

Study Design & Reliability

NYComputational Psychiatry Workshop

Alisa Loosen, PhD

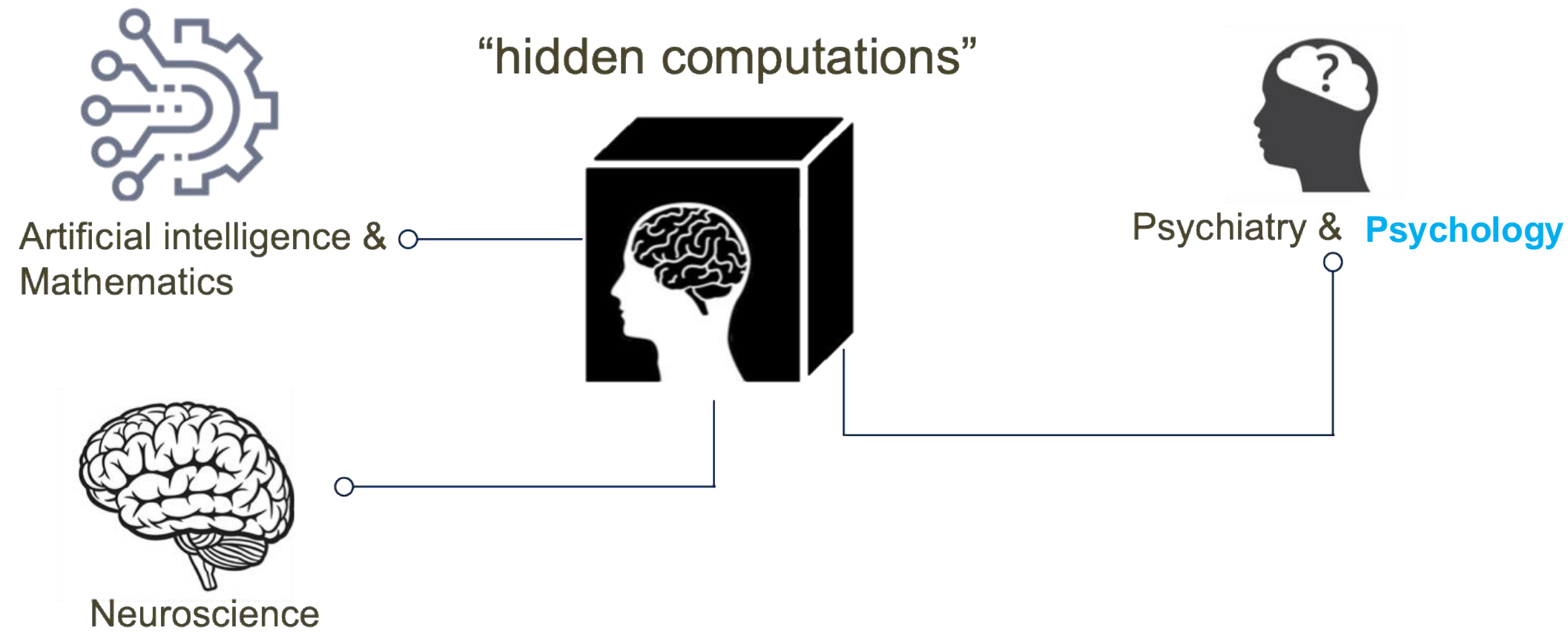
22nd October, 2024



**Mount
Sinai**

*Center for
Computational
Psychiatry*

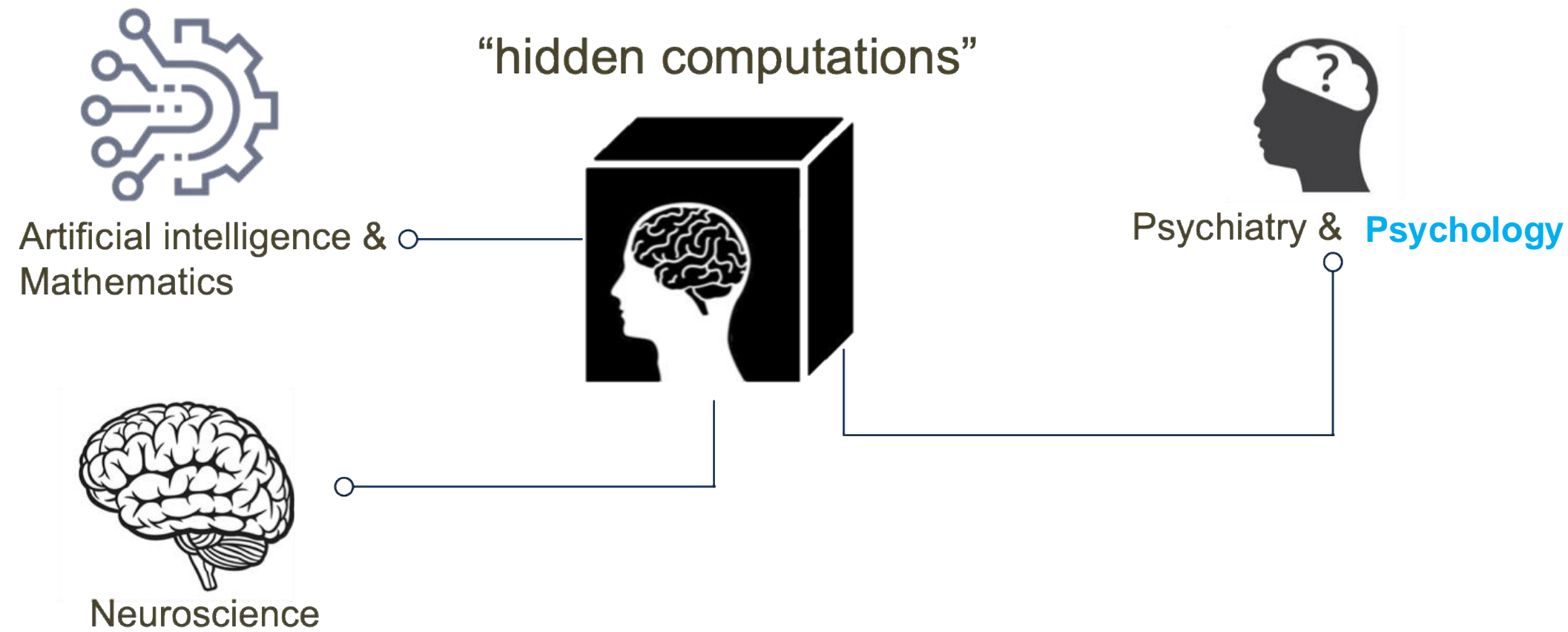
Computational Psychiatry & Safety Checks in Science



Open Science Collaboration, 2015, *Science*

How can we test and ensure the robustness of our measures?

