A large-scale analysis of test-retest reliabilities of self-regulation measures

Zeynep Enkavi
Stanford University

https://zenkavi.github.io/SRO_Retest_Analyses/output/reports/SRO_Report.nb.html

Motivation

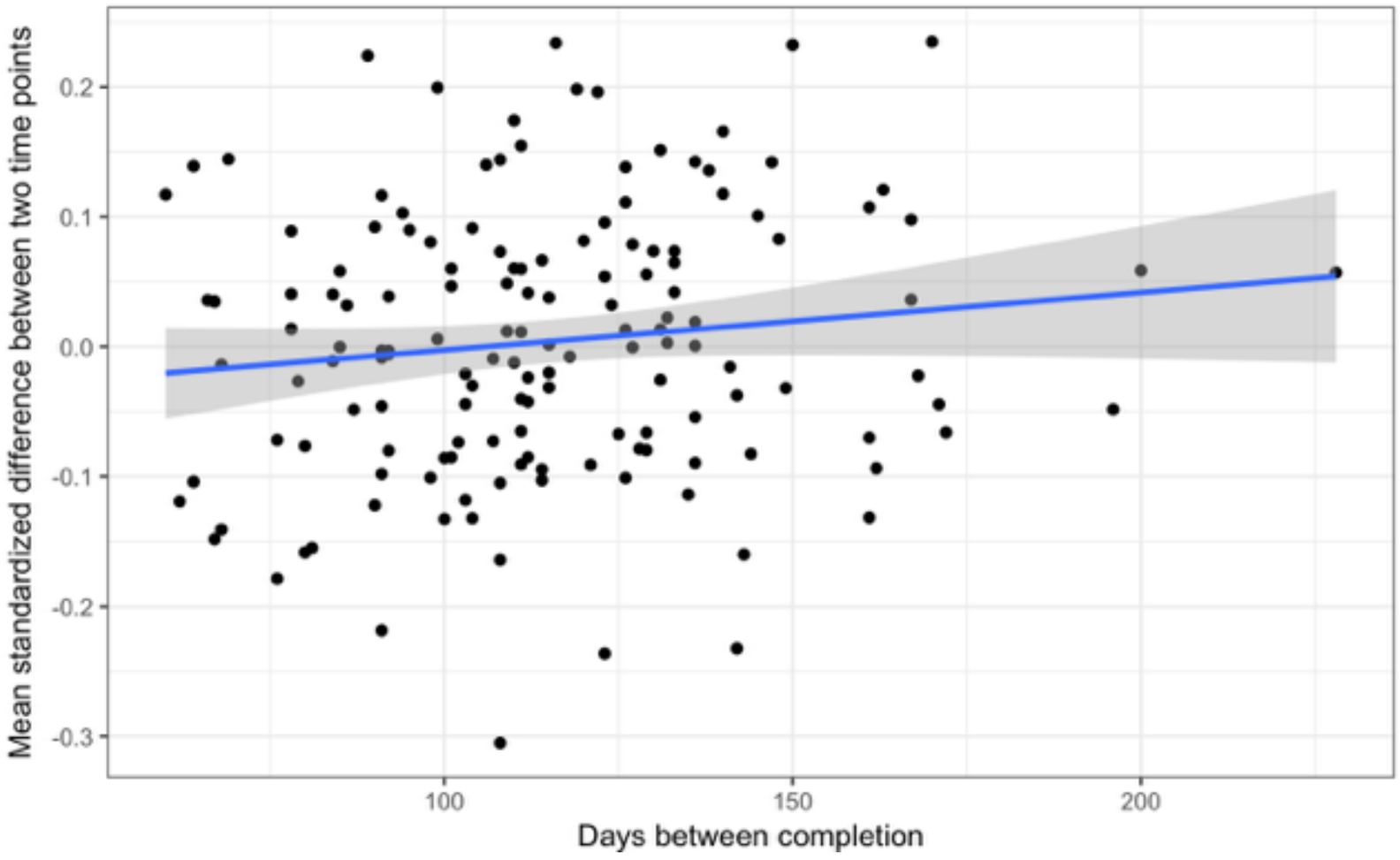
- The literature on retest reliabilities for task measures of self-regulation is not coherently and comprehensively summarized

- This hinders their evaluation as behavioral assays suitable for individual difference analyses

