Analysis II: Correlation and Regression

PNI Summer Internship 2020
Mai Nguyen

Overview

- Today's data set
- Correlation
- Linear regression
- HW #4

Today's data set

- Data from Sarah Wilterson, graduate student in Jordan Taylor's Lab at Princeton

J Neurophysiol 120: 2640–2648, 2018.
First published September 12, 2018; doi:10.1152/jn.00283.2018.

RESEARCH ARTICLE | *Control of Movement*

Relative sensitivity of explicit reaiming and implicit motor adaptation

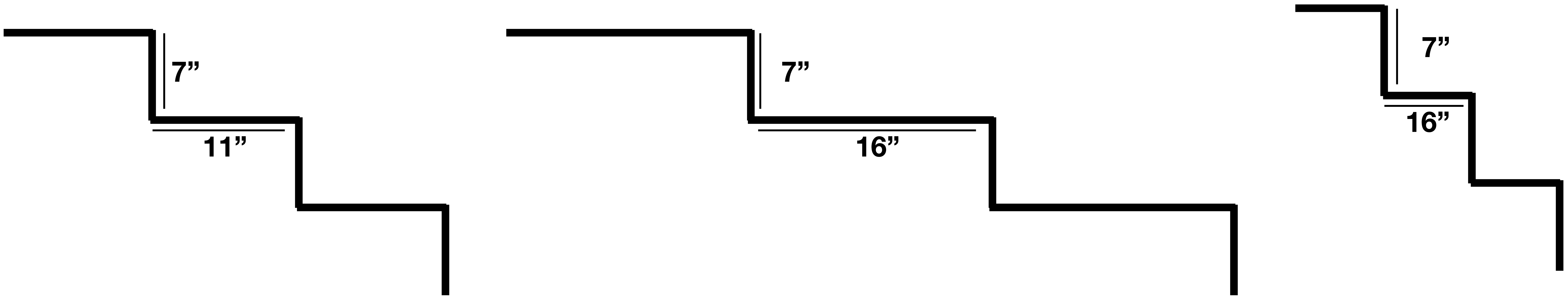
Sarah A. Hutter[1,2] **and** **Jordan A. Taylor**[1,2]
[1]*Department of Psychology, Princeton University, Princeton, New Jersey; and* [2]*Princeton Neuroscience Institute, Princeton University, Princeton, New Jersey*

Submitted 30 April 2018; accepted in final form 11 September 2018

Today's data set

- **Motor adaptation**: process of adapting motor control to changing environment

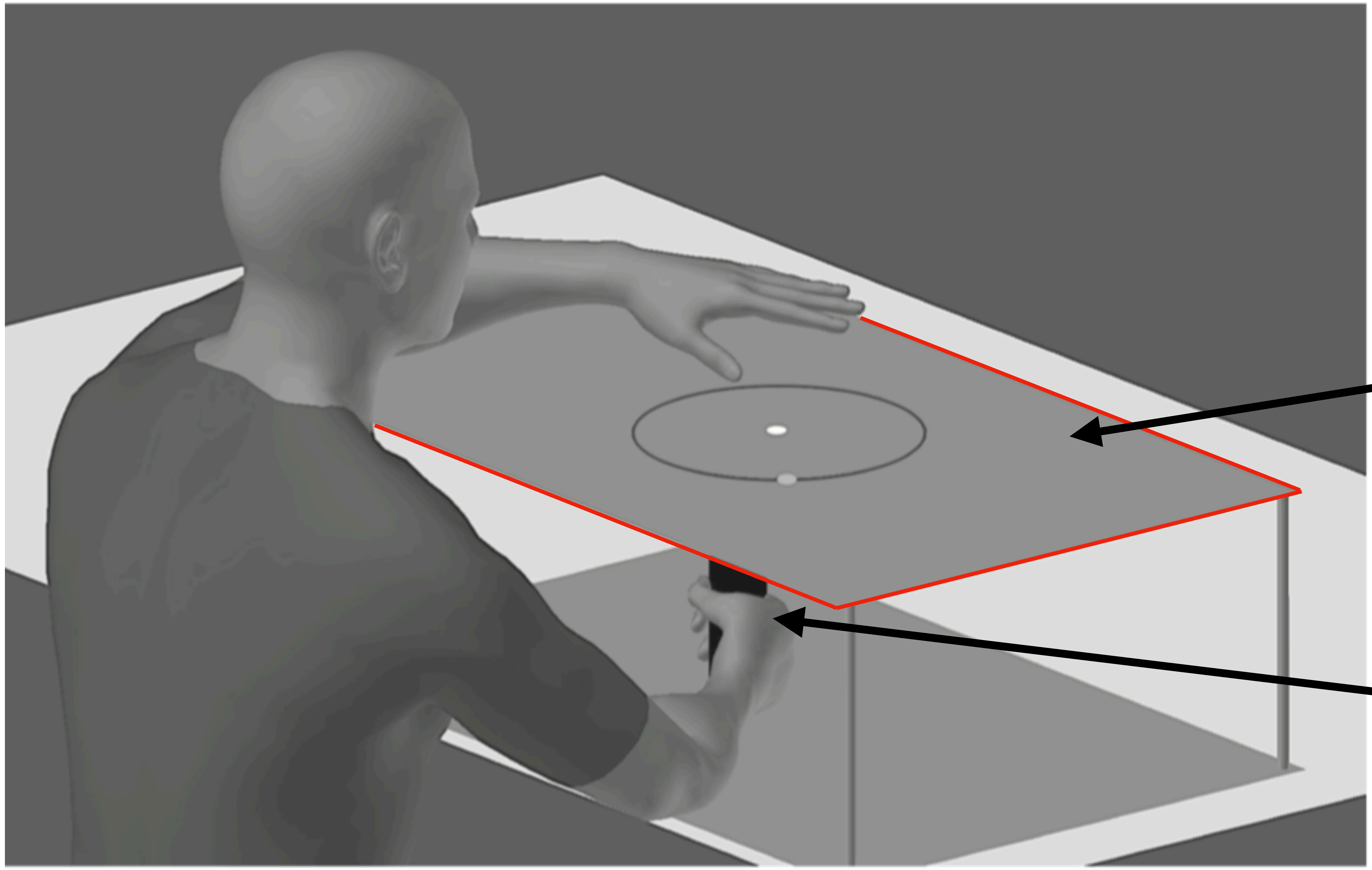
- Example: adjusting to different stairs



Today's data set

- What contributes to motor adaptation?
    - **Implicit adaptation**: unconsciously adjust motor movements following feedback (e.g. stub your toe on stairs and so reflexively change gait)
    - **Explicit re-aiming:** consciously change motor movement (e.g. see weird stairs and take smaller or larger steps to adapt)
- Hutter et al 2018: How do implicit adaptation and explicit re-aiming work depending on the level of error?

