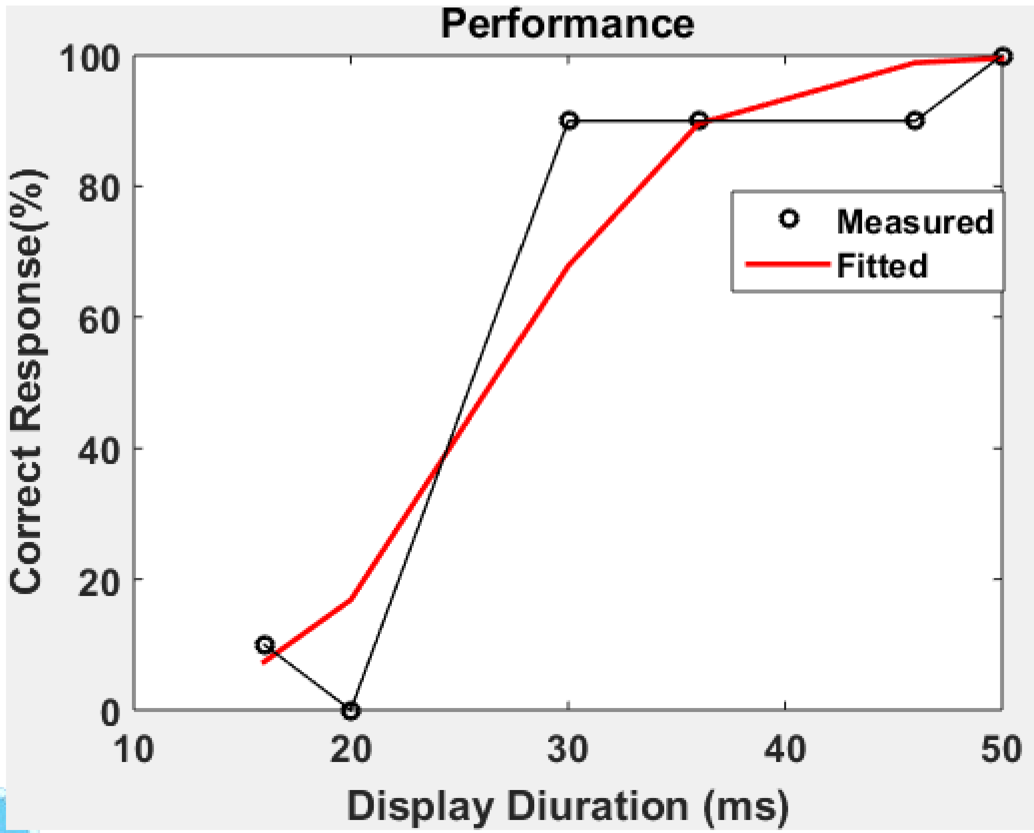

Performance
Percent correct (%)
Coherency (%)