Procedure

- N = 150; 37 tasks; 23 surveys; 3 demographics
- Invited participants were chosen randomly and only subsets of them were invited for a given batch with the intention to avoid a potential oversampling and bias towards “high self regulators”.
- In total 242 participants were invited, 175 began the battery, 157 completed the battery and 150 provided data that passed qc for both time points.
- Data collection took place on average 115 number of days after the completion of the initial battery with a range of 60 to 228 days

Results



b = 0.0004, t = 1.55, p = 0.12

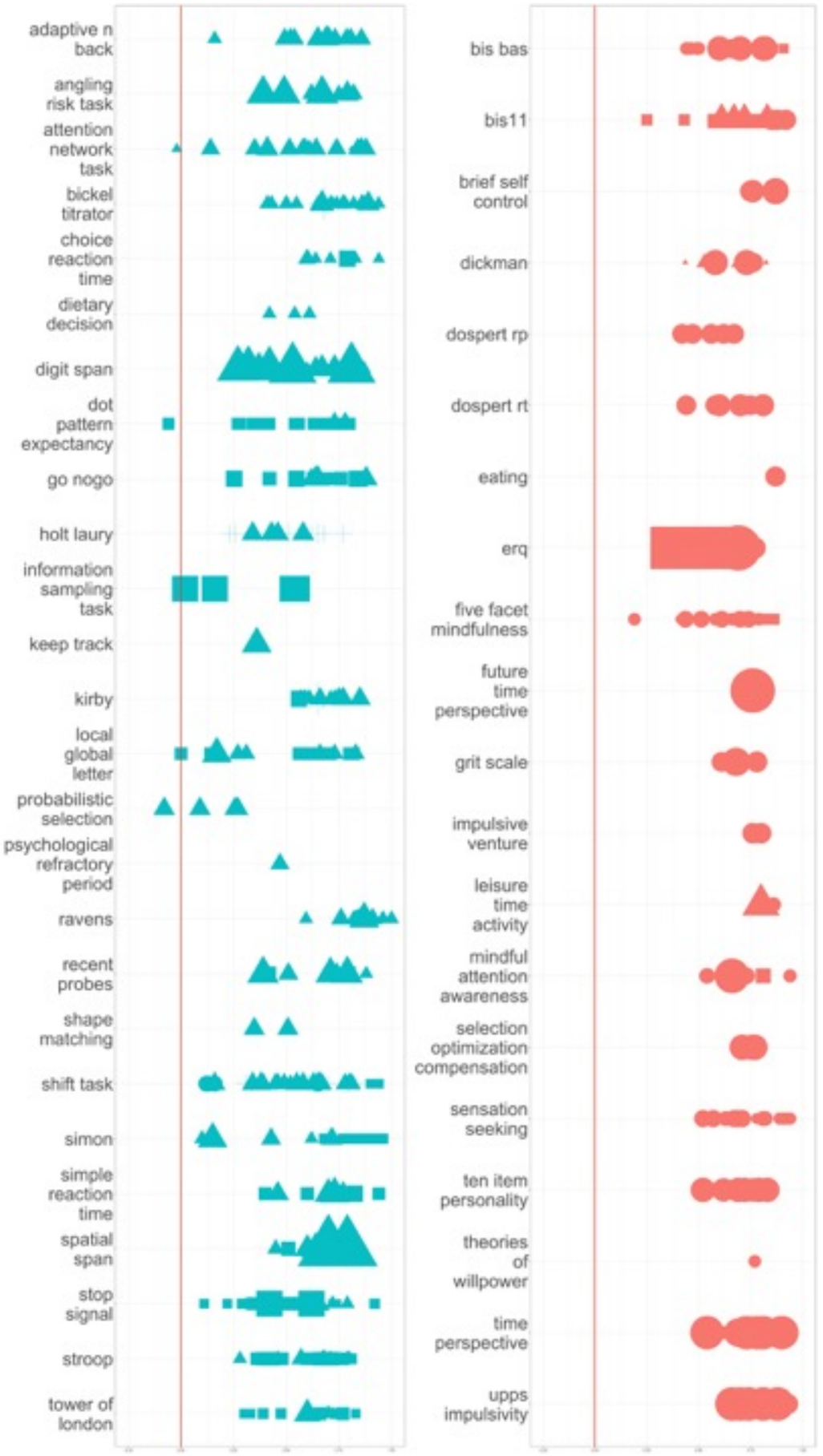
Lit review

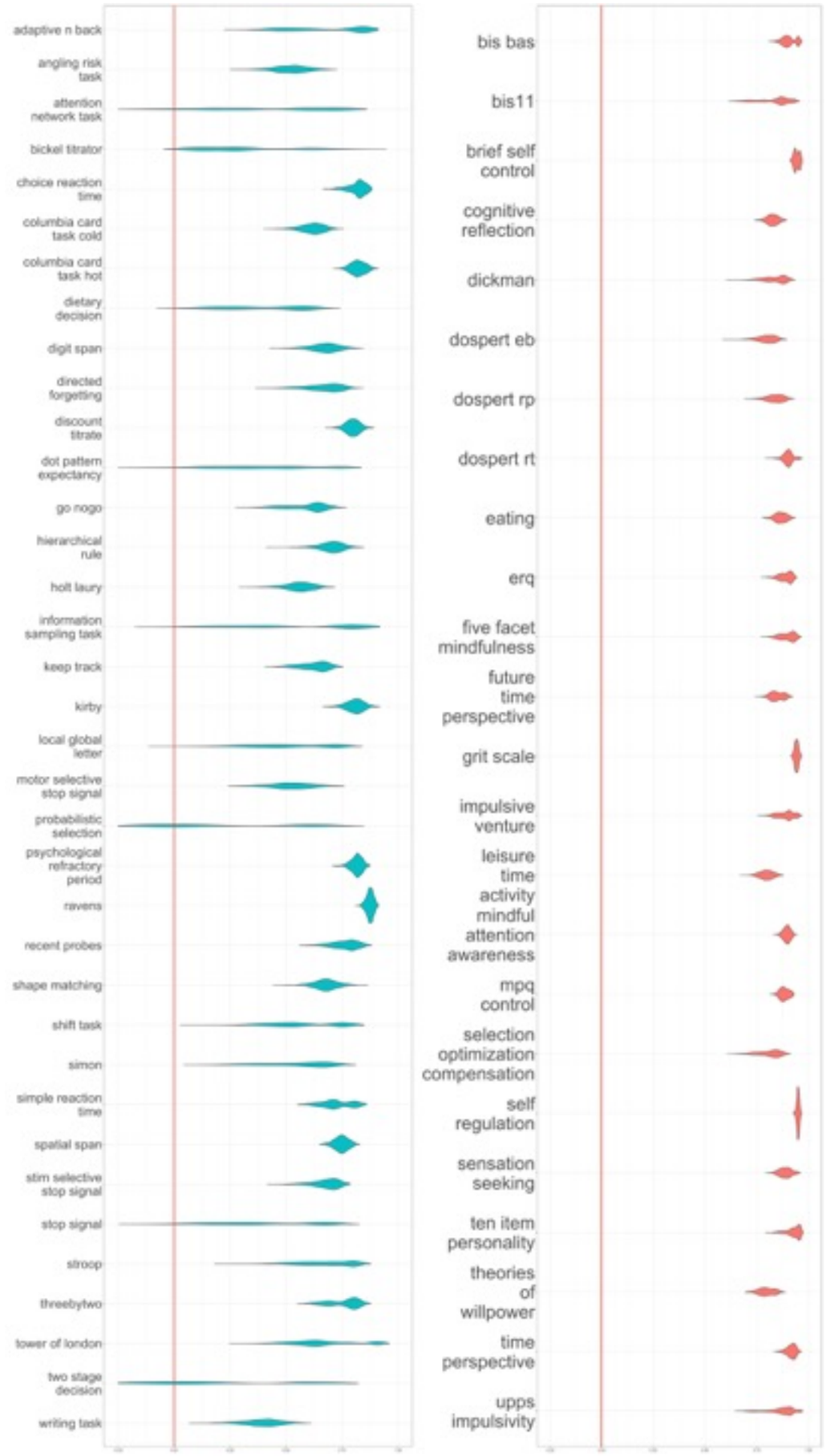

adaptive n back
angling risk task
attention network task
cicled titrator
choice reaction time
dietary decision
digit span
dot pattern expectancy
go nogo
holt laury
information sampling task
keep track
kirby
local global letter
probabilistic selection
psychological refractory period
ravens
recent probes
shape matching
shift task
simon
simple reaction time
spatial span
stop signal
stroop
tower of london
bis bas
bis11
brief self-control
dospert rp
dospert rt
eating
erq
five facet mindfulness
future time perspective
grit scale
impulsive venture
leisure time activity
mindful attention awareness
selection optimization compensation
sensation seeking
ten item personality
theories of willpower
time perspective
upps impulsivity
-0.25
0.00
0.25
0.50
0.75
1.00
sample_size
300
600
900
type
ICC
Kendall's Tau
Pearson
Spearman