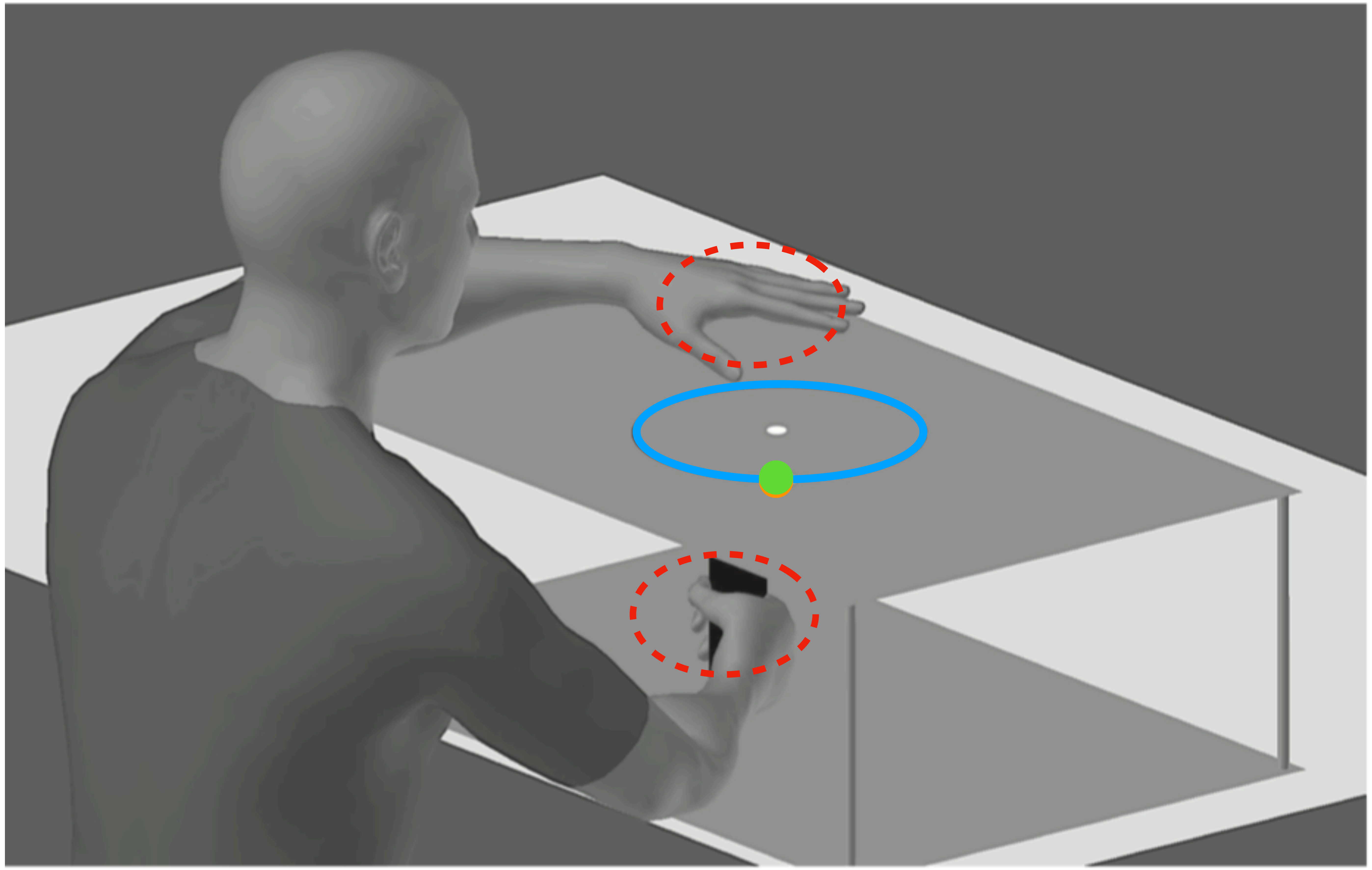
Today's data set



Digital screen occludes right hand

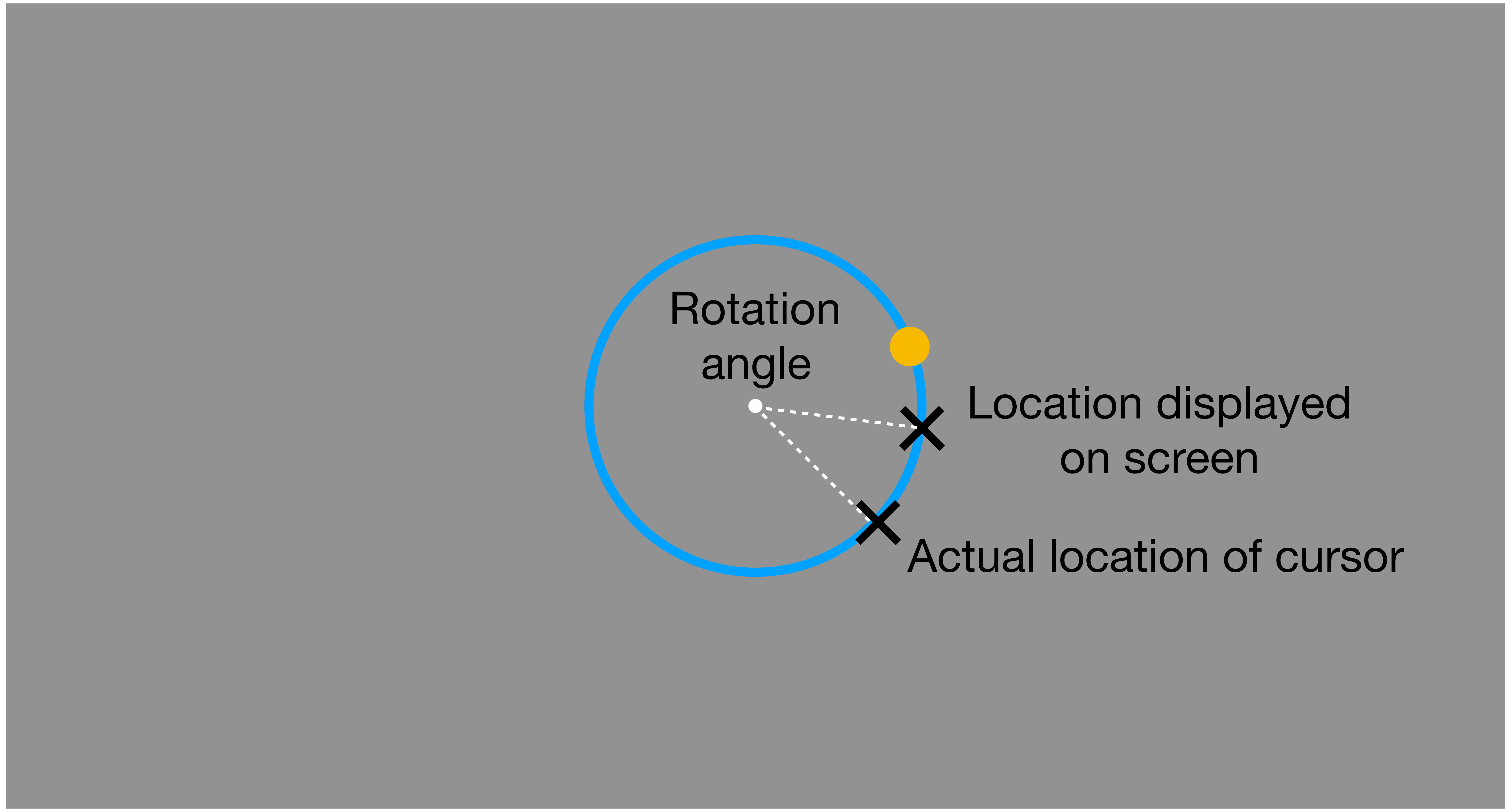
Right hand holds digital pen that controls cursor on digital screen

Today’s data set



1. Subject moves digital pen to center
2. Orange target dot appears anywhere on circle with 7cm radius from center
3. Subj taps where they will aim with left hand (explicit aiming)
4. Subj moves digital pen with right hand to hit target with cursor
5. Target turns green when subject hits target

Today's data set



- Position of cursor on the screen is different from actual position
- Rotate between -16 to +16 degrees -> manipulates amount of error

Today's data set

- N = 80 subjects, 20 in each consistency group
- Two manipulations
    - Rotation angle: -16, -8, -4, -2, 0, +2, +4, +8, +16
    - Consistency: keep rotation same rotation angle for 1, 2, 3, or 7 trials in a row
- Measures
    - Explicit re-aiming angle: angle between location of target and where subjects tap with LH at start of each trial
    - Implicit adaptation angle: Hand angle - aiming angle
- Questions:
    - How do implicit adaptation and explicit re-aiming change with error (rotation angle)?
    - Are there differences in adaptation and re-aiming with consistency?