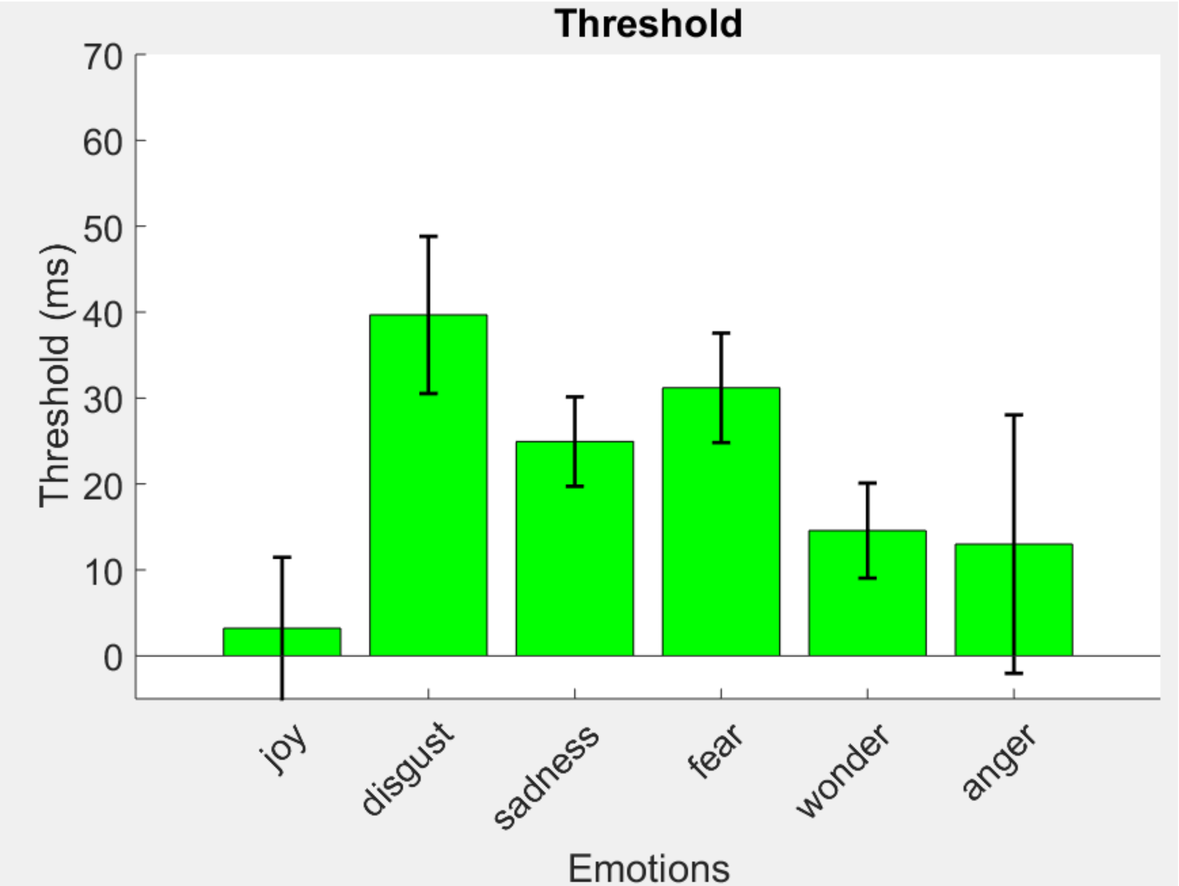

Performance
Correct Response(%)
Coherency (%)
Measured
Fitted



Threshold
Sensitivity
Threshold (ms)
500
400
300
200
100
0
0
2
4
0.014
0.012
0.01
0.008
0.006
0.004
0.002
0
0
2
4