Lit review

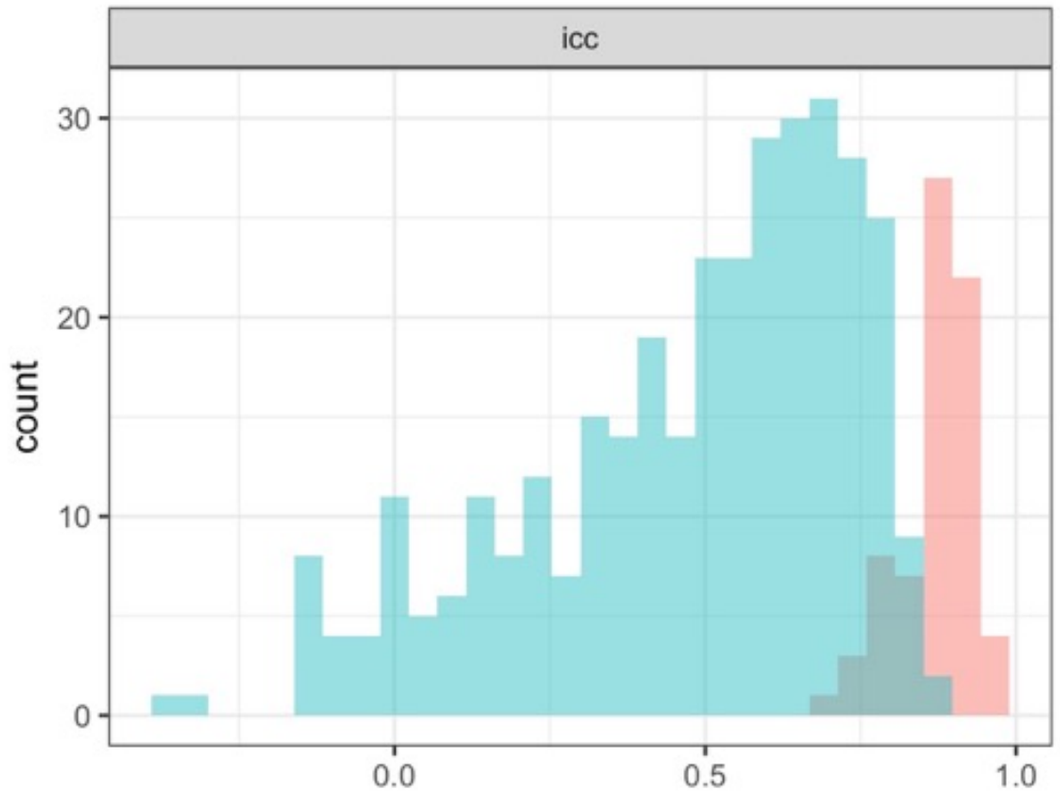

adaptive n back
angling risk task
attention network task
bickel titrator
choice reaction time
dietary decision
digit span
dot pattern expectancy
go nogo
holt laury
information sampling task
keep track
probabilistic selection
psychological refractory period
ravens
recent probes
shape matching
shift task
simon
simple reaction time
spatial span
stop signal
stroop
tower of london
bis bas
bis11
brief self control
dospert rp
dospert rt
eating
mindful attention awareness
selection optimization compensation
sensation seeking
ten item personality
theories of willpower
time perspective
upps impulsivity
sample_size 300 600 900
type ICC Kendall's Tau Pearson Spearman
-0.25 0.00 0.25 0.50 0.75 1.00


adaptive n back
angling risk task
attention network task
bickel titrator
choice reaction time
dietary decision
digit span
dot pattern expectancy
go nogo
holt laury
information sampling task
keep track
kirby
local global letter
probabilistic selection
psychological refractory period
ravens
recent probes
shape matching
shift task
simon
simple reaction time
spatial span
stop signal
stroop
tower of london
bis bas
bis11
brief self control
dickman
dospert rp
dospert rt
eating
erq
five facet mindfulness
future time perspective
grit scale
impulsive venture
leisure time activity
mindful attention awareness
selection optimization compensation
sensation seeking
ten item personality
theories of willpower
time perspective
upps impulsivity


adaptive n back
angling risk task
attention network task
bickel titrator
choice reaction time
dietary decision
digit span
dot pattern expectancy
go nogo
holt laury
information sampling task
keep track
kirby
local global letter
probabilistic selection
psychological refractory period
ravens
recent probes
shape matching
shift task
simon
simple reaction time
spatial span
stop signal
stroop
tower of london
bis bas
bis11
brief self control
dickman
dospert rp
dospert rt
eating
erq
five facet mindfulness
future time perspective
grit scale
impulsive venture
leisure time activity
mindful attention awareness
selection optimization compensation
sensation seeking
ten item personality
theories of willpower
time perspective
upps impulsivity