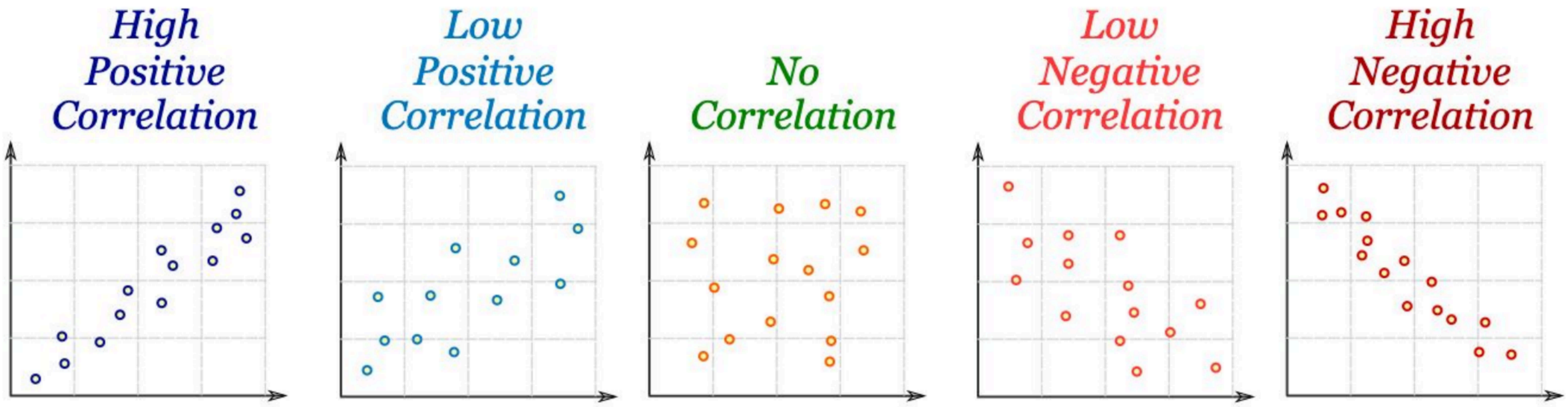
Today’s data set

- N = 80 subjects, 20 in each consistency group
- Two manipulations
    - Rotation angle: -16, -8, -4, -2, 0, +2, +4, +8, +16
    - Consistency: keep rotation same rotation angle for 1, 2, 3, or 7 trials in a row
- Measures
    - Explicit re-aiming angle: angle between location of target and where subjects tap with LH at start of each trial
    - Implicit adaptation angle: Hand angle - aiming angle
- Questions:
    - **How do implicit adaptation and explicit re-aiming change with error (rotation angle)?**
    - Are there differences in adaptation and re-aiming with consistency?

Correlation

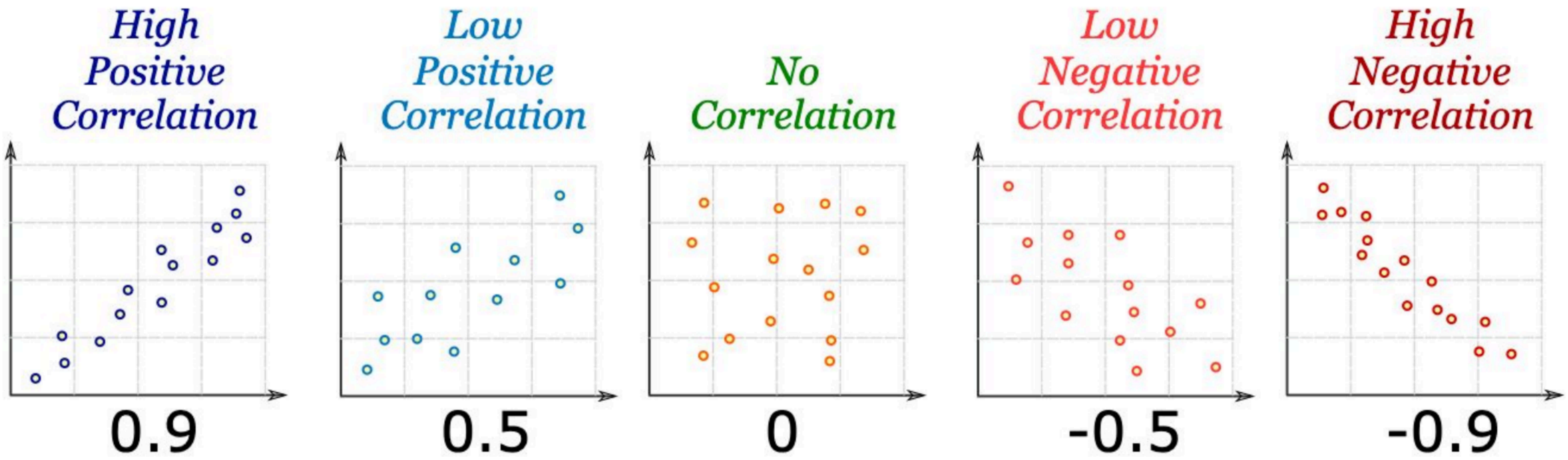
- Previously:
    - t-tests look for differences in means between groups/conditions
    - e.g. Do people learn faster in high risk versus low risk condition?
- Correlation: look for relationship between two continuous variables
    - e.g. Do these two things go up or down together?

Correlation



- **Pearson's correlation, r**
- Measures relationship between x and y
- Always ranges between -1 and 1

Correlation



- **Pearson's correlation, r**
- Measures relationship between x and y
- Always ranges between -1 and 1

Pearson's correlation, r

- Two variables: x and y with means $\bar{x}$ and $\bar{y}$

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

$$r = \frac{\textbf{covariance of x and y}}{\textbf{var in x and y}}$$

Testing for significance of r

- Null hypothesis: r = 0, no relationship between x and y
- Alternate hypothesis: r is not 0, there is a relationship between x and y
- P-values
    - MATLAB assumes r is drawn from a t-distribution (various assumptions)
    - Same logic as for t-tests, ANOVA

Pearson’s correlation r

- In MATLAB:
    - `[r, p] = corr(data1, data2)`

Linear regression

- Model relationship between x and y (e.g. rotation angle and implicit adaptation angle)

- Find the “best fit line”

