

Using performance discontinuities to estimate individual Working-Memory Capacities in serial recall tasks

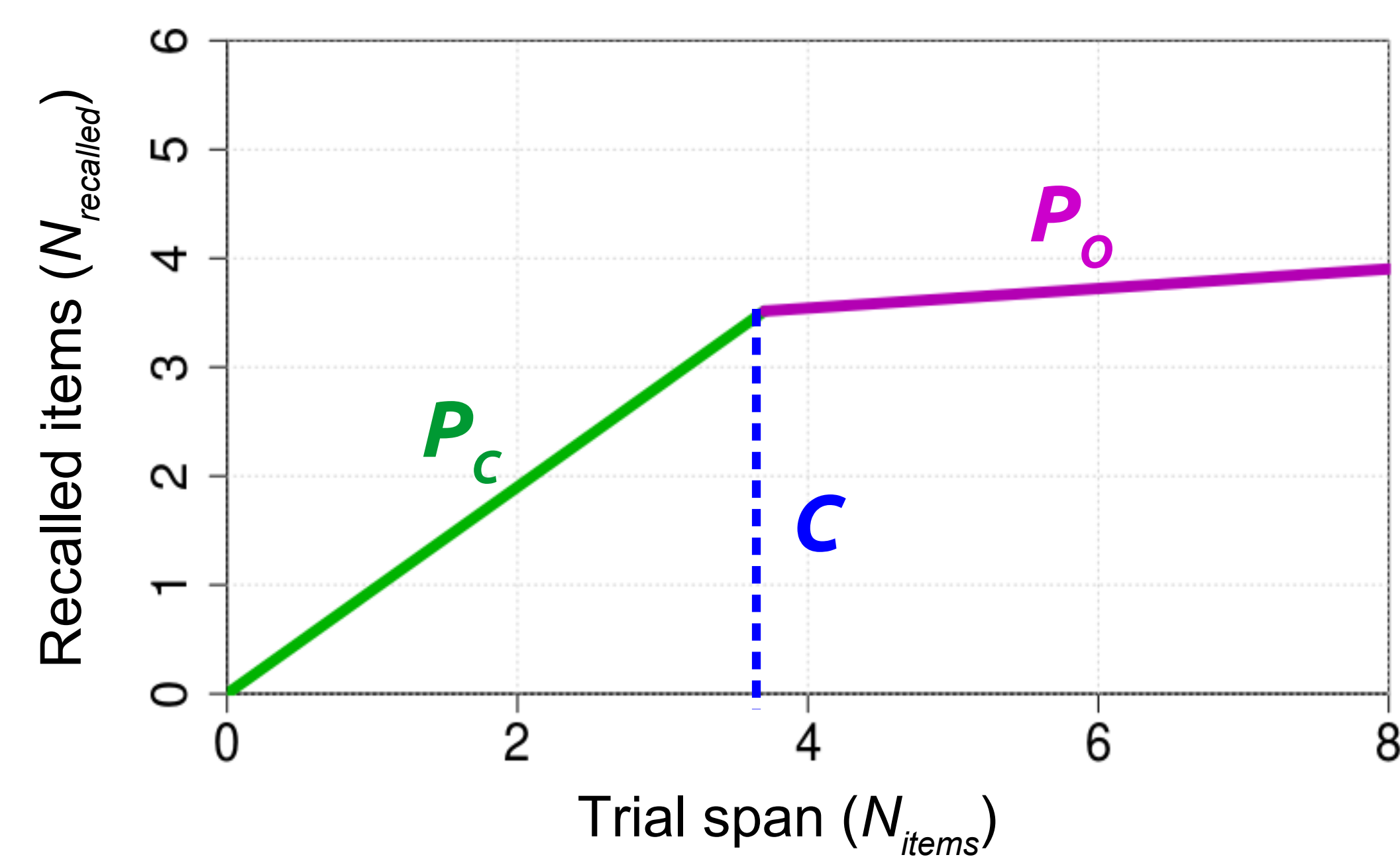
Jonas Kristoffer Lindeløv, Centre for Cognitive Neuroscience, Aalborg University, Denmark, jonas@hum.aau.dk

Introduction

Complex Span tasks and other serial recall tasks are typically scored using **simple indices**, such as the proportion of correctly recalled items (e.g., Conway, 2005). These indices **are not interpretable** as Working Memory Capacity.