adaptive n back
angling risk task
attention network task
bickel titrator
choice reaction time
columbia card task cold
columbia card task hot
dietary decision
digit span
directed forgetting
discount titrate
dot pattern expectancy
go nogo
hierarchical rule
holt laury
information sampling task
keep track
kirby
local global letter
motor selective stop signal
probabilistic selection
psychological refractory period
ravens
recent probes
shape matching
shift task
simon
simple reaction time
spatial span
stim selective stop signal
stop signal
stroop
threebytwo
tower of london
two stage decision
writing task
bis bas
bis11
brief self control
cognitive reflection
dickman
dospert eb
dospert rp
dospert rt
eating
erq
five facet mindfulness
future time perspective
grit scale
impulsive venture
leisure time activity
mindful attention awareness
mpq control
selection optimization compensation
self regulation
sensation seeking
ten item personality
theories of willpower
time perspective
upps impulsivity

Results



b = -0.41, t(344) = -12.6, p < 0.001

task	median_spearman	median_icc	median_pearson	median_eta_sq	median_sem	num_vars
survey	0.787	0.883	0.791	0.003	0.369	69
task	0.386	0.526	0.369	0.006	0.318	277

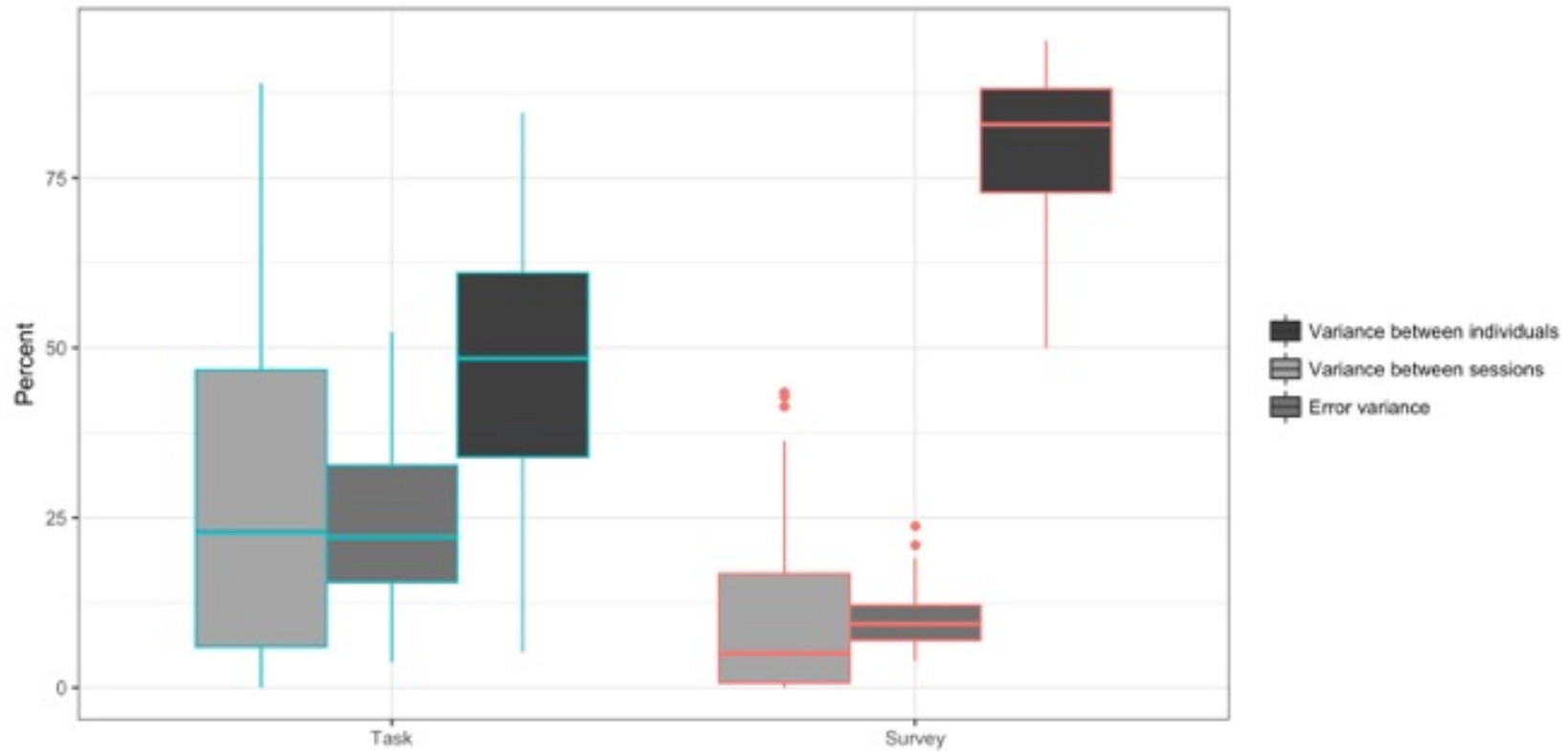
Results

$$ICC = \frac{Variance\ \ between\ \ individuals}{Variance\ \ between\ \ individuals\ \ + Error\ \ variance\ \ + Variance\ \ between\ \ sessions}$$

Hedge, Powell, Sumner (2017)