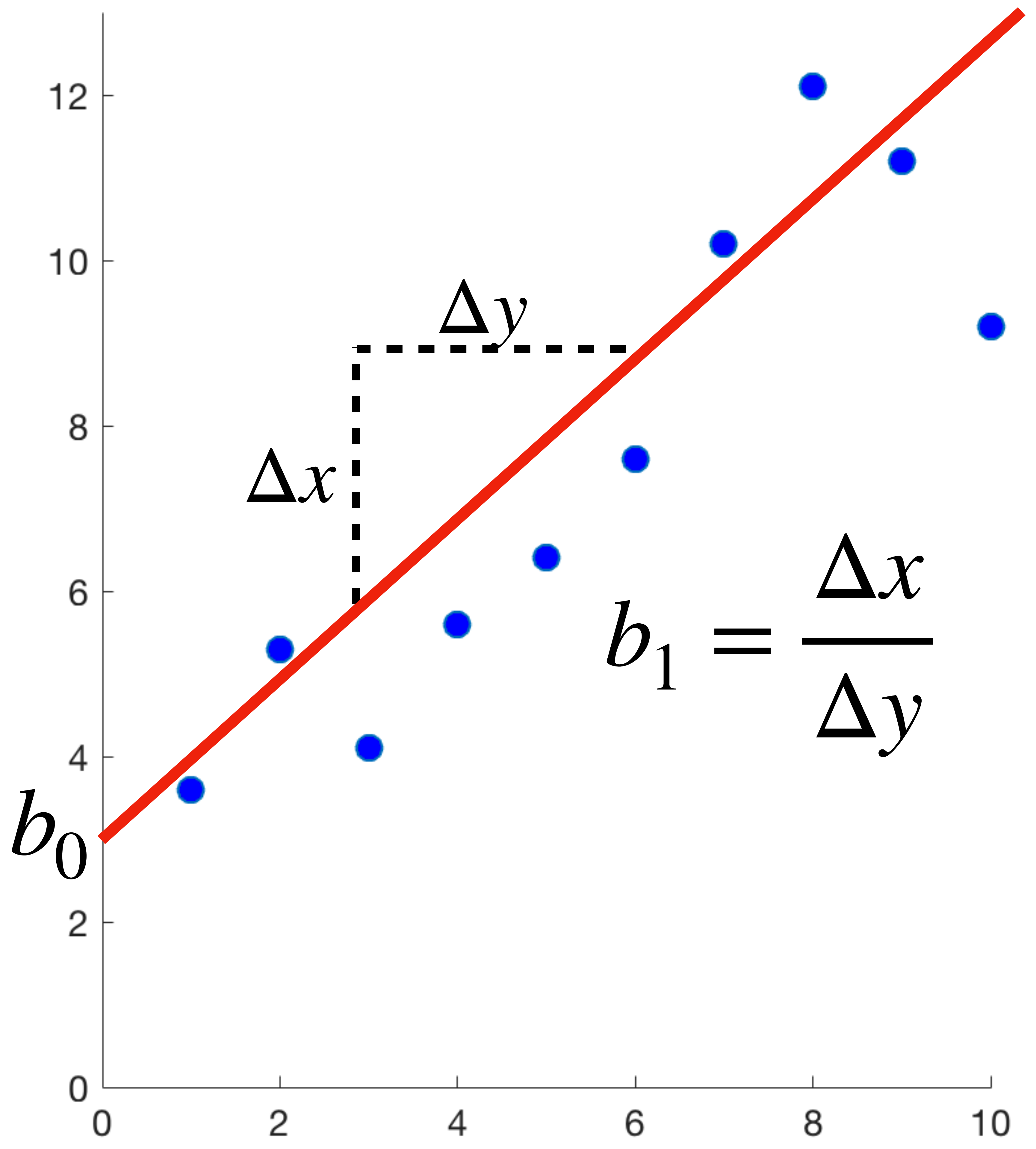
$$y = b_0 + b_1 x$$

DV, Outcome

y-intercept

slope, correlation coefficient

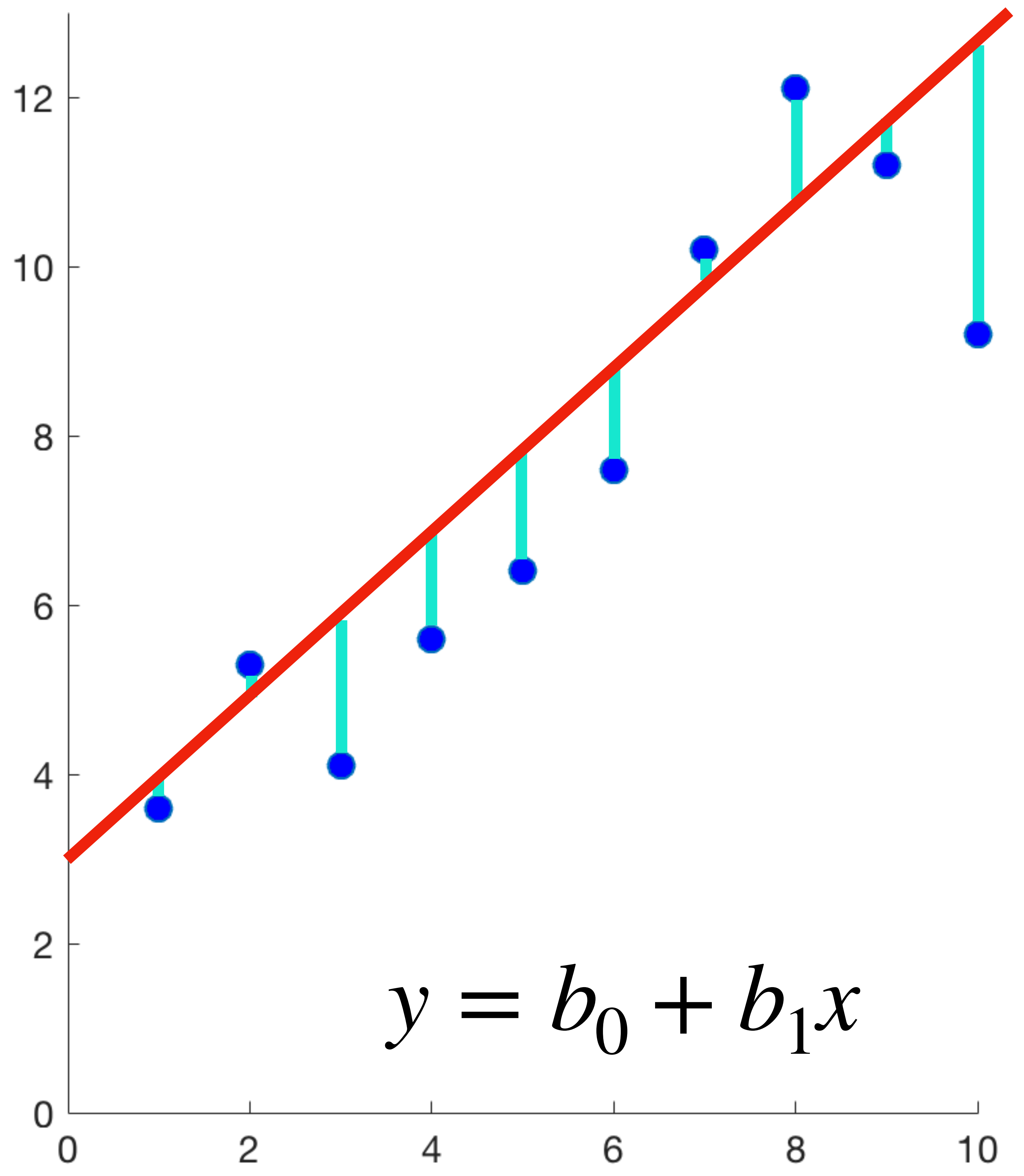
Predictors, Features



Linear regression

- Fit model with Least Squares Estimation

- Minimizes difference between observed data and the best-fit line



Linear regression

- Significant of overall model: many different ways
    - F-stat and p-value: does model have coefficients that are different from 0? (e.g. do the model predictors have any relationship to outcomes?)
    - R^2: proportion of variance explained by model predictors
- Significance of b_0 and b_1: many different ways
    - Approach today: get b_0 and b_1 for each subject, then use t-tests to test that coefficients are different from 0