The performance discontinuity model

... consists of **just three informative parameters**: C_i the subject's capacity, P_C a high proportion of correctly recalled items within C_i , and P_O a lower proportion for further presented items. For a given trial with $N_{recalled}$ correctly recalled out of N_{items} presented items:



$$N_{recalled_{i,j}} = \begin{cases} P_C N_{items_{i,j}} & \text{if } N_{items_{i,j}} \leq C_i \\ P_C C_i + P_O (N_{items_{i,j}} - C_i) & \text{if } N_{items_{i,j}} > C_i \end{cases}$$

Method

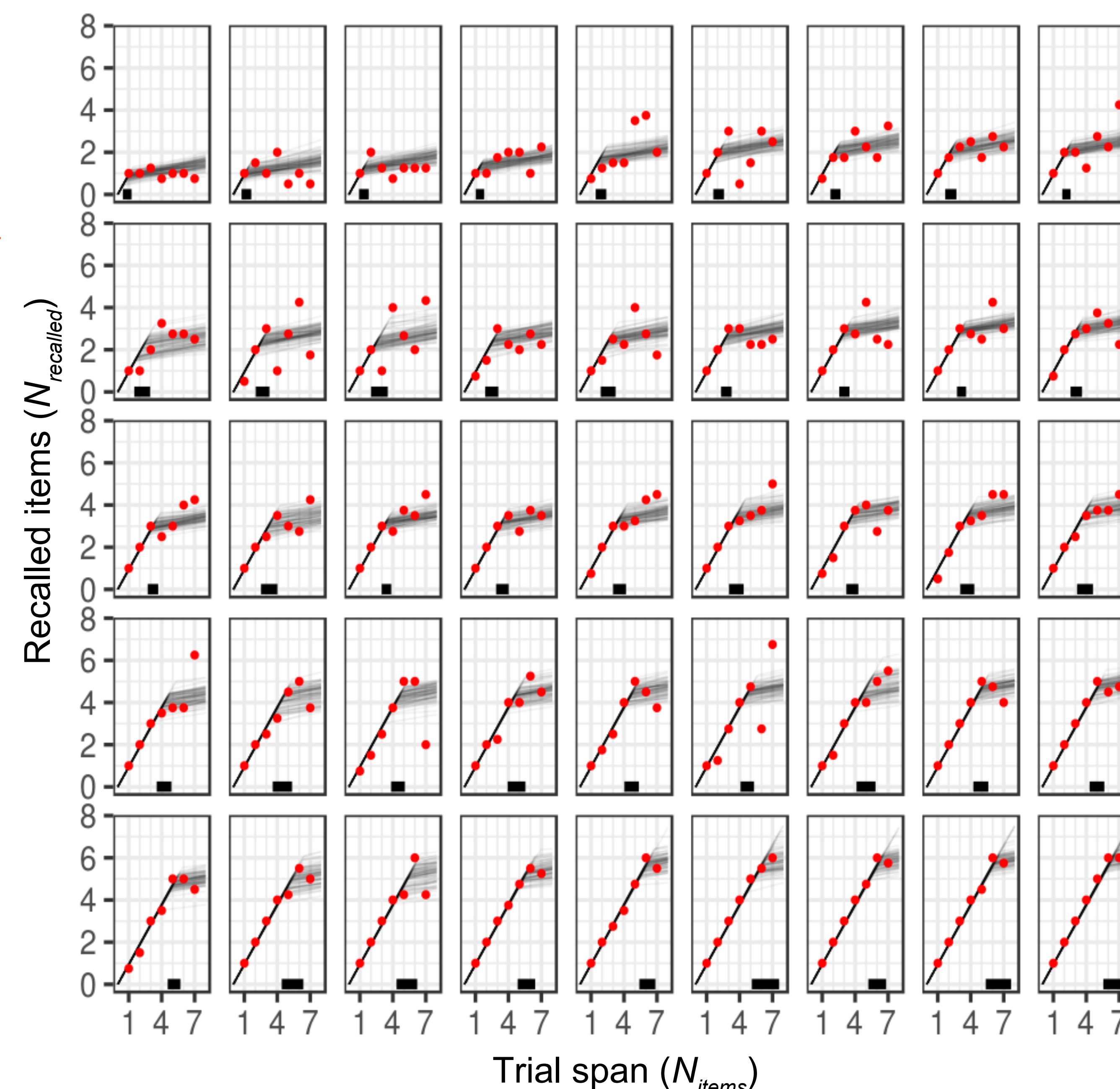
We estimated the parameters using JAGS for Bayesian inference (Plummer, 2003) on **Operation Span data from 46 subjects** who did four repetitions of spans one through seven, i.e., 28 trials.

Results: parameter estimates

The population C is 3.7, congruent with previous estimates (Cowan, 2001). C marks a sharp discontinuity between an almost perfect rate of recall in working memory and barely any recalled items when C is exceeded.