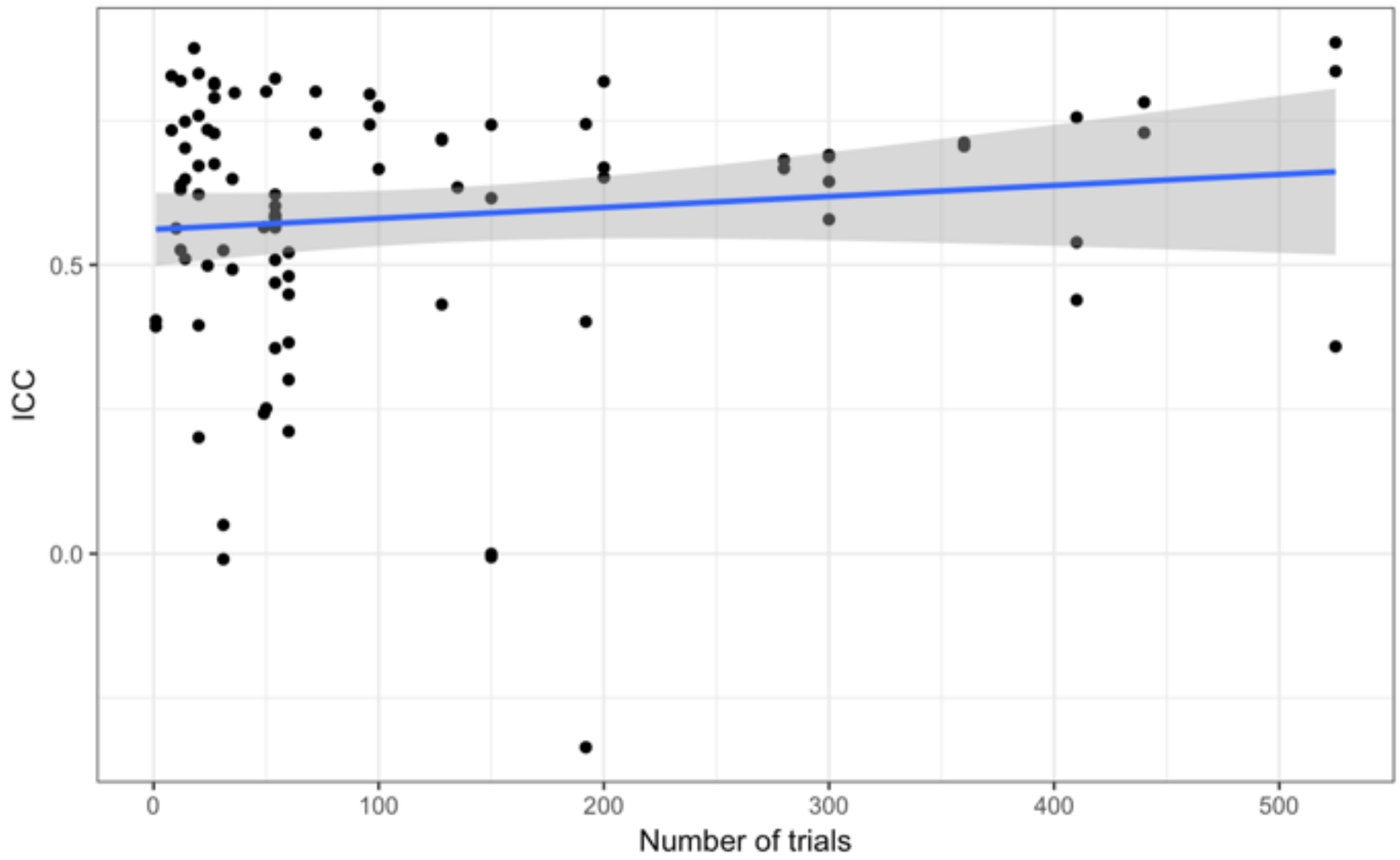

adaptive n back
avg rt
mean load
acc
angling risk task
release loss percent
release adjusted clicks
release score
keep score
keep adjusted clicks
keep loss percent
attention network task
congruent rt
orienting rt
neutral rt
acc
incongruent rt
orienting acc
conflict acc
alerting acc
alerting rt
conflict rt
bickel titrator
auc small
hyp discount rate medium
hyp discount rate small
hyp discount rate large
choice reaction time
avg rt
acc
columbia card task cold
information use
probability sensitivity
pain sensitivity
avg cards chosen
loss sensitivity
columbia card task hot
avg cards chosen
information use
loss sensitivity
pain sensitivity
probability sensitivity
dietary decision
prop healthy choice
health sensitivity
taste sensitivity
digit span
reverse span
forward span
directed forgetting
acc
proactive interference acc
proactive interference rt
avg rt
discount titrate
percent patient
dot pattern expectancy
avg rt
AY BY rmss
BX BY acc
AX BX acc
BX BY rt
AY BY rt
dprime
acc
go nogo
go rt
bias
commission errors
dprime
rtci
omission errors
hierarchical rule
score
acc
avg rt
holt laury
risky choices
safe choices
risk aversion
information sampling task
Fixed Win motivation
Fixed Win rt correct
Fixed Win acc
Fixed Win boxes opened
Decreasing Win rt correct
Decreasing Win boxes opened
Decreasing Win acc
Decreasing Win motivation
keep track
score
kirby
exp discount rate small
hyp discount rate medium
hyp discount rate small
exp discount rate medium
exp discount rate large
hyp discount rate large
local global letter
local incongruent rt
incongruent rt
avg rt
congruent rt
local congruent rt
conflict rt
local conflict rt
global incongruent rt
conflict acc
global conflict rt
incongruent rtrmss acc
global incongruent errors
global error cost
incongruent rtrmss rt
local congruent errors
congruent facilitation acc
local incongruent errors
global congruent errors
global congruent rt
switch cost rt
congruent facilitation rt
local error cost
switch cost acc
motor selective stop signal
ignore rt error std
SSRT
ignore rt error
ignore acc
selective proactive control
reactive control
probabilistic selection
avoid trial rt
approach trial rt
approach trial acc
value sensitivity
avoid trial acc
positive learning bias
overall task acc
psychological refractory period
PRP slope
PRP slowing
task1 acc
task2 acc
ravens
score
recent probes
acc
avg rt
proactive interference acc
proactive interference rt
shape matching
acc
negative priming
avg rt
shift task