Linear regression in MATLAB

- `[b, ci, r, rint, stats] = regress(Y, X)`
    - Y = outcomes or DV (implicit angle, aim angle)
    - X = predictors or IV (rotation)

```
x =

     1   -16
     1    -8
     1    -4
     1    -2
     1     0
     1     2
     1     4
     1     8
     1    16
```

predictors

```
y =

    3.6000
    5.3000
    4.1000
    5.6000
    6.4000
    7.6000
   10.2000
   12.1000
   11.2000
    9.2000
```

outcomes

Linear regression in MATLAB

- `[b, ci, r, rint, stats] = regress(Y, X)`

```
b =

    2.6400
    0.8891
```

b_0
b_1

```
ci =

    0.3951    4.8849
    0.5273    1.2509
```

95% CI for b0
95% CI for b1

```
resid =

    0.0709
    0.8818
   -1.2073
   -0.5964
   -0.6855
   -0.3745
    1.3364
    2.3473
    0.5582
   -2.3309
```

difference between observed y and model y

```
stats =

    0.8006   32.1133    0.0005    2.0308
```

R^2 (var exp)	F-stat	p-value	err var

Multiple regression

$$y = b_0 + b_1 x + b_2 x + b_1 b_2 x$$

HW #4

- Replicate subset of findings from my preprint (link)

Teacher-student neural coupling during teaching and learning

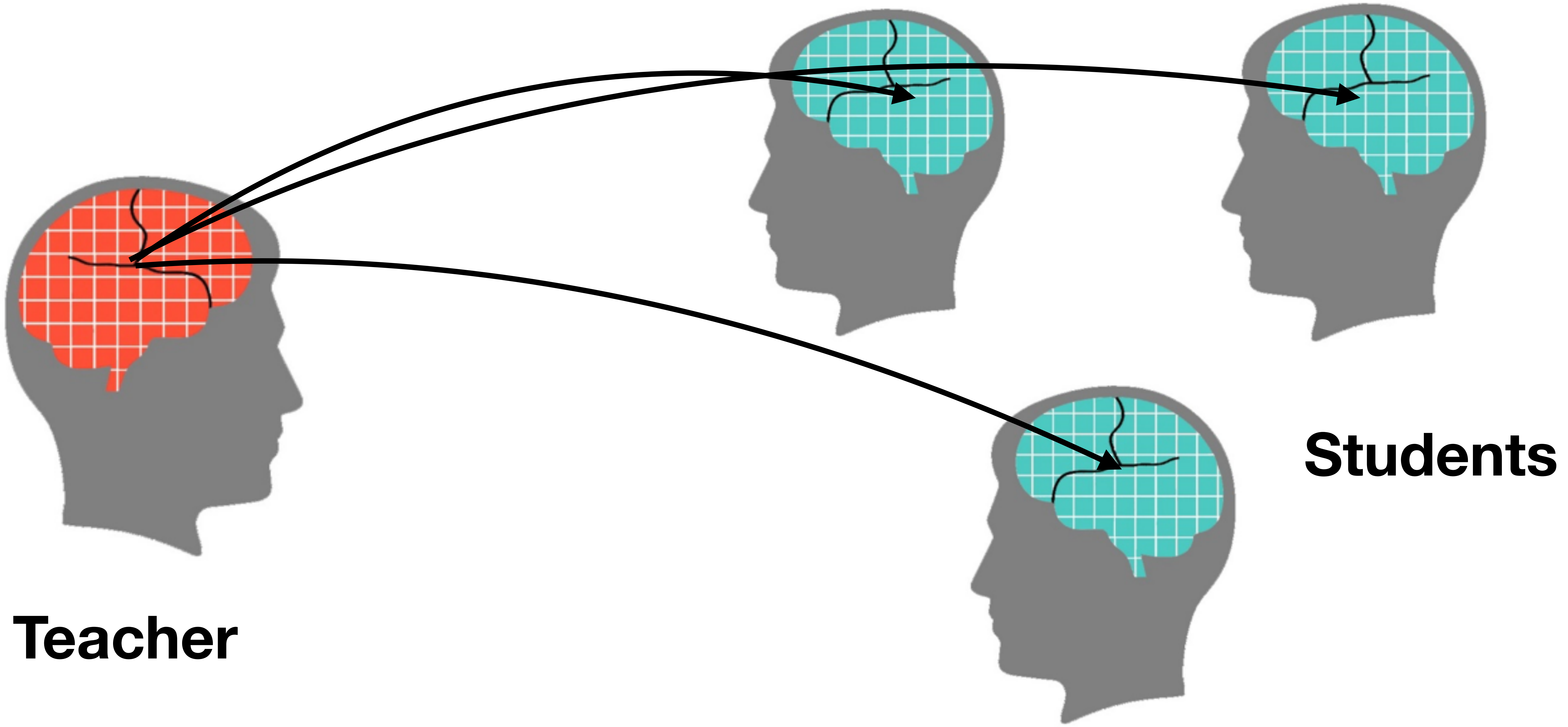
Mai Nguyen[1], Ashley Chang[1], Emily Micciche[2], Meir Meshulam[3], Samuel A. Nastase[3], & Uri Hasson[1,3]