Computational Psychiatry & Safety Checks in Science



Open Science Collaboration, 2015, *Science*

Control mechanisms

Open-science practices

Pre-registration

Preventative mechanisms

Careful study design

Assessment of validity

Assessment of reliability

Overview

Study design

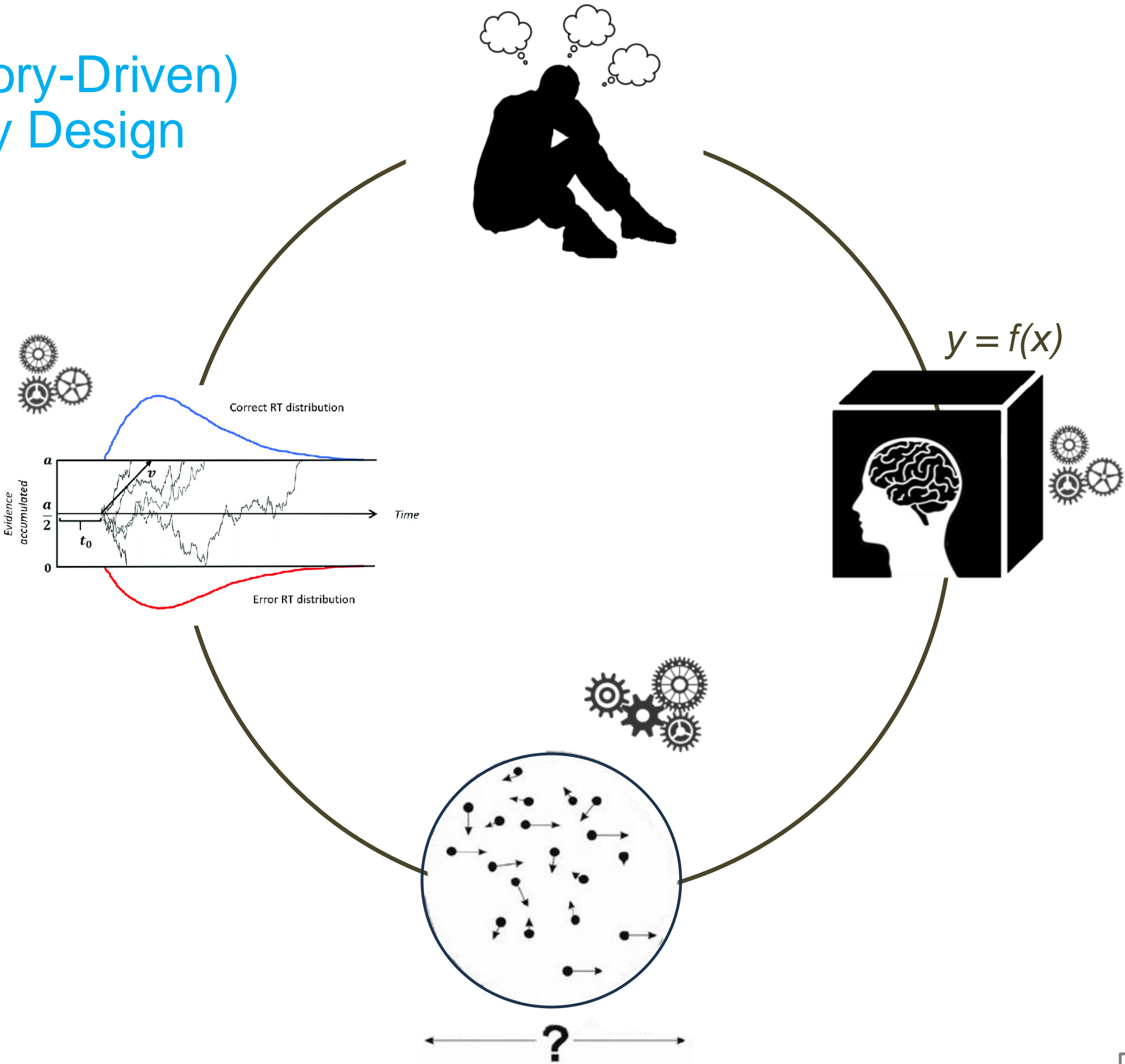
Validity assessments

Reliability assessments

Example study: assessment of a widely-used decision-making task

Dos and Don'ts

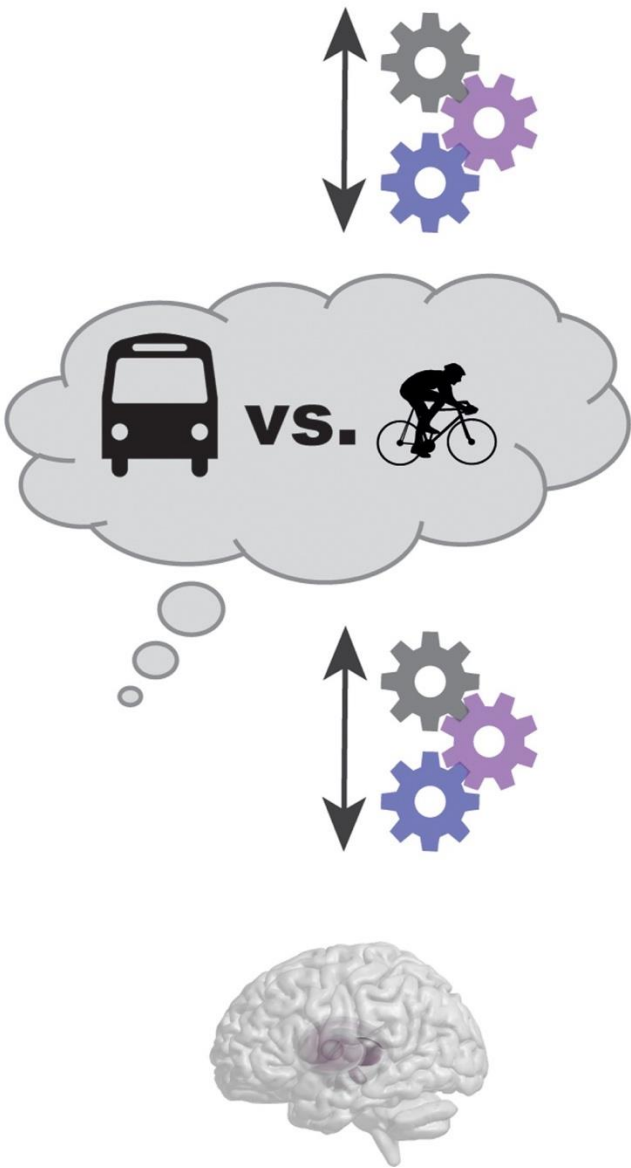
(Theory-Driven)
Study Design



Symptoms



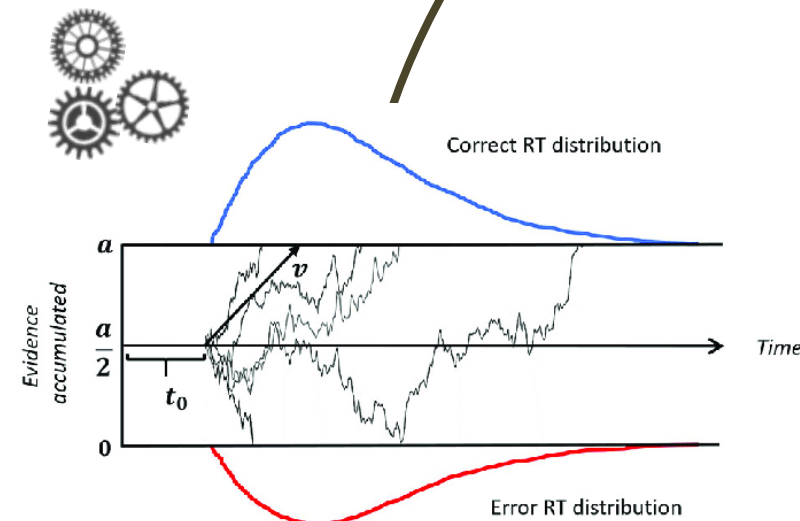
Cognition



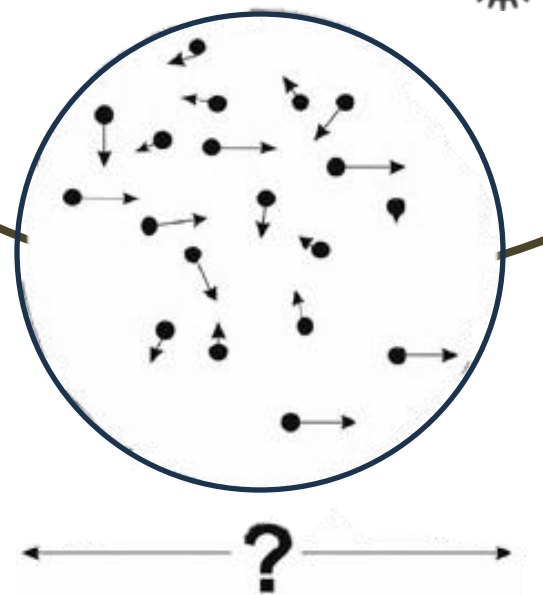
Brain

(Theory-Driven) Study Design

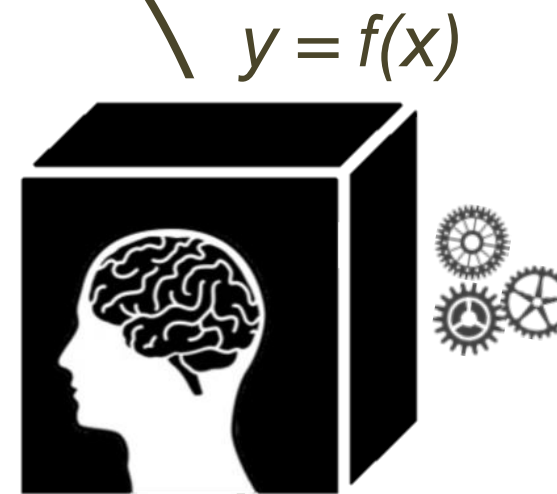
Can we answer
our questions?



- Simulate
- Do not disregard simple analyses



- Base study on clinical symptoms and behaviors
- Target specific cognitive processes affected



- Generate precise, testable questions
- Aim for clear mechanistic, neuroscientifically and clinically informed hypotheses

Design Experiment

Build Models

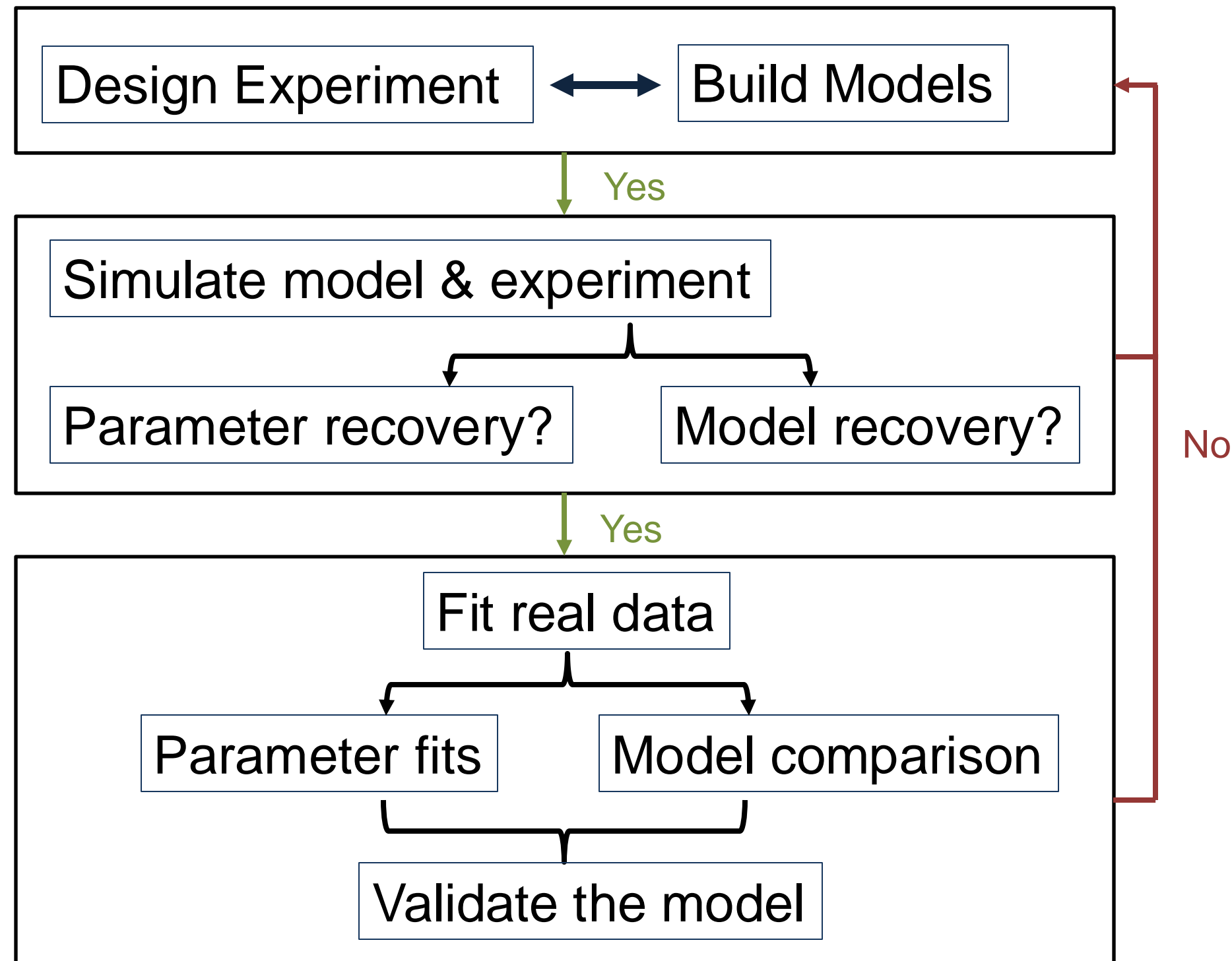


- Use appropriate tasks for specific hypotheses tailored to your model
- Think about identifiability of parameters

(Theory-Driven) Study Design

Can we answer
our questions?

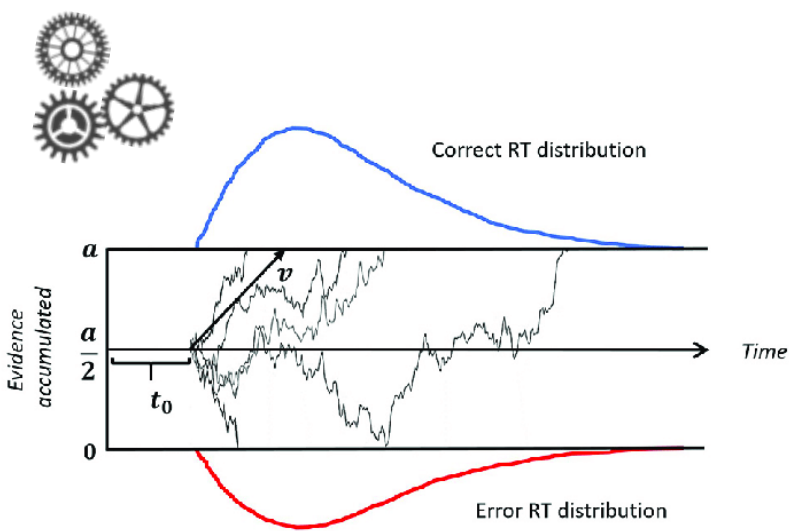
Can the model
account for our
data?



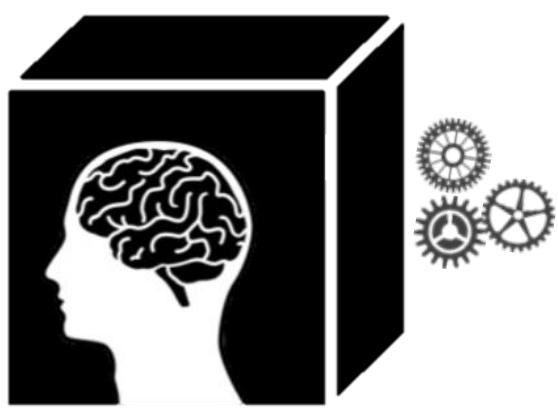
Measurement Quality



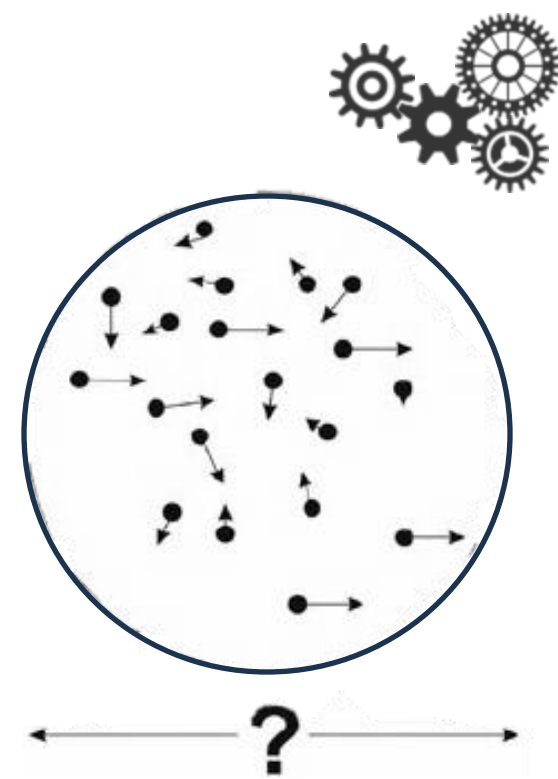
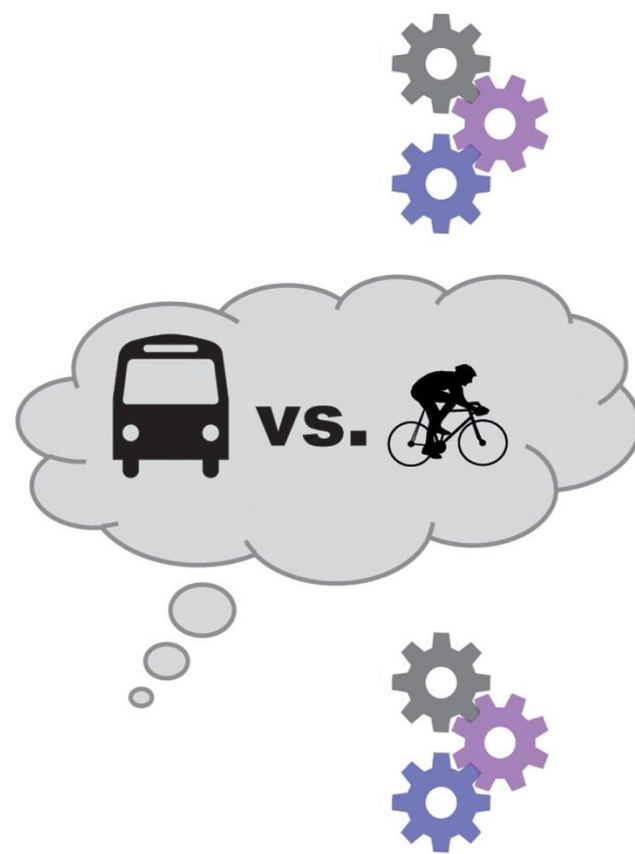
Symptoms