avg rt
fail to maintain set
conceptual responses
num cat achieved
perseverative responses
perseverative errors
learning to learn
learning rate
acc
total errors
nonperseverative errors
simon
incongruent sd rt
simon
congruent acc
simon acc
congruent avg rt
avg rt
incongruent acc
incongrdent avg rt
acc
congruent sd rt
simple reaction time
avg rt
spatial span
forward span
reverse span
stim selective stop signal
SSRT
ignore acc
ignore rt error
ignore rt error std
stop signal
total errors
omission errors
go rt
stop rt error
commission errors
stop acc
proactive SSRT slowing
SSRT high
SSRT
SSRT low
go rt sd error
proactive slowing
inhibition slope
stroop
congruent rt
incongruent errors
avg rt
incongruent rt
stroop acc
stroop rt
errors
acc
congruent errors
threebytwo
avg rt
acc
cue switch cost rt 100 0
task switch cost rt 100 0
cue switch cost rt 900 0
task switch cost rt 900 0
tower of london
planning time
num extra moves
num optimal solutions
avg move time
writing task
neutral probability
positive probability
0
25
50
75
100
bis bas
BIS
BAS reward responsiveness
BAS drive
BAS total
BAS fun seeking
bis11
Attentional
Nonplanning
total
Motor
brief self control
self control
cognitive reflection
intuitive proportion
correct proportion
dickman
dysfunctional
functional
dospert eb
social
financial
ethical
health safety
recreational
dospert rp
health safety
ethical
social
financial
recreational
dospert rt
health safety
financial
recreational
social
ethical
eating
emotional eating
cognitive restraint
uncontrolled eating
erq
suppression
reappraisal
five facet mindfulness
nonjudge
observe
describe
act with awareness
nonreact
future time perspective
future time perspective
grit scale
grit
impulsive venture
venturesomeness
impulsiveness
leisure time activity
activity level
mindful attention awareness
mindfulness
mpq control
control
selection optimization compensation
compensation
elective selection
loss based selection
optimization
self regulation
control
sensation seeking
disinhibition
experience seeking
total
thrill adventure seeking
boredom susceptibility
ten item personality
extraversion
openness
agreeableness
emotional stability
conscientiousness
theories of willpower
endorse limited resource
time perspective
past positive
past negative
present hedonistic
present fatalistic
future
upps impulsivity
negative urgency
positive urgency
lack of perseverance
sensation seeking
lack of premeditation
0
25
50
75
100
Variance between individuals
Variance between sessions
Error variance