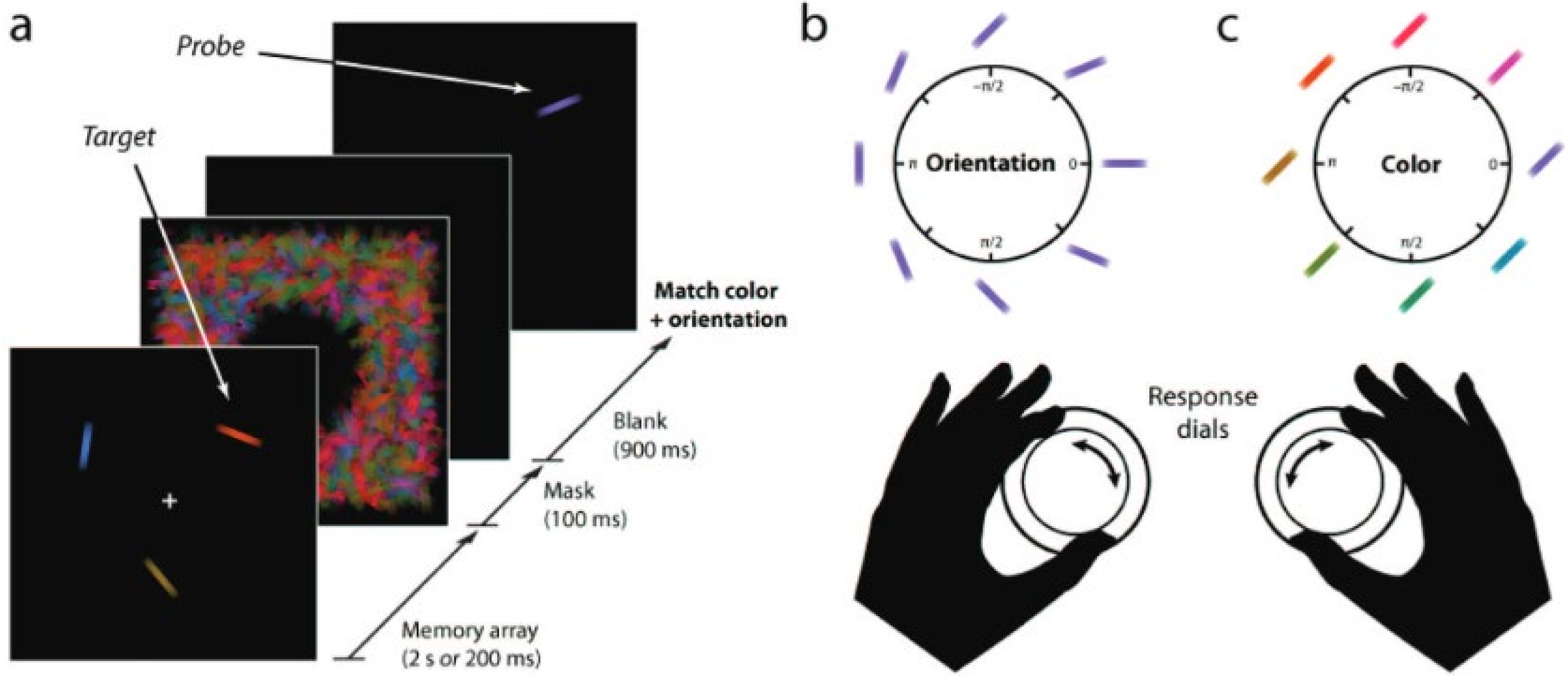

Fixation
Mask
Sample
Mask
Delay
Choice
شادی
خشم
تنفر
غم
تعجب
ترس


Performance
100
80
60
40
20
0
10
20
30
40
50
Correct Response(%)
Display Diuration (ms)
Measured
Fitted

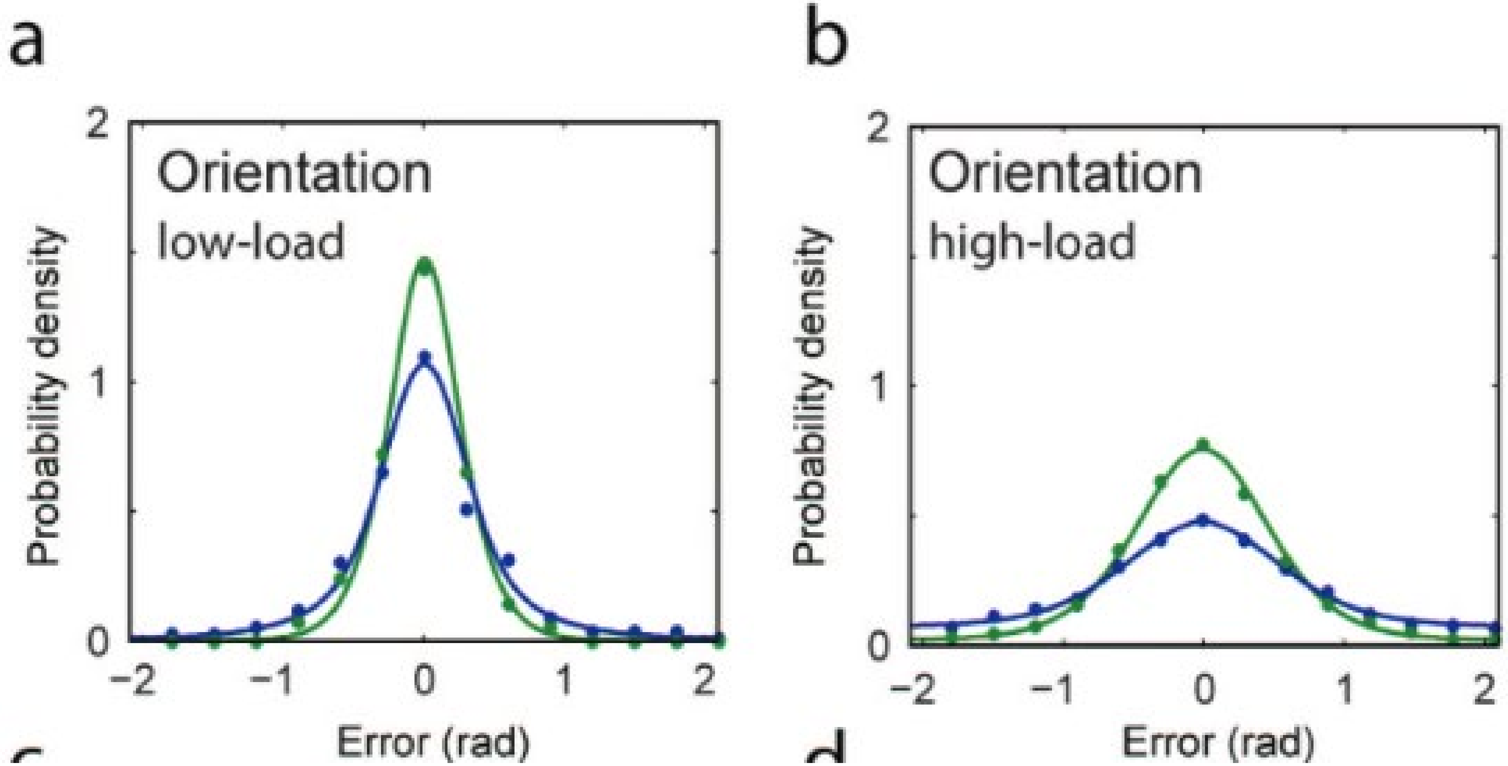

Threshold
70
60
50
40
30
20
10
0
Threshold (ms)
joy
disgust
sadness
fear
wonder
anger
Emotions