$$y = f(x)$$



Cognition



Brain

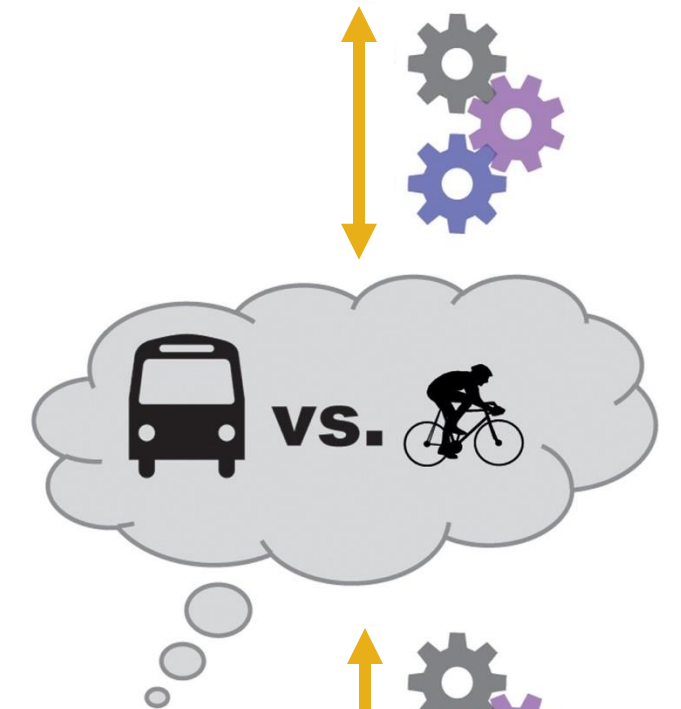


Measurement Quality

Symptoms



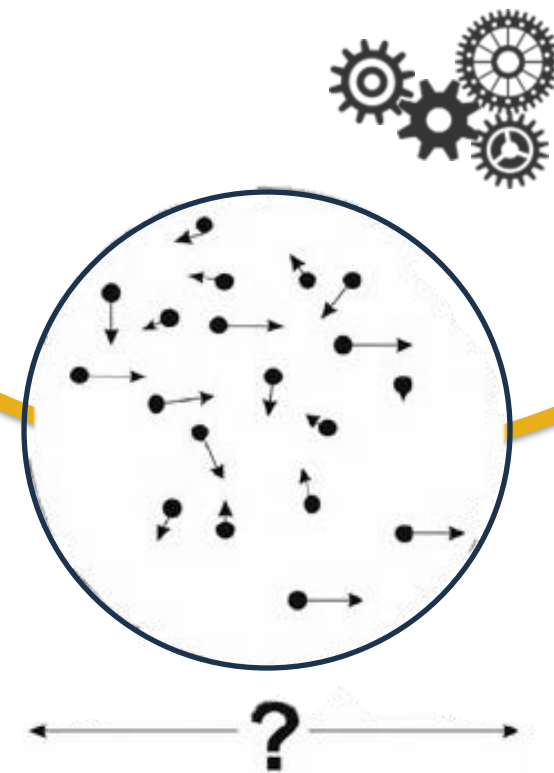
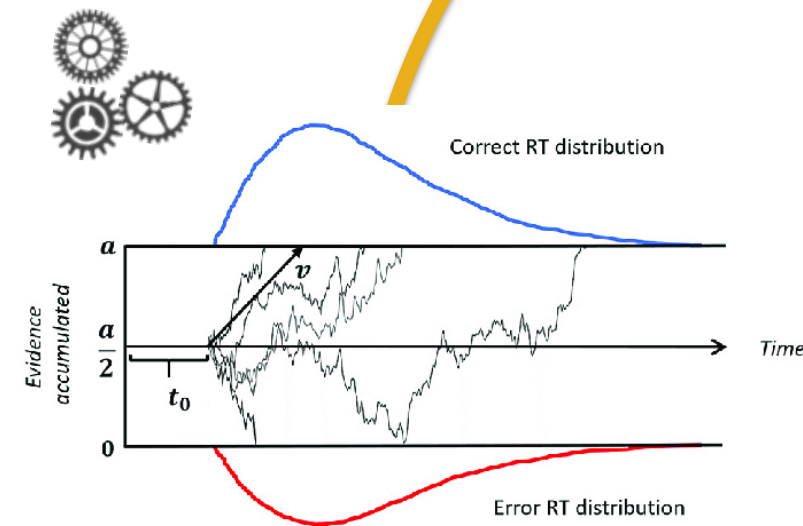
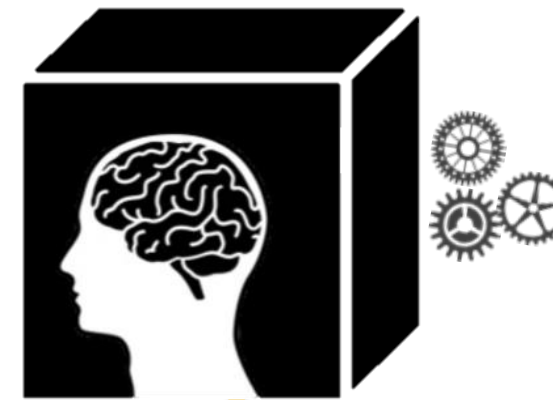
Cognition



Brain



$$y = f(x)$$



Overview

Study design

Validity assessments

Reliability assessments

Validity

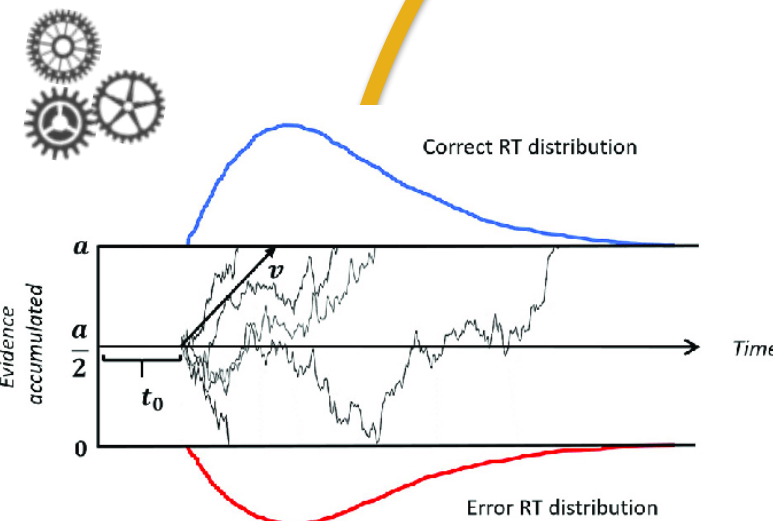
Construct Validity

Do our measures accurately represent the process we want to measure?

$$y = f(x)$$

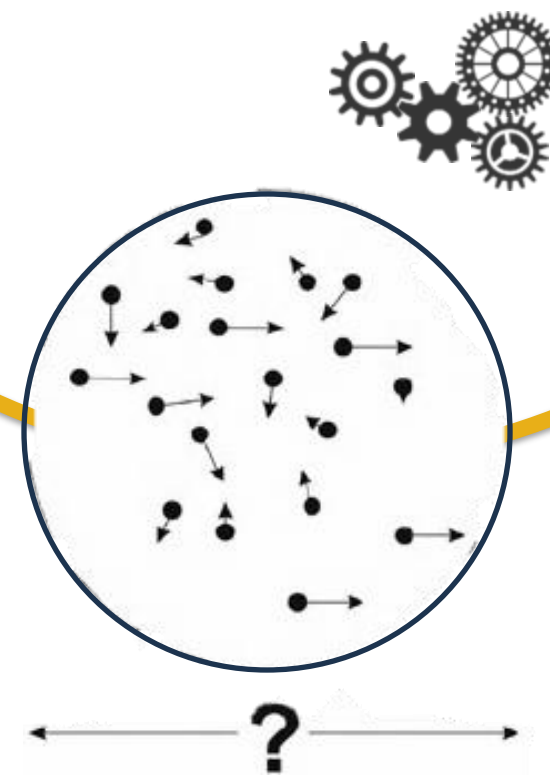
Neuroscientific & Clinical Relevance

Are we capturing neural/ behavioral systems that reflect the psychiatric symptom or process?



Ecological Validity

Does the study accurately simulate real-life scenarios?



Psychometric Quality Assessments - Validity

Validity

- *Does the task measure what it is supposed to measure?*
- *The accuracy of a measure in capturing the intended construct.*

Reliability

- *Does the measure yield consistent results over time?*
- *The consistency of a measure across different time frames.*

Psychometric Quality Assessments - Reliability

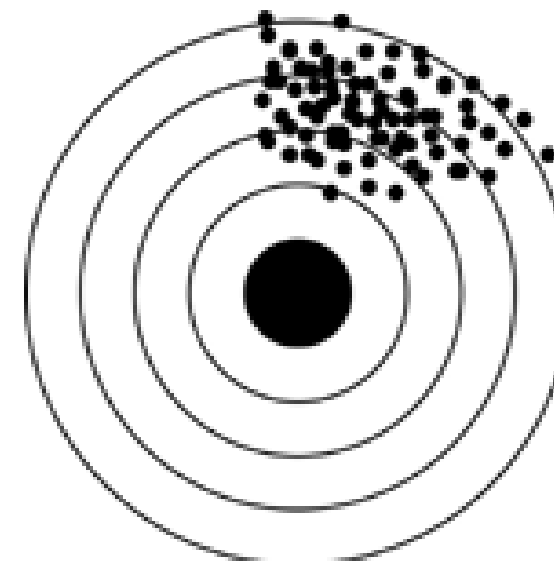
Internal consistency

- The degree to which items within a test consistently measure the same underlying construct.
- *Commonly assessed in questionnaires (e.g., using Cronbach's alpha).*

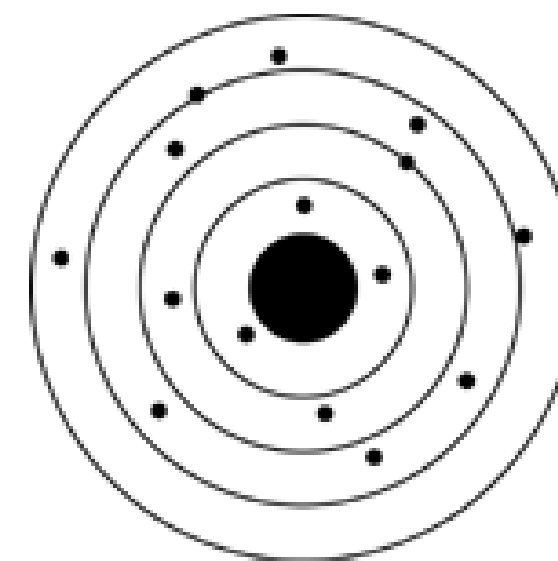


Test-retest reliability

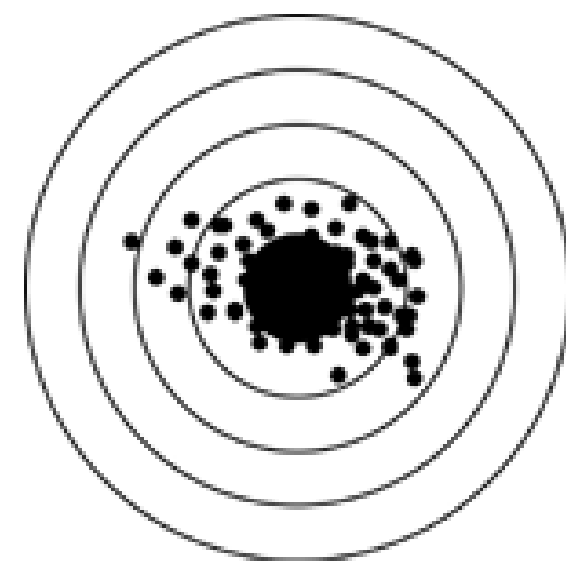
- The stability of test results when the same test is administered to the same individuals under similar conditions at different times.
- *Ensures consistency over time.*



Reliable but Not Valid



Valid but Not Reliable



Valid and Reliable

Overview

Study design

Validity assessments

Reliability assessments

Example study: assessment of a widely-used decision-making task

What might be challenging when assessing the psychometric properties of task measures?