[1]Department of Psychology, Princeton University

[2]Department of Psychology, Vanderbilt University

[3]Princeton Neuroscience Institute, Princeton University

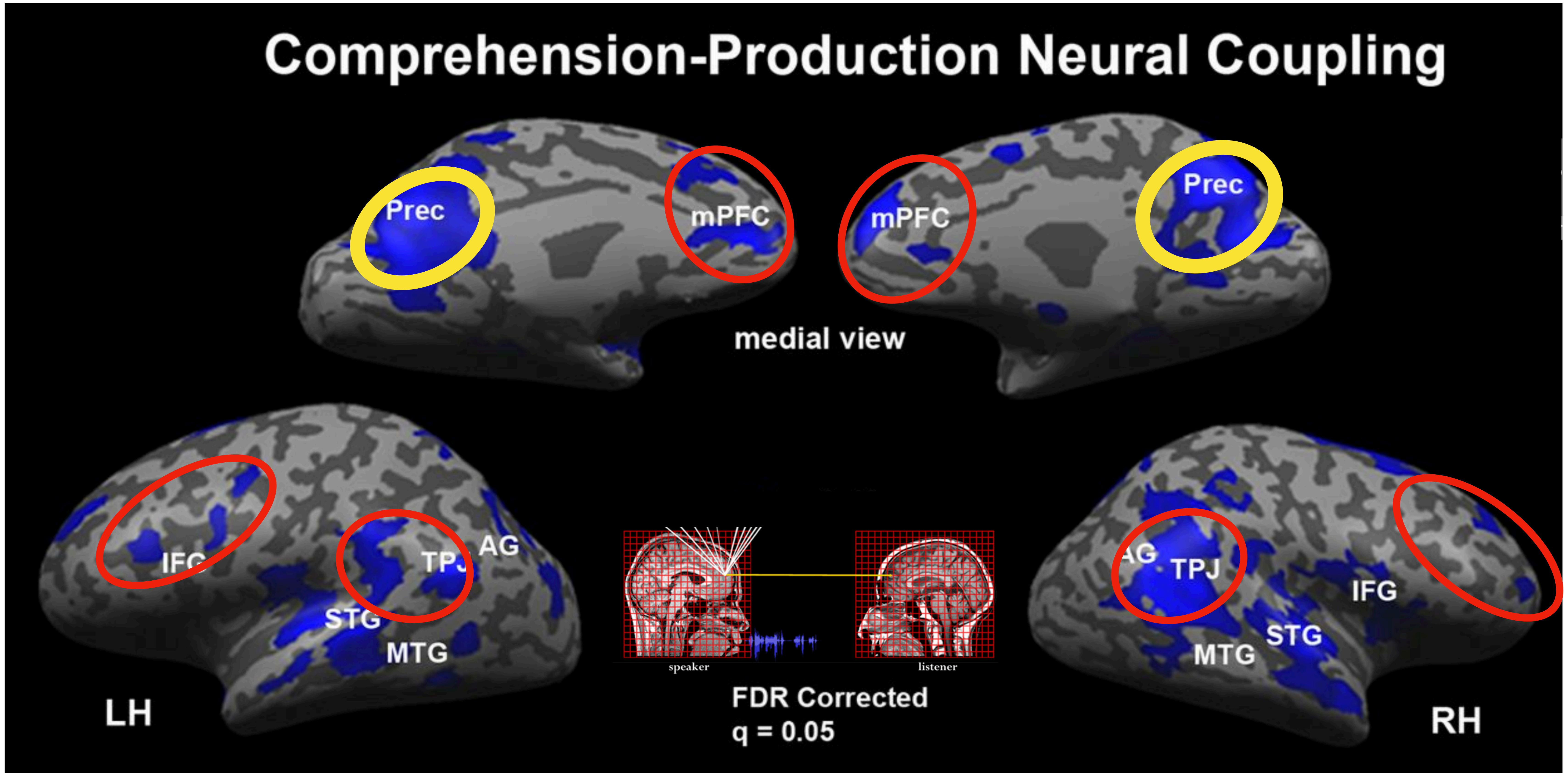
HW #4

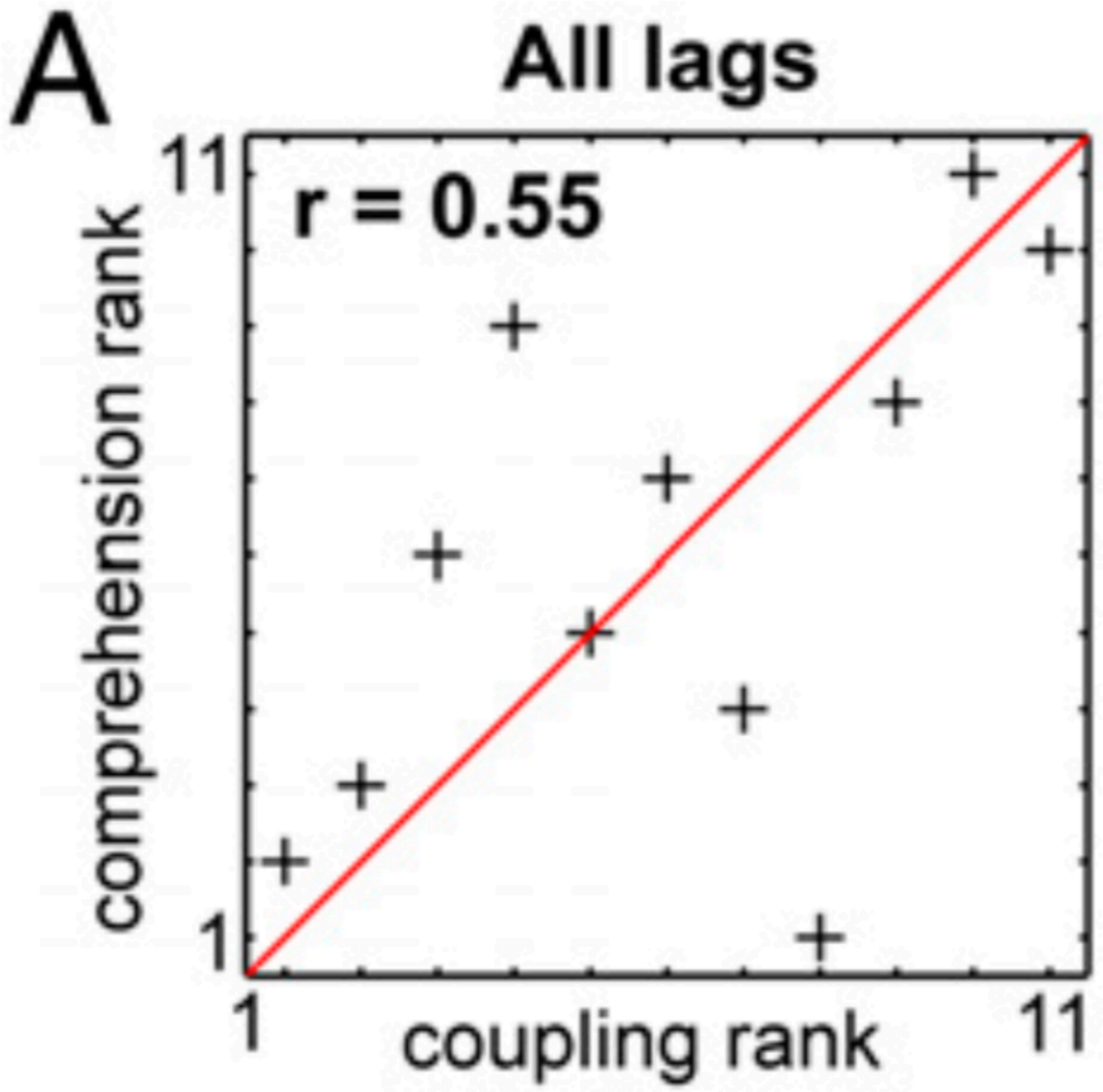


How are individual differences in learning reflected in teacher-driven neural responses?