Results



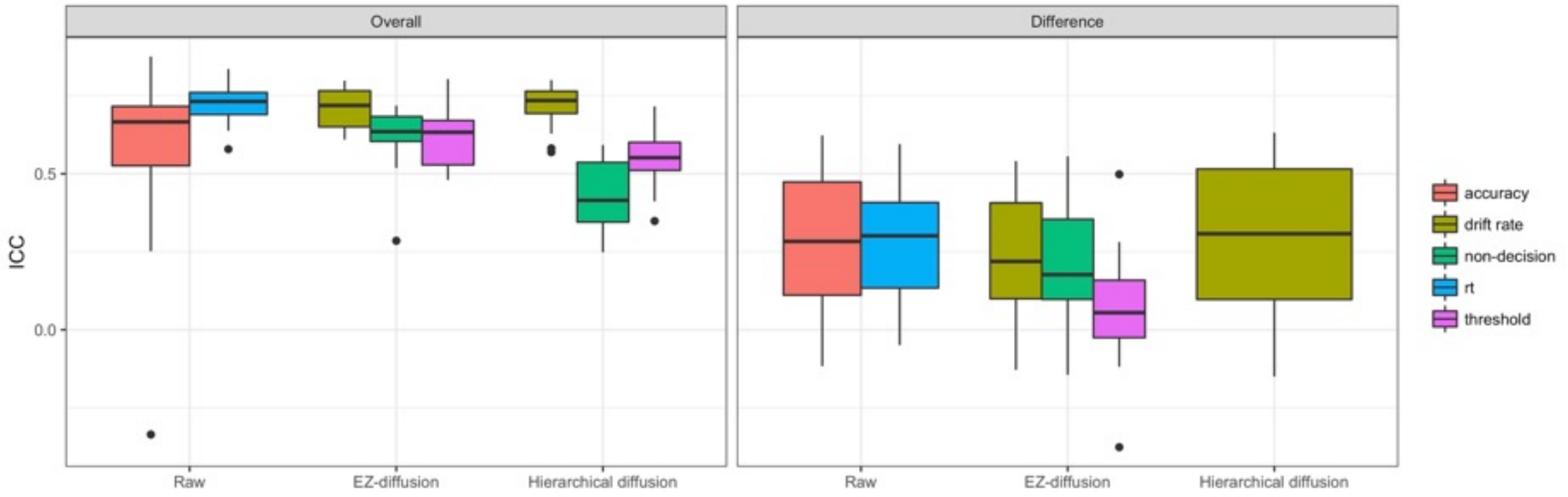
task*variance between individuals:
b = -49.074, t(1032) = -14.93, p<0.001

Results



b = 0.0002, t(88) = 1.12, p = 0.27

Results


Overall
Difference
ICC
0.5
0.0
Raw
EZ-diffusion
Hierarchical diffusion
Raw
EZ-diffusion
Hierarchical diffusion
accuracy
drift rate
non-decision
rt
threshold

Conclusions

- Dependent variables from cognitive tasks show larger variability and lower reliability compared to measures from surveys
  - Due to lower between subject variability
- Looking closer at task variables:
  - Number of trials does not have an effect on reliability
  - Drift diffusion parameters show similar reliability to RT and accuracy

Open Questions

- How to utilize reliability estimates for construct validity

  - Attenuation correction?

$$\text{"True" correlation}(x,y) = \frac{\text{Sample correlation}(x,y)}{\sqrt{\text{Reliability}(x).\text{Reliability}(y)}}$$