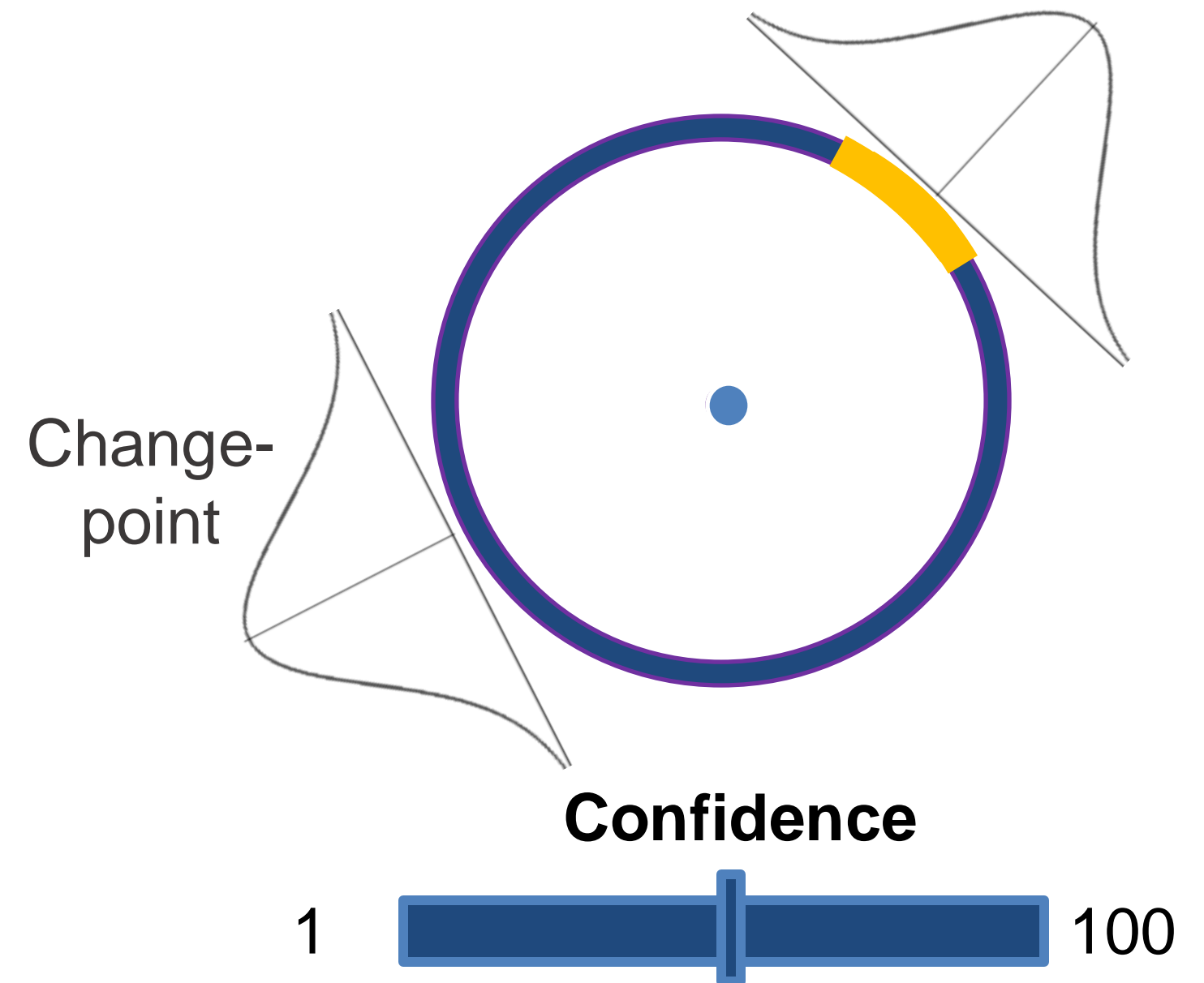
Example: The Predictive-Inference (“Helicopter”) Task

Confidence & Action-update alterations in an uncertain, changing environment

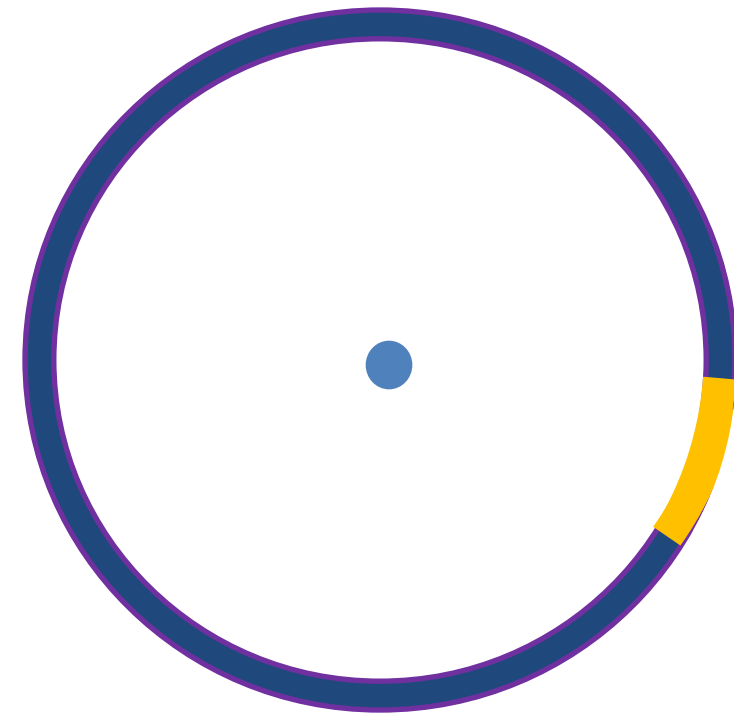
Obsessive-compulsive disorder

Schizophrenia

Transdiagnostic investigations



e.g., Vaghi et al., 2017; Seow & Gillan, 2020; Nassar et al., 2010, 2012, 2016, 2019, 2020; McGuire et al., 2014; Bruckner et al., 2020; Jepma et al., 2016



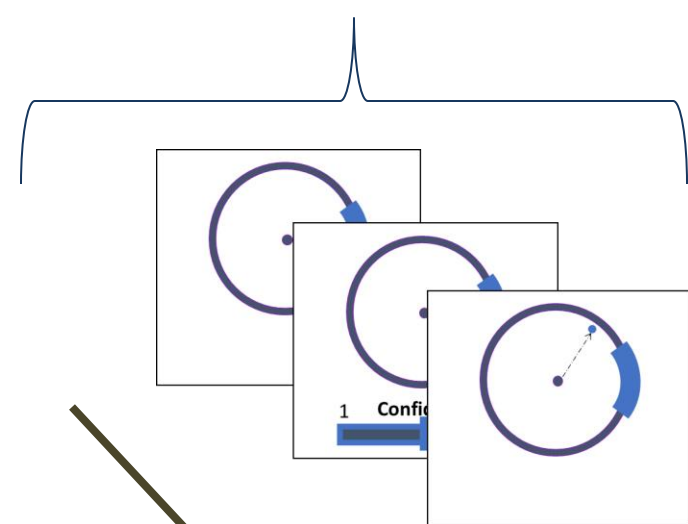
UK Public (Prolific)

Time point 1: $N = 329$
Time point 2: $N = 219$ } ~3 months in between

Reliability Measures

Internal consistency

Measures task part 1

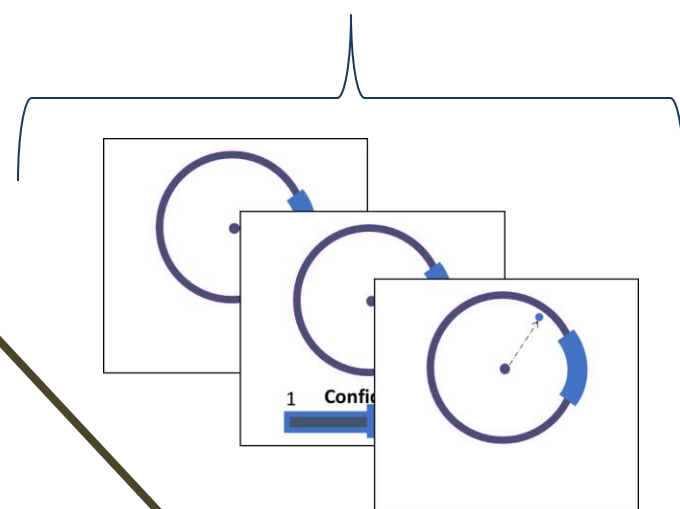


=

Measures task part 2

...

Trials

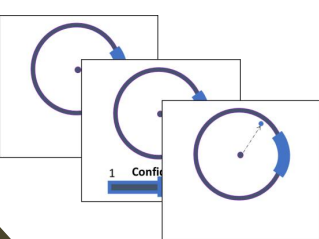
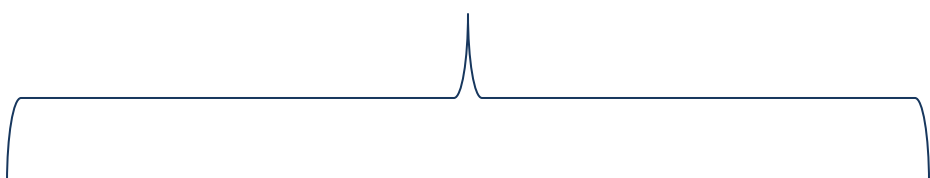
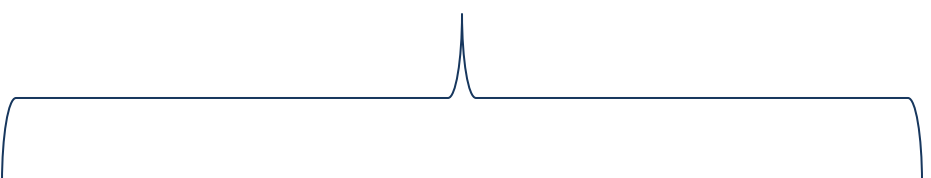


Test-retest reliability

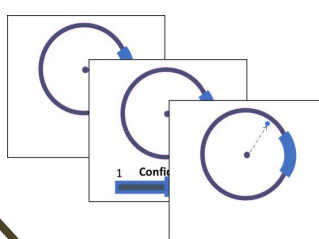
Measures time point 1

=

Measures time point 2



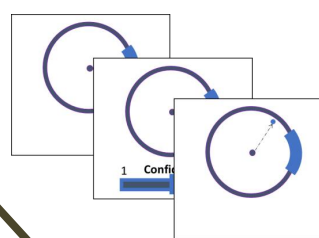
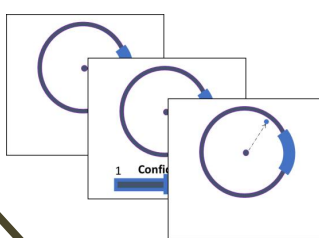
...



...