Nguyen et al., 2020. *bioRxiv*

Neural responses in DMN are coupled between speaker and listener during storytelling

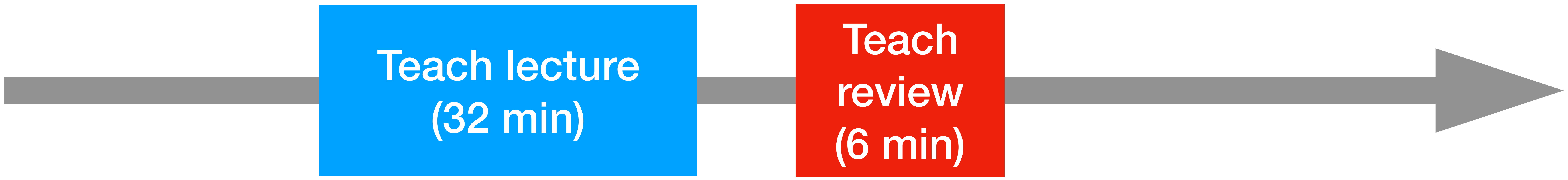




Stephens, Silbert, & Hasson, 2010. *PNAS*
Silbert et al., 2014. *PNAS.*

Design: Scan teacher in fMRI

Teacher
(n=1)