Continues response task

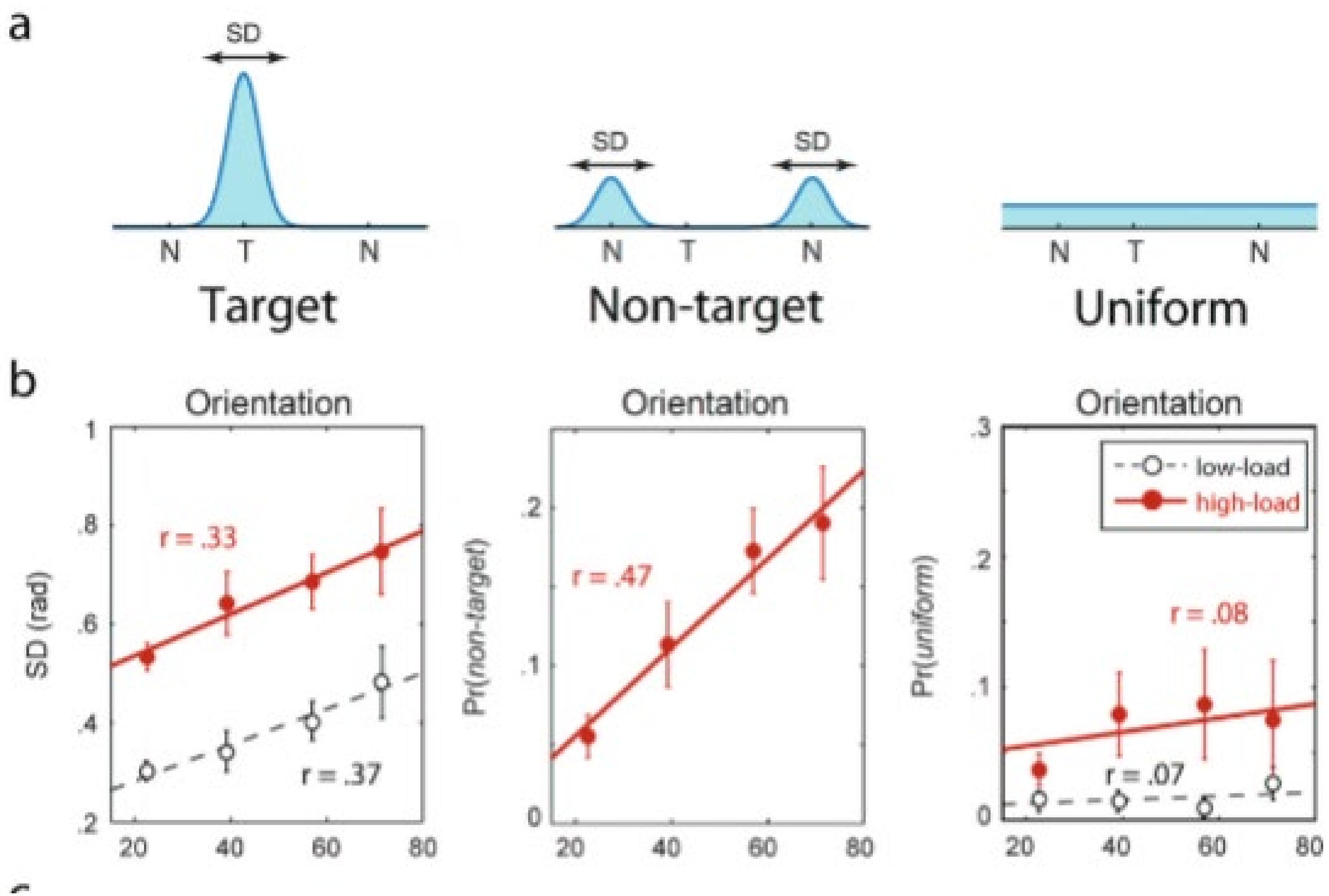
- https://psycnet.apa.org/fulltext/2013-29995-001.pdf



Error distribution



Error components



Assignment#14

- Generate three random normal distribution with different parameters (mean and variance).
- Plot the distributions.
- Using LSE and MLE method fit a Gaussian function to your data.
- Compare your fitted value with defined variables