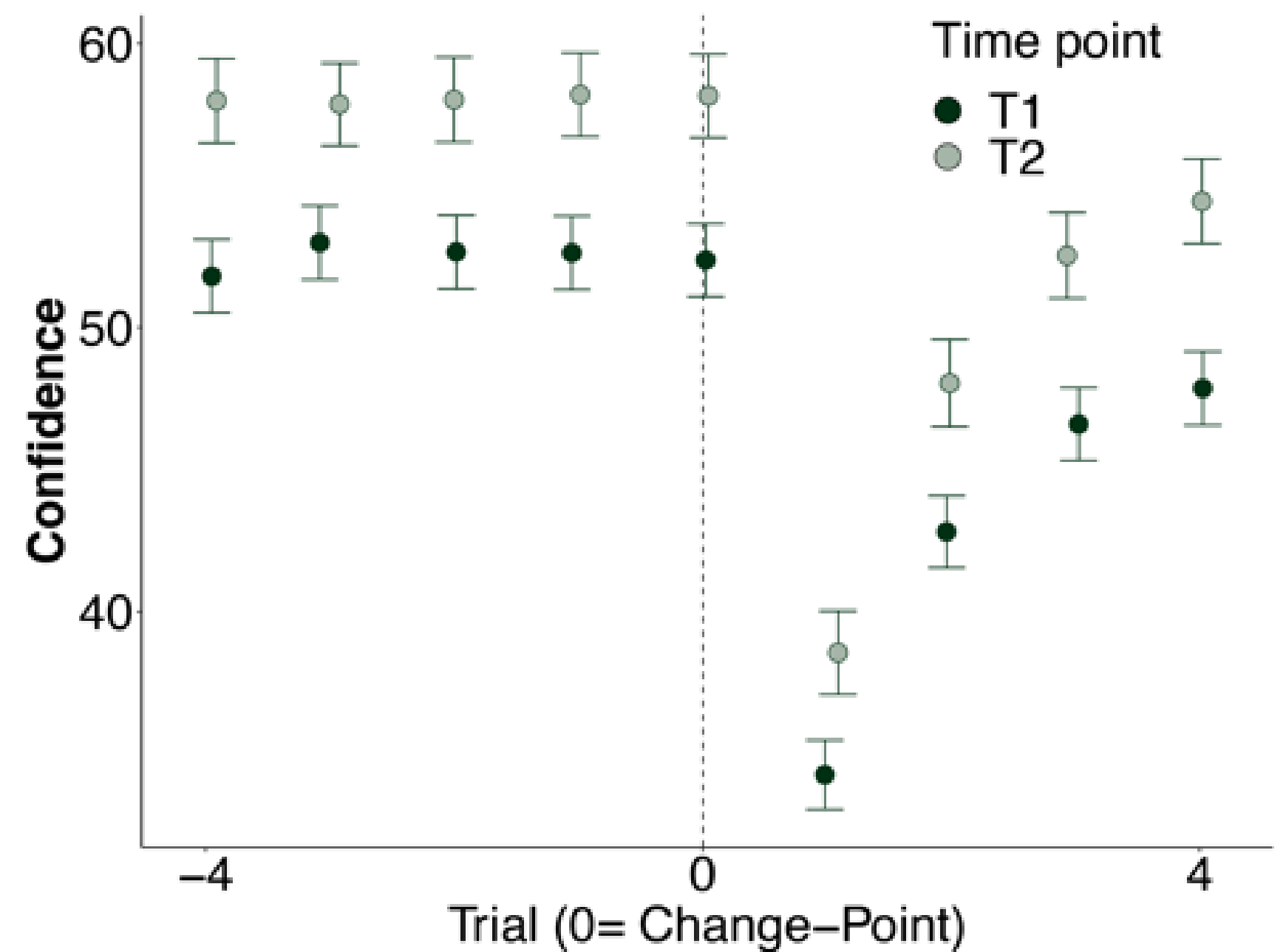
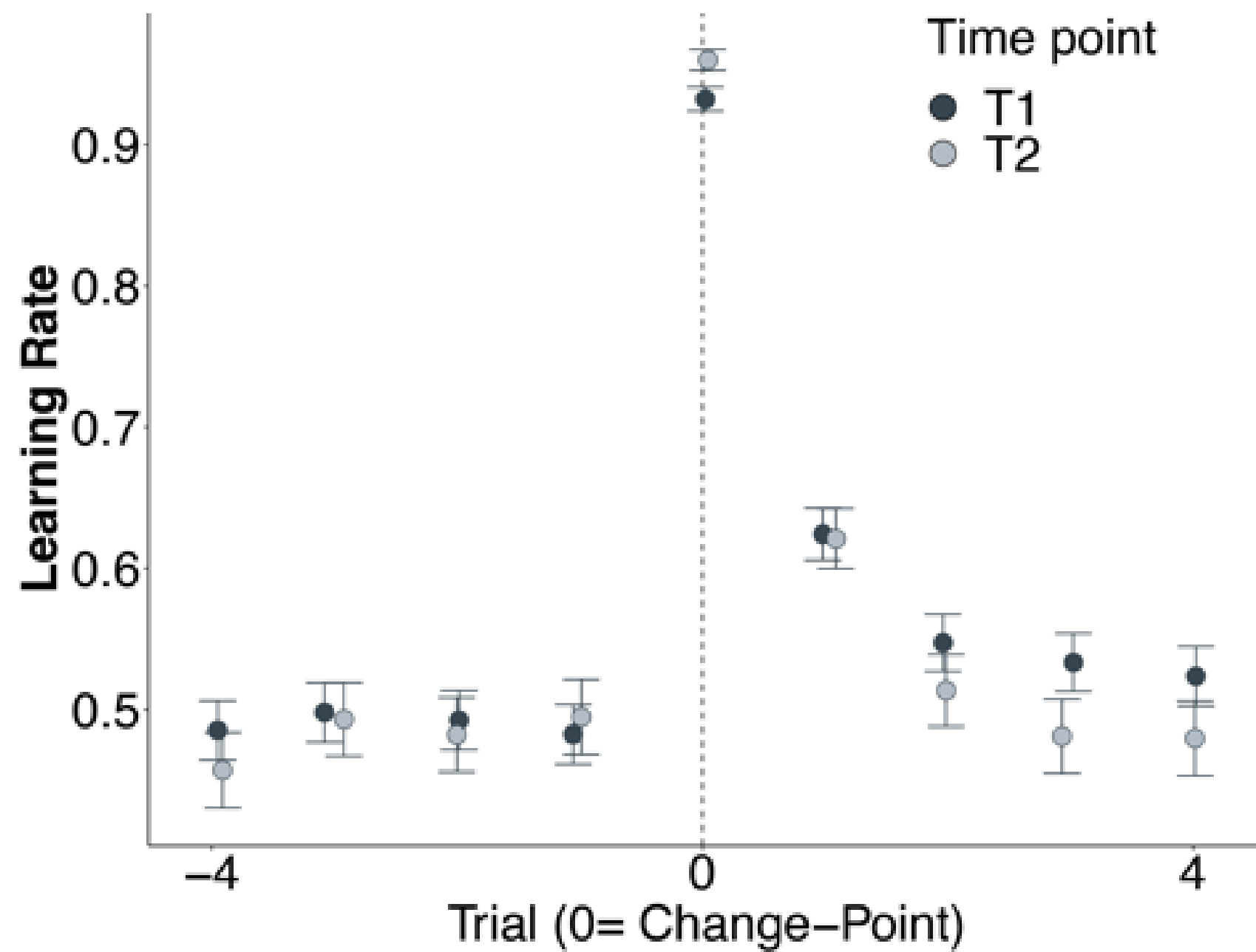
Trials

Trials



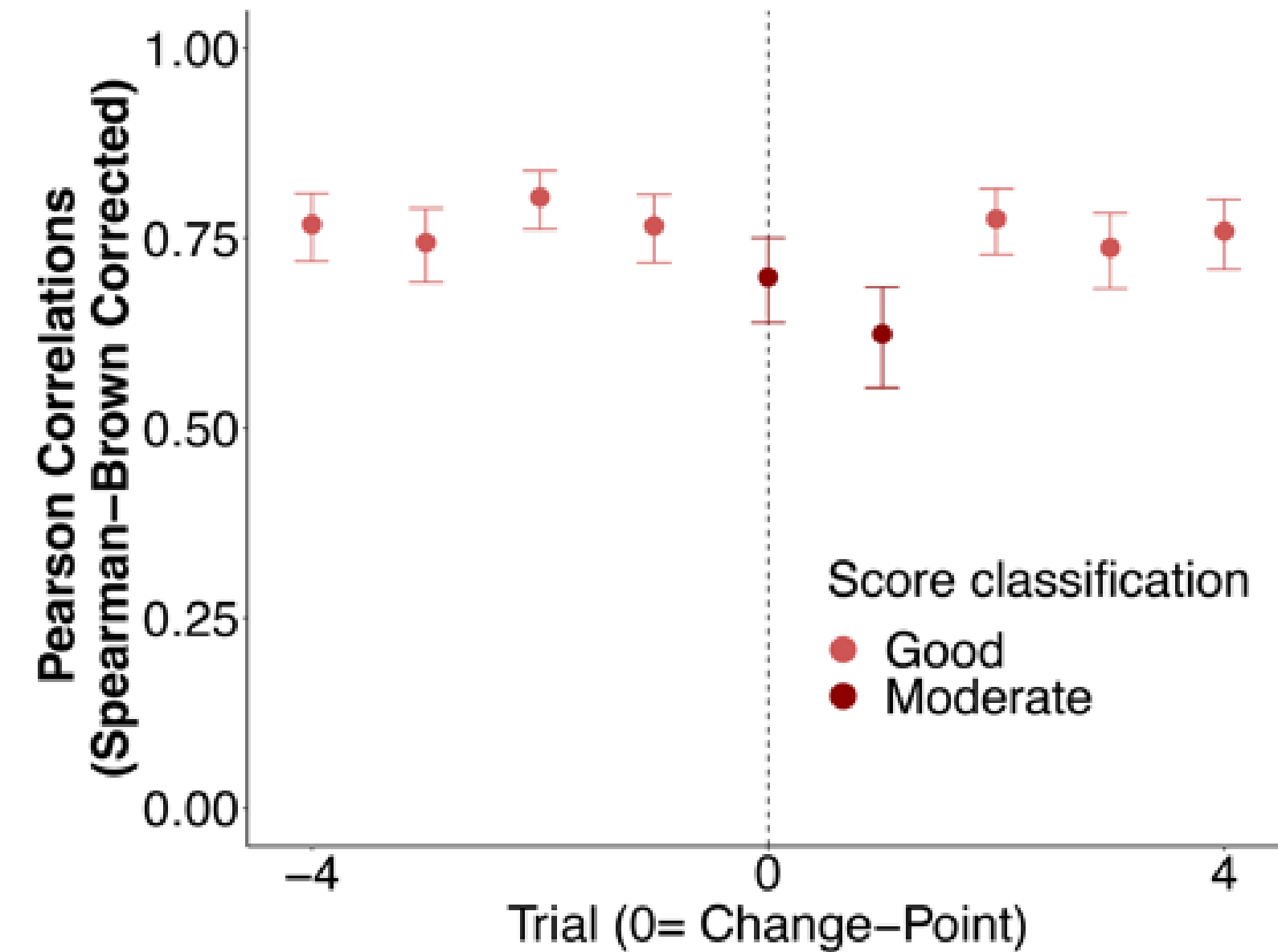
~3 Months

Raw Behavioral Measures

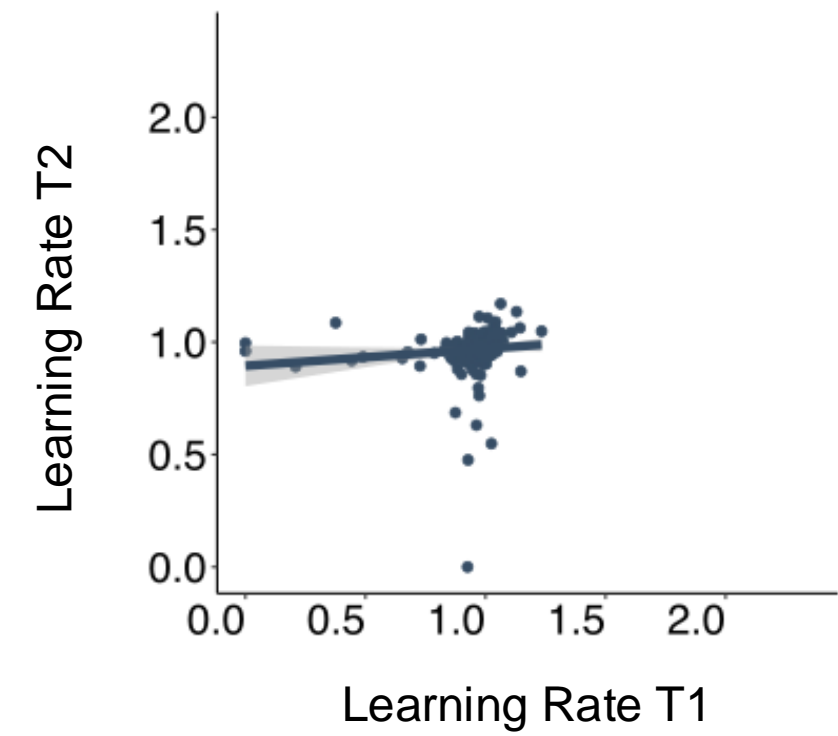
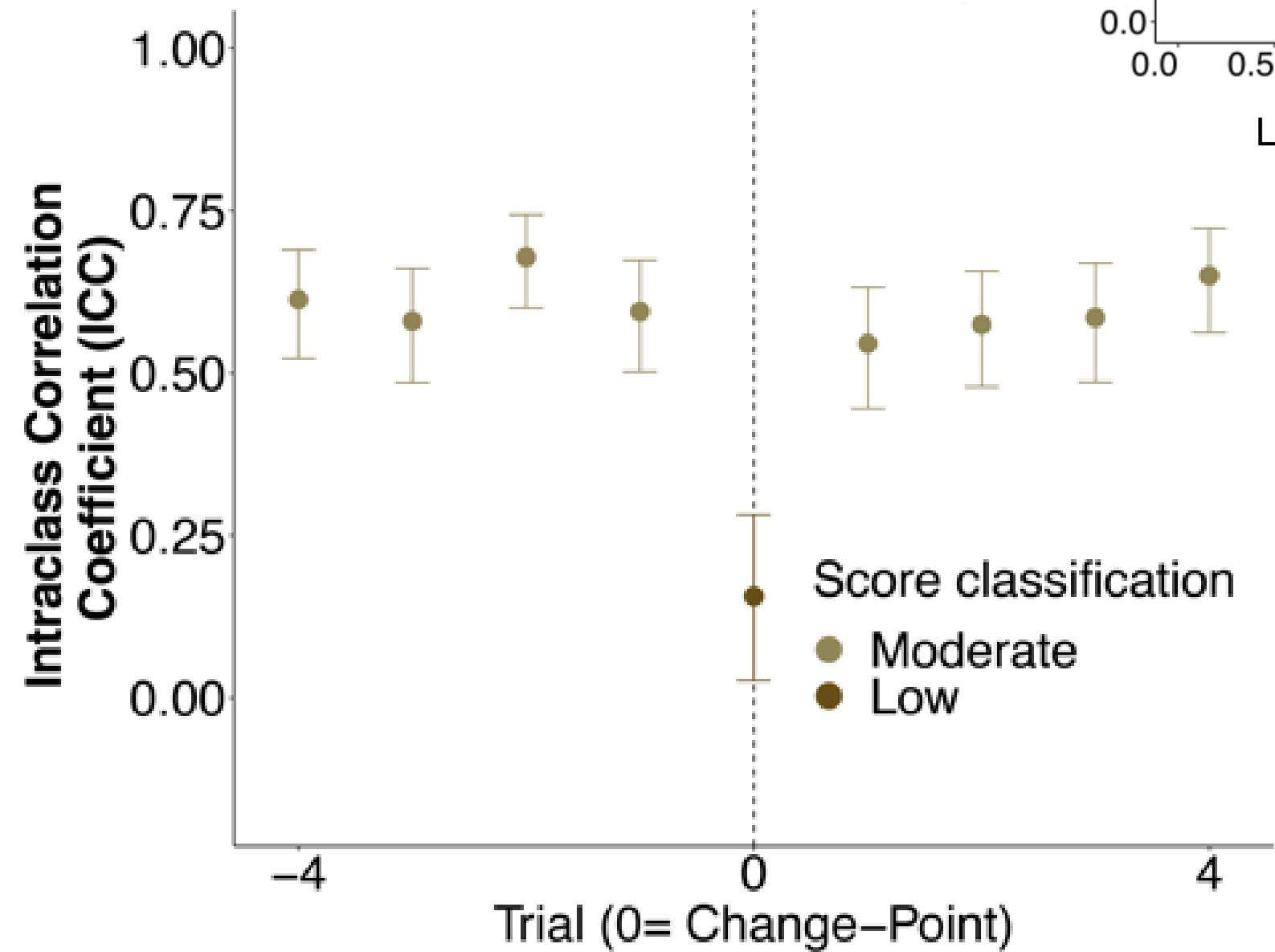