MR safe microphone

Button box

Nguyen et al., 2020. *bioRxiv*

Teach fMRI in fMRI

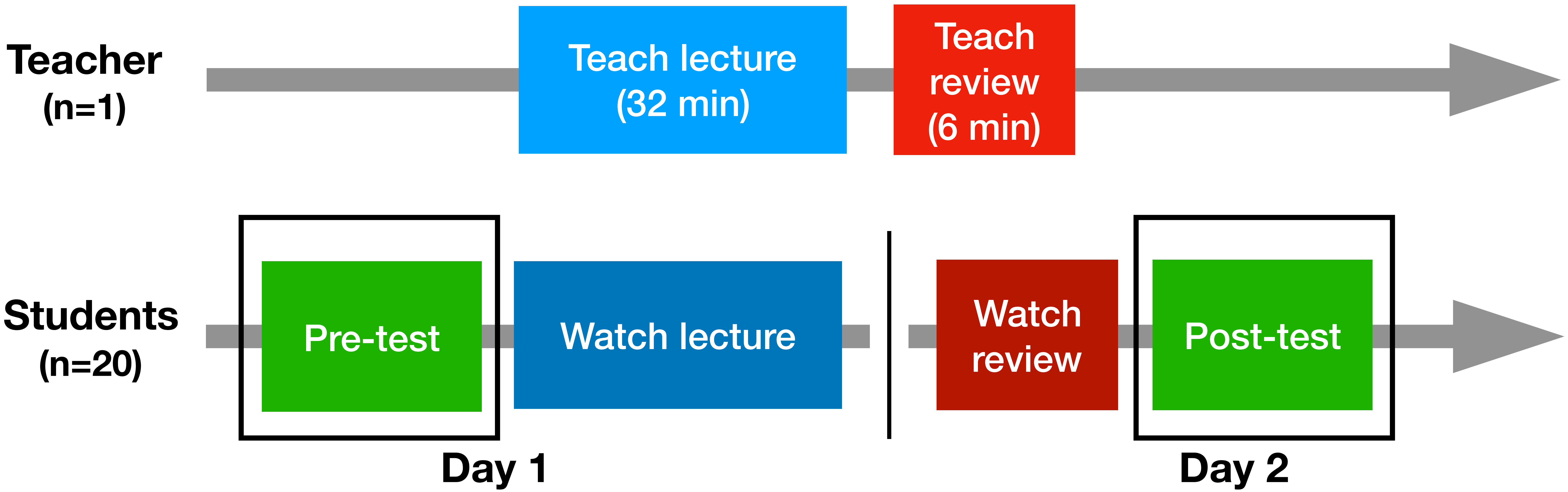
The plan

- FMRI basics
- FMRI scanner
- FMRI physics
- From neurons to fMRI

Introduction to fMRI, ~32 min

Nguyen et al., 2020. *bioRxiv*

Design: scan students in fMRI

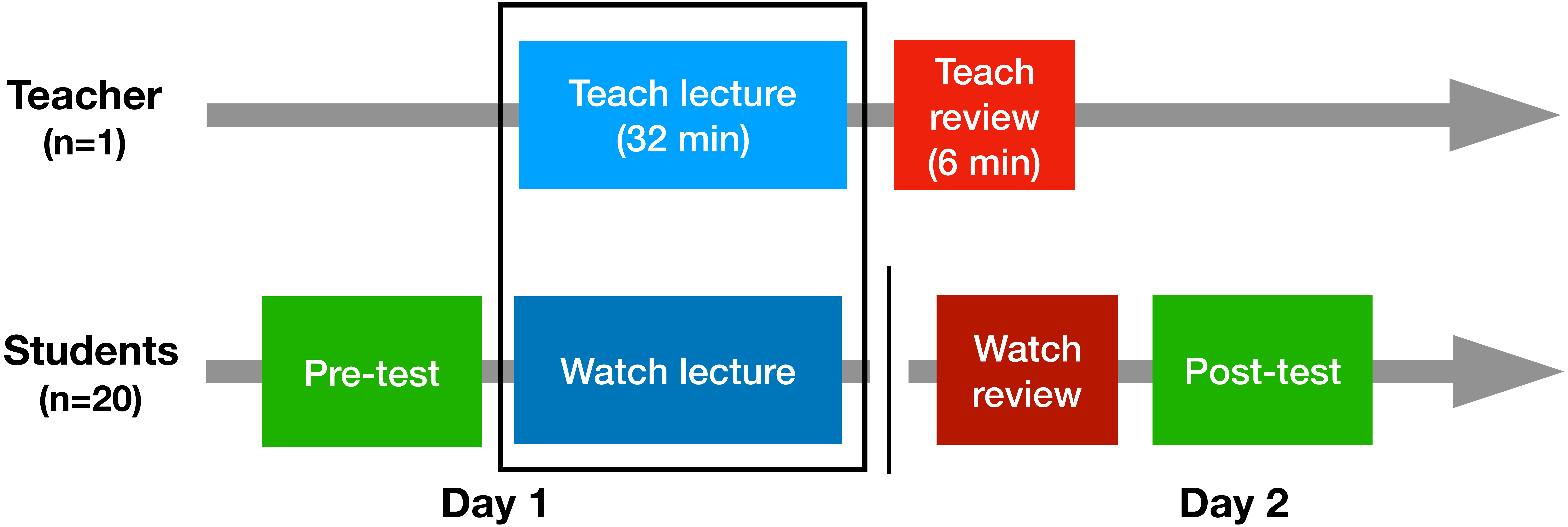


Nguyen et al., 2020. *bioRxiv*

HW #4, Part 1

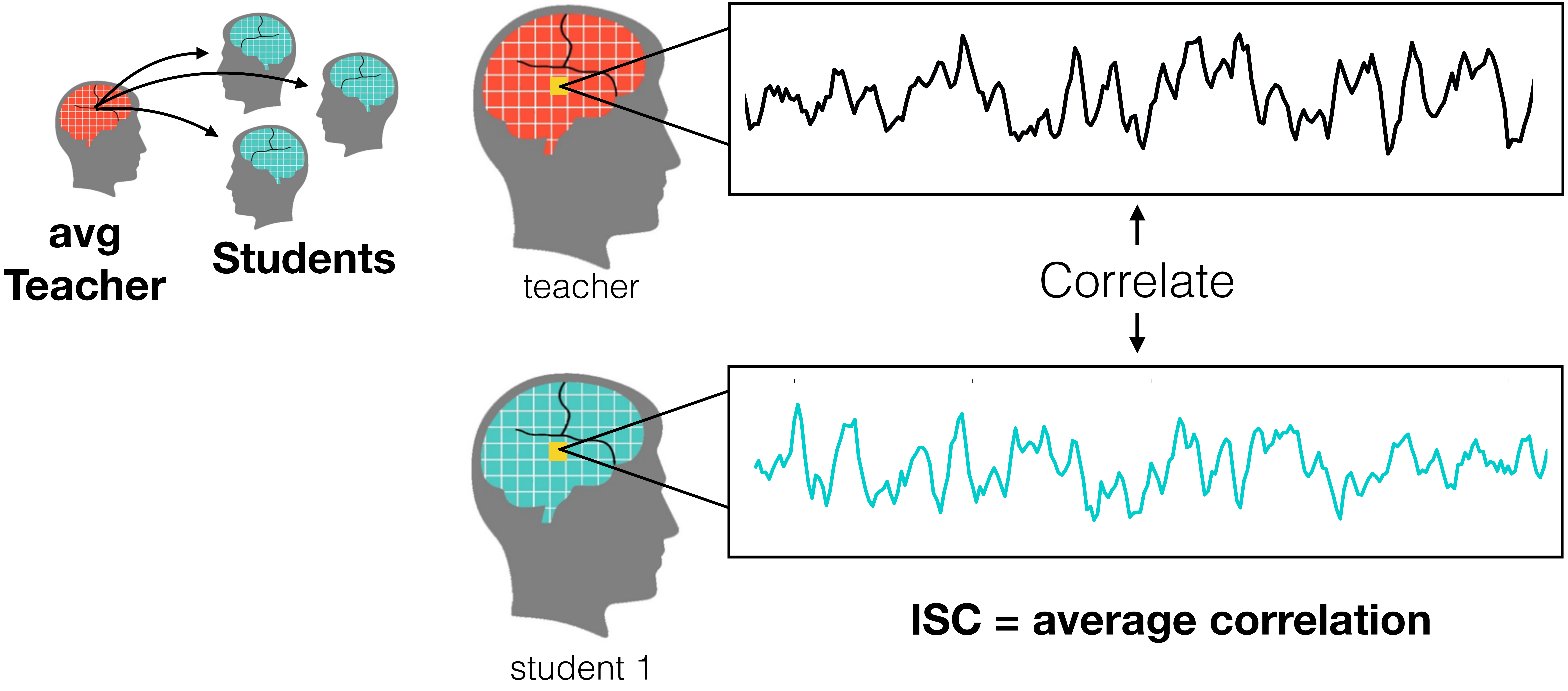
- Compare pre- and post-test scores using the appropriate statistical test
- Make an appropriate visualization
- Describe your results

HW #4, Part 2



Nguyen et al., 2020. *bioRxiv*

HW #4, Part 2: Measure shared neural response between teacher and students using ISC



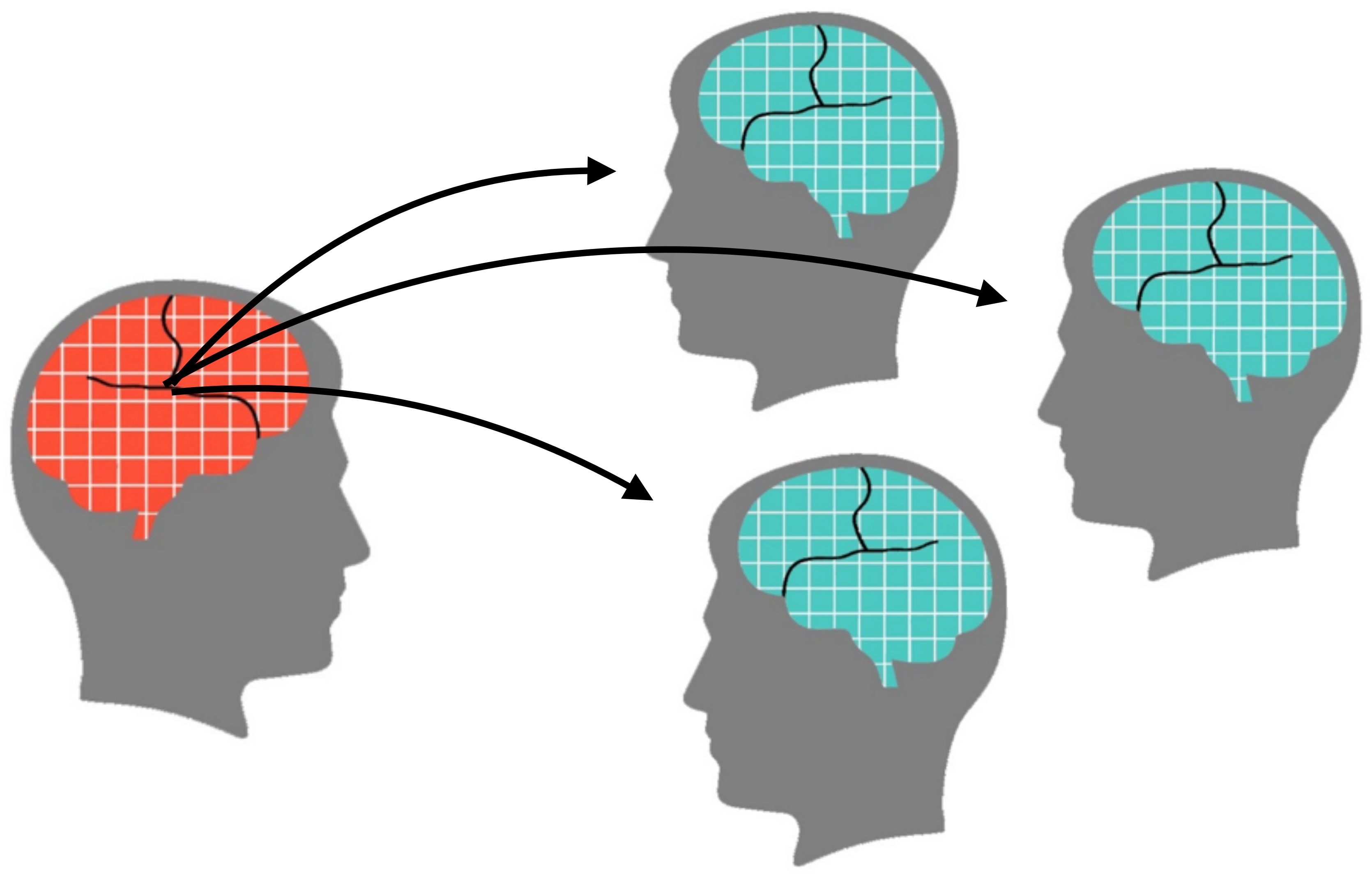
Nguyen et al., 2020. *bioRxiv*

HW #4, Part 2

- Calculate shared neural response between the teacher and student using ISC

- Visualize ISC by plotting the average student time course and the teacher time course on the same figure. Include appropriate legend, labels, and display the ISC value.

HW #4, Part 3



Teacher - Student ISC

Post-test score - **Pre-test score**

Difference learning score

HW #4, Part 3

- Calculate the correlation between Student-Teacher ISC and improvement in quiz score (posttest - pretest)
- Visualize the results with a figure. Be sure to include appropriate labels, titles, etc and include the correlation calculated above
- Write up your results.

HW #4: Bonus

- Repeat Part 3 using the shuffling analysis described in the paper
- Essentially:
    - Randomly shuffle the test scores so that they are out of order (e.g. instead of test scores for [sub1 sub2 sub3 sub4], randomly reorder to something like [sub3 sub1 sub4 sub2]
    - Calculate the correlation between the shuffled test scores and the original ISC scores. Store this “null” correlation in a vector
    - Repeat above two steps 1000x
    - Compare the actual observed correlation to the distribution of null correlations. Calculate p-values by counting the number of null correlations are smaller than the observed correlation and subtracting from 1000.
    - Use this p-value to evaluate significance of the ISC-behavior correlation