Raw Behavioral Measures – Learning Rate

Internal Consistency

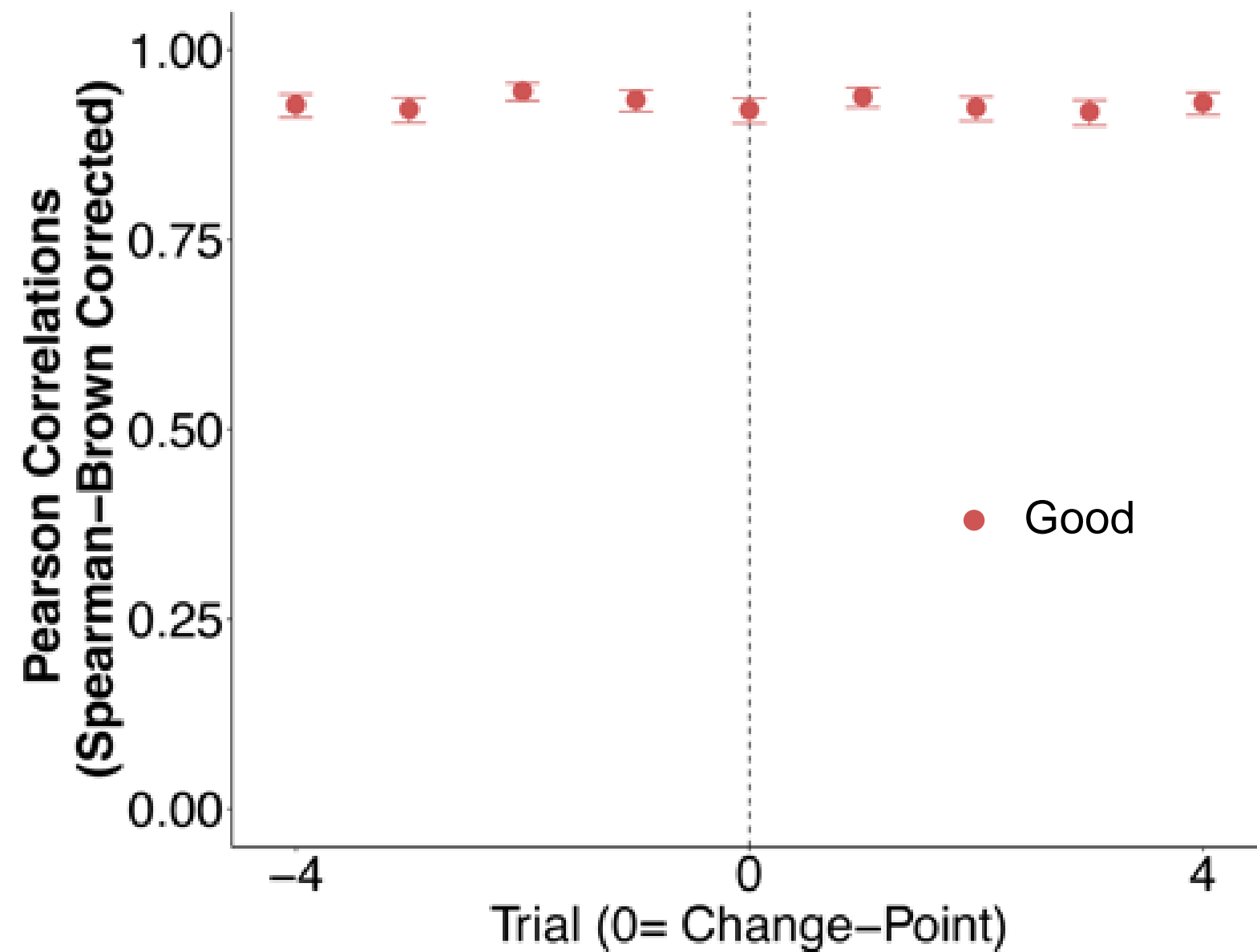


Test-retest reliability

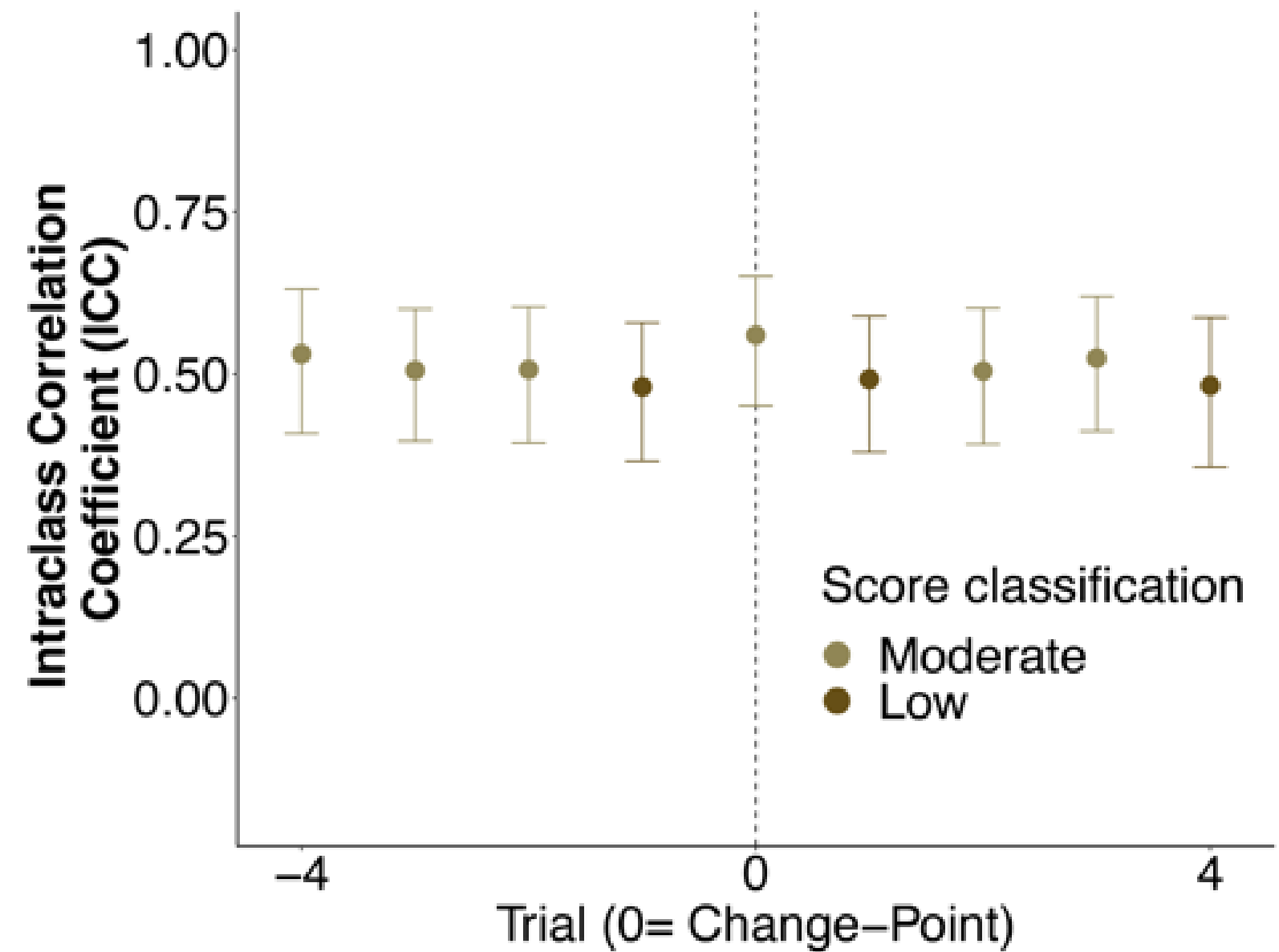


Raw Behavioral Measures - Confidence

Internal Consistency



Test-retest reliability



Complex Measures - Bayesian Learner

Bayesian Learner's

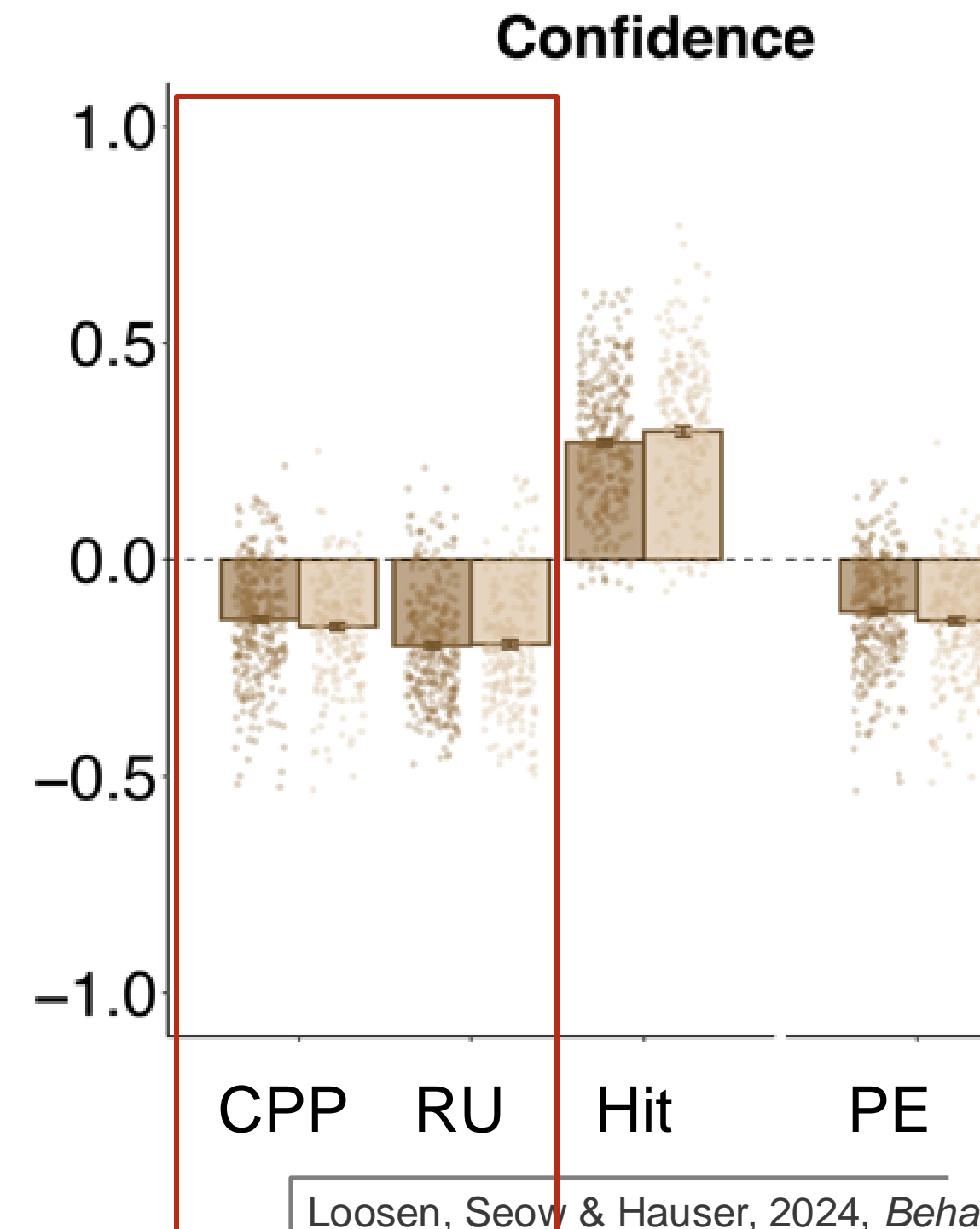
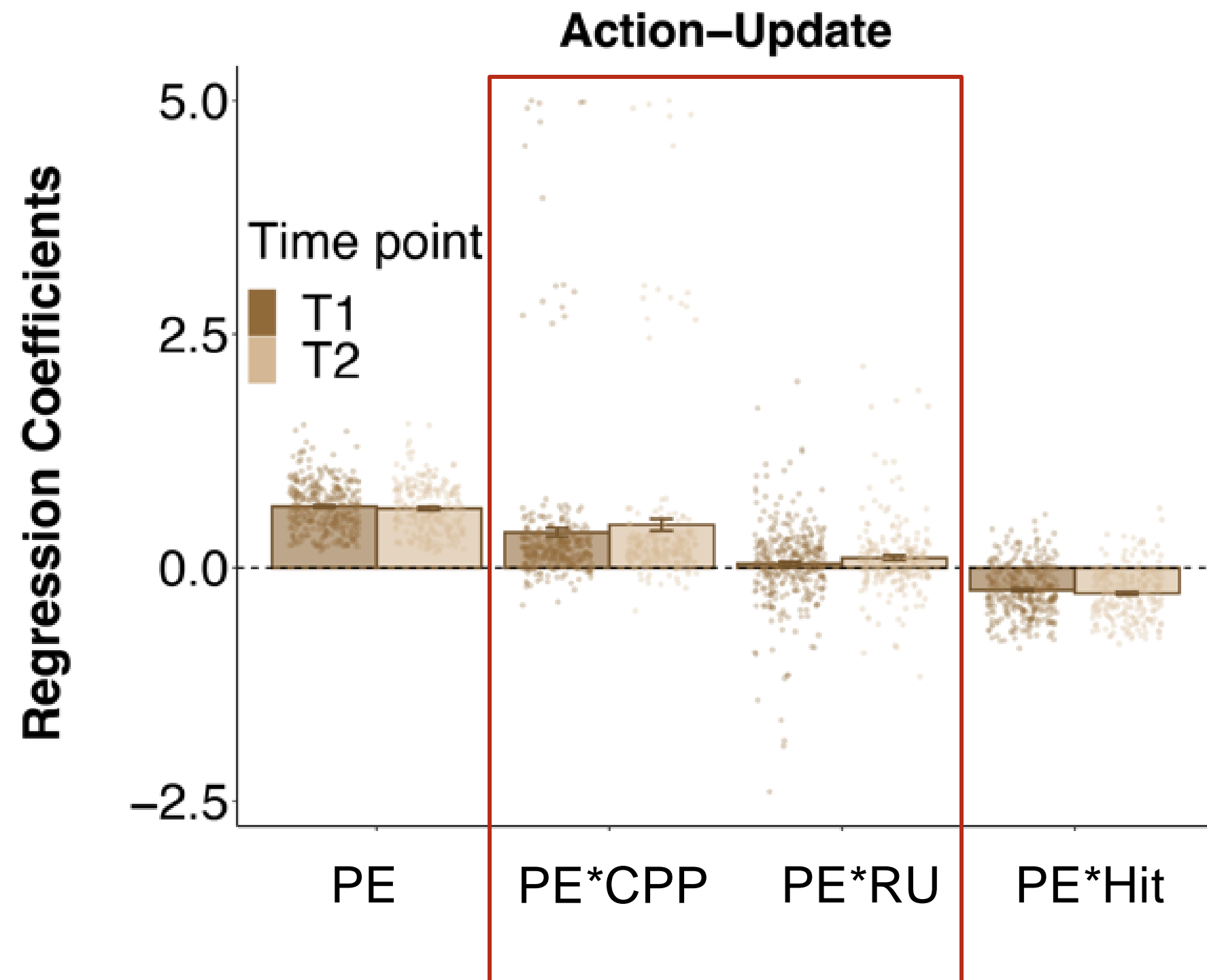
1. Change-point probability,

an approximation of the probability that a change-point has occurred

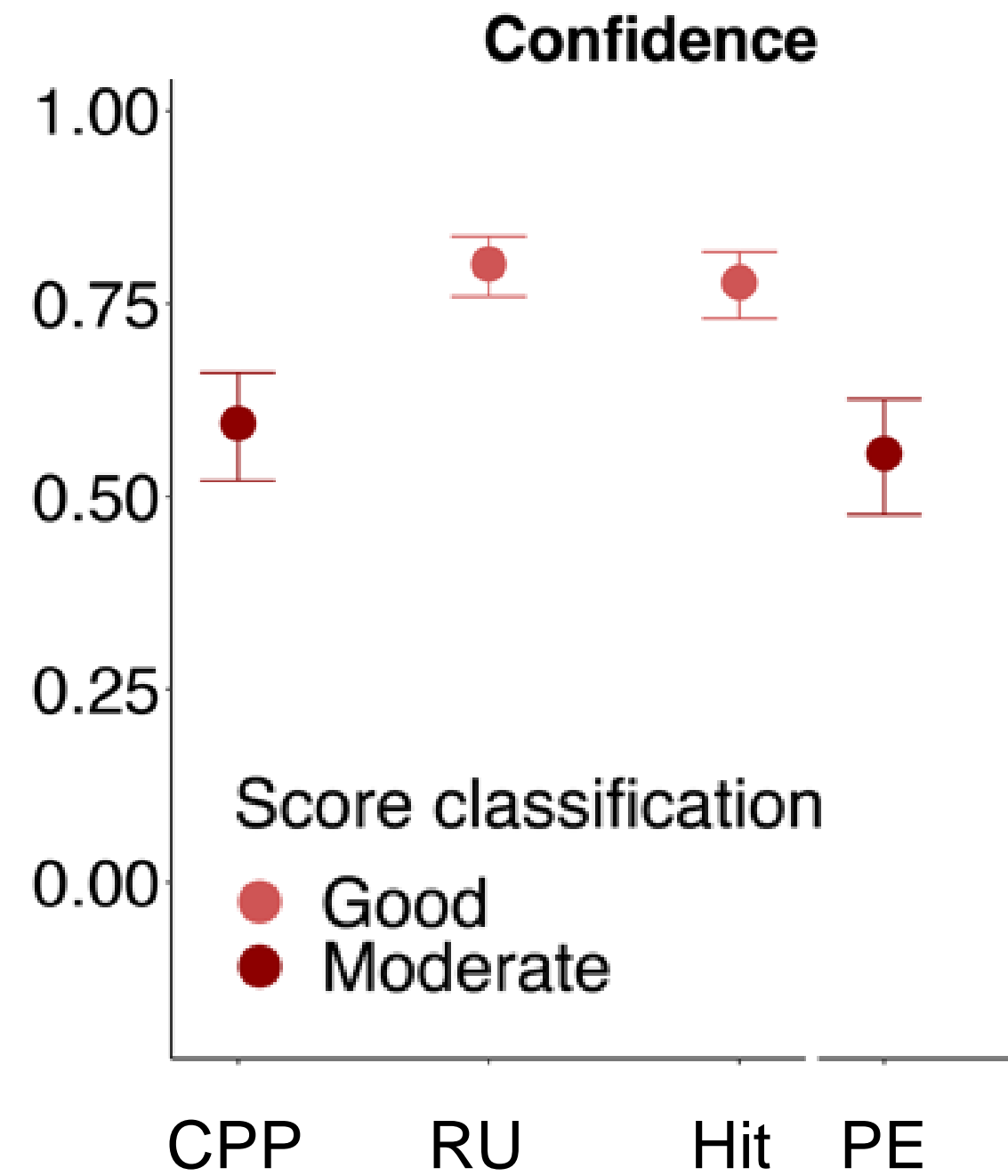
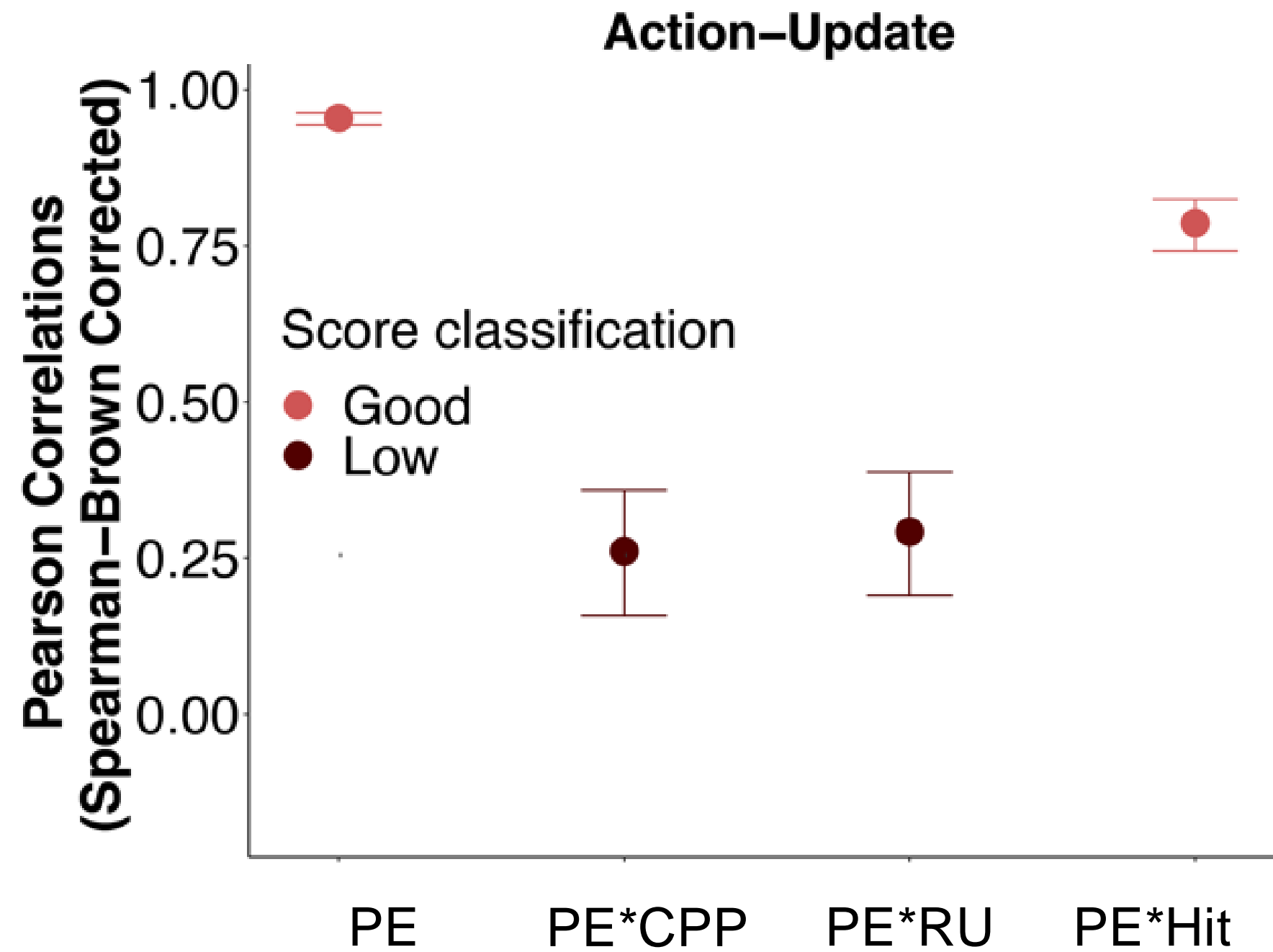
2. Relative uncertainty

in the belief about the mean of the distribution determining the particle landing location

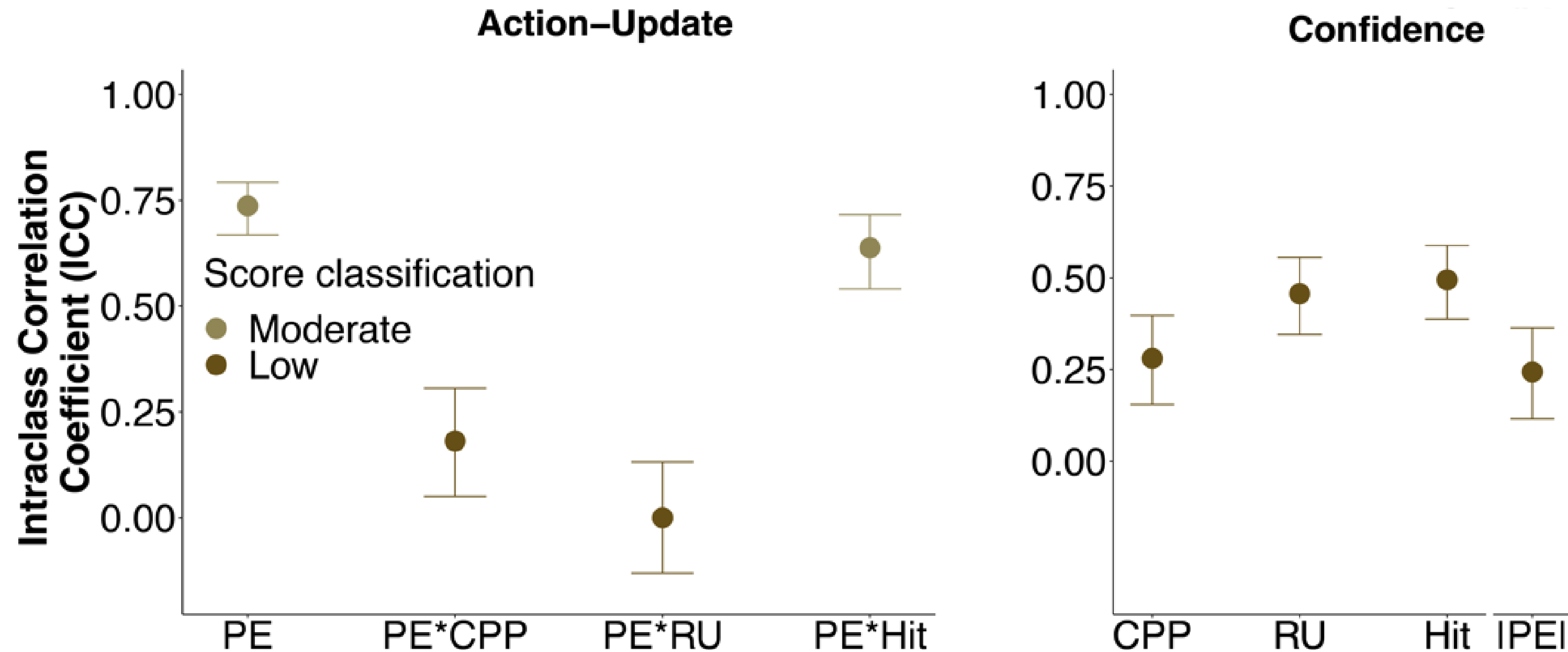
Example: The Predictive-Inference (“Helicopter”) Task



Internal Consistency – Bayesian Learner & Behavior

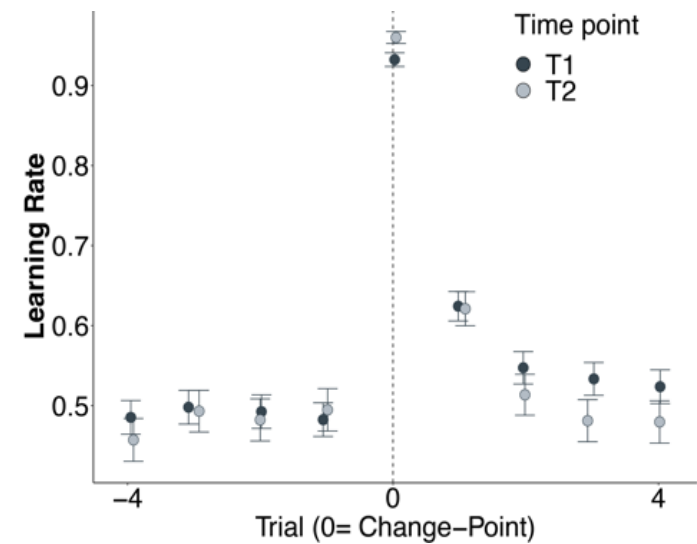


Test-Retest Reliability – Bayesian Learner & Behavior

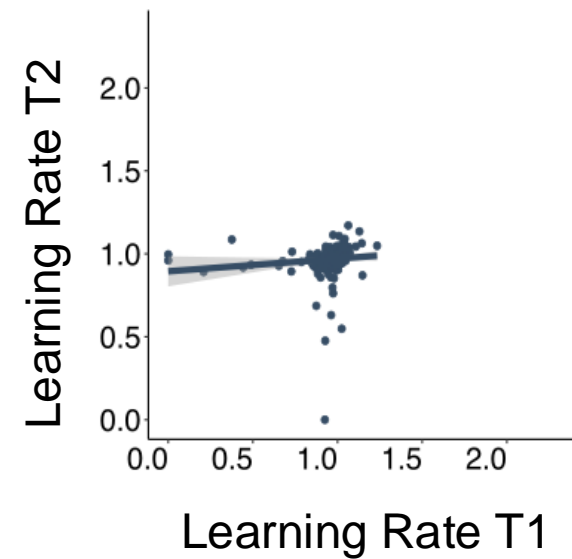


Example: The Predictive-Inference (“Helicopter”) Task

Focus on points in the task are important.

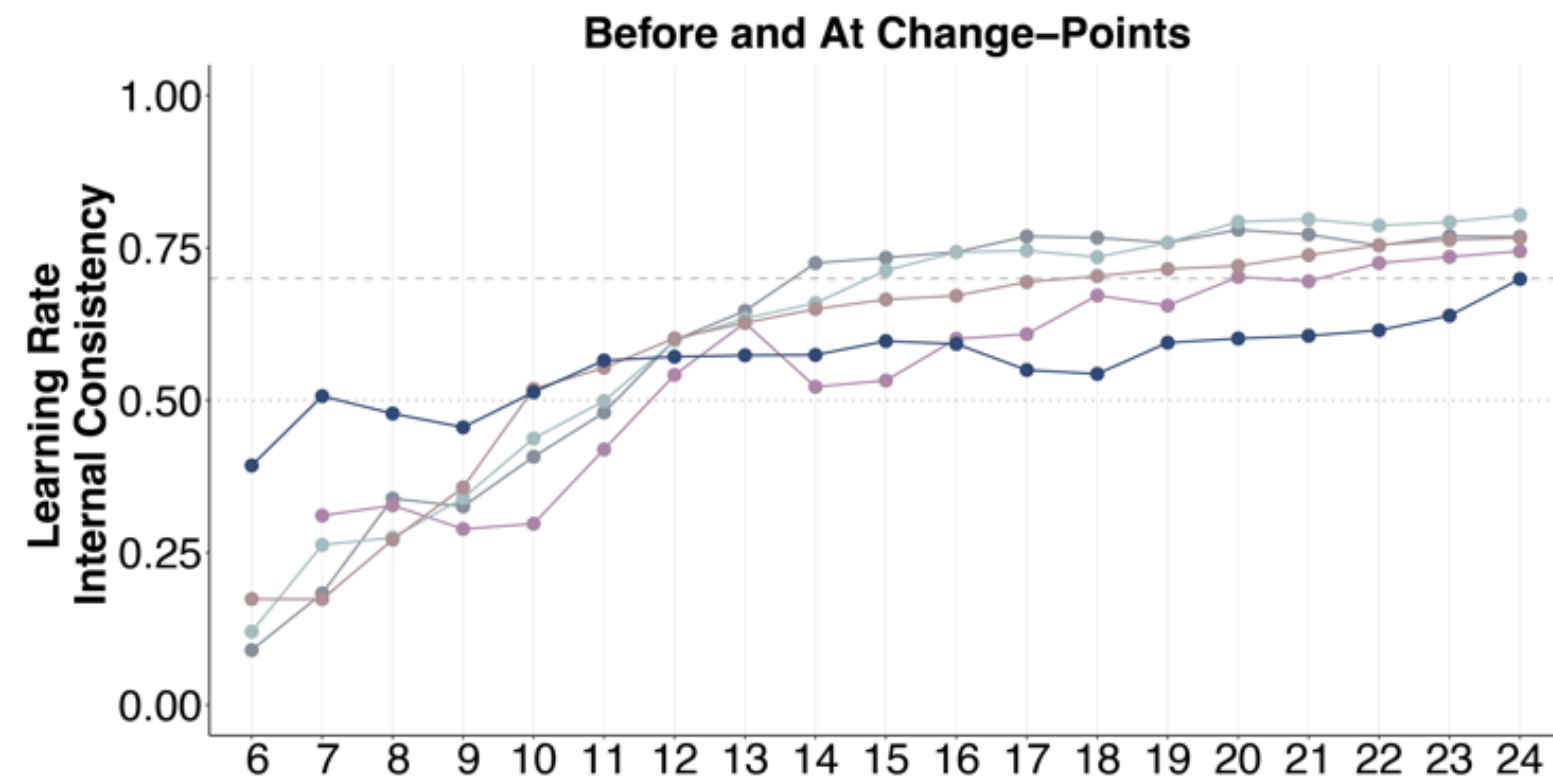


Zoom in! Low scores might be driven by low between participant variability that might fit some studies.



Check correlations, identifiability of your parameters etc.!

- If necessary, mend your task and model
- Alterations like hazard rate, trial number etc. can influence reliability estimates



Overview

Study design

Validity assessments

Reliability assessments

Example study: assessment of a widely-used decision-making task

Dos and Don'ts

Dos and Don'ts



Dos

Specific research question & hypothesis that are clinically and neuroscientifically informed if theory-driven approach

Know your research population

Specific model and task design built in tandem

Simulate, simulate, simulate!

Run pilots and quality checks

Don'ts

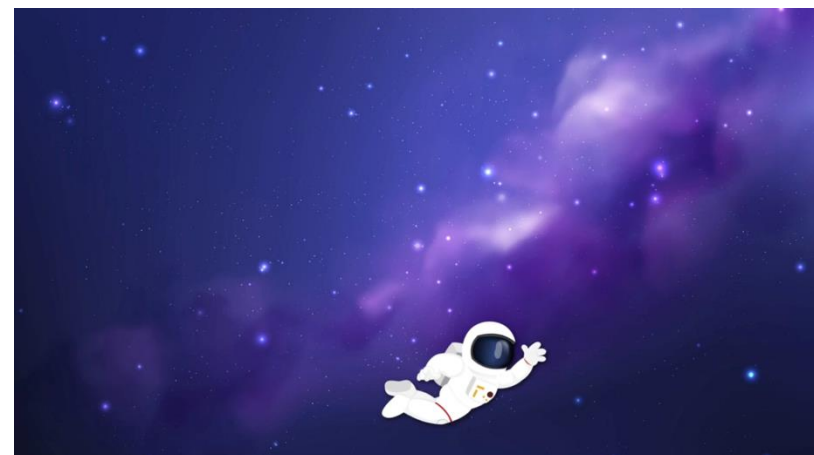
Taking any task for any population

Not talking to your participants

Overcomplicate your analysis

No simulations and model recoverability checks etc.

Overly long or boring tasks



"i thoroughly enjoyed playing the game!"

"This was so fun it was like playing a guessing game. I loved it so much and hope more games like this. My hand was tired but I was able to keep up. Thank you for this fun experiment. Have a wonderful day!"



https://github.com/christinamaher/Canva_VideoInstructions_Tutorial

Overview

Study design

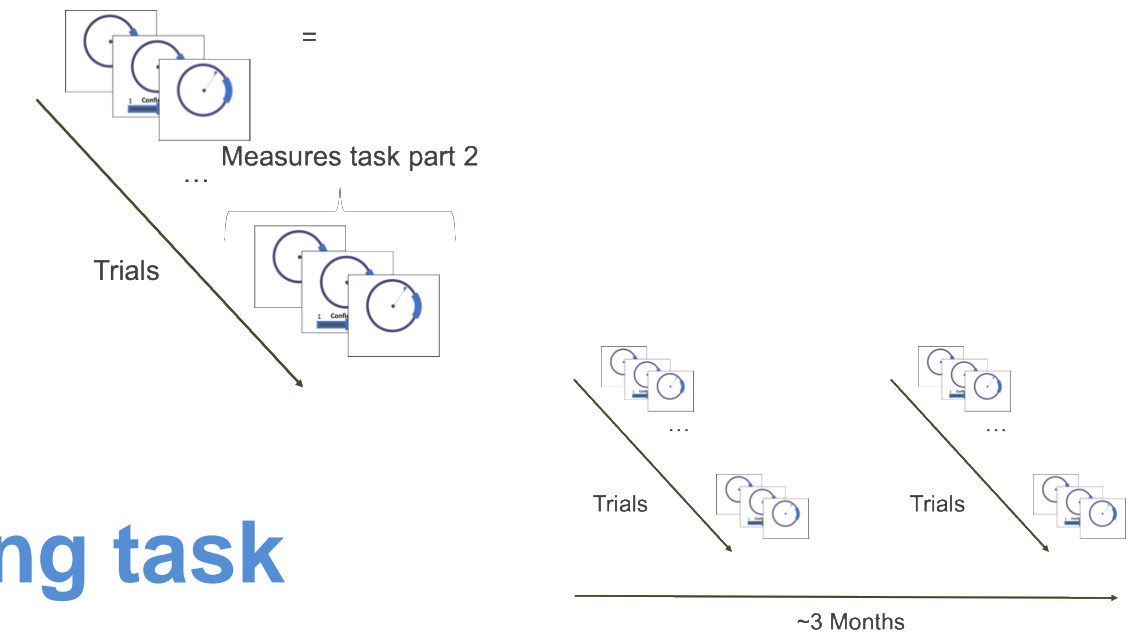


Validity assessments

Reliability assessments

Example study: assessment of a widely-used decision-making task

Dos and Don'ts



Thank you!



@AlisaLoosen
Alisa.loosen@mssm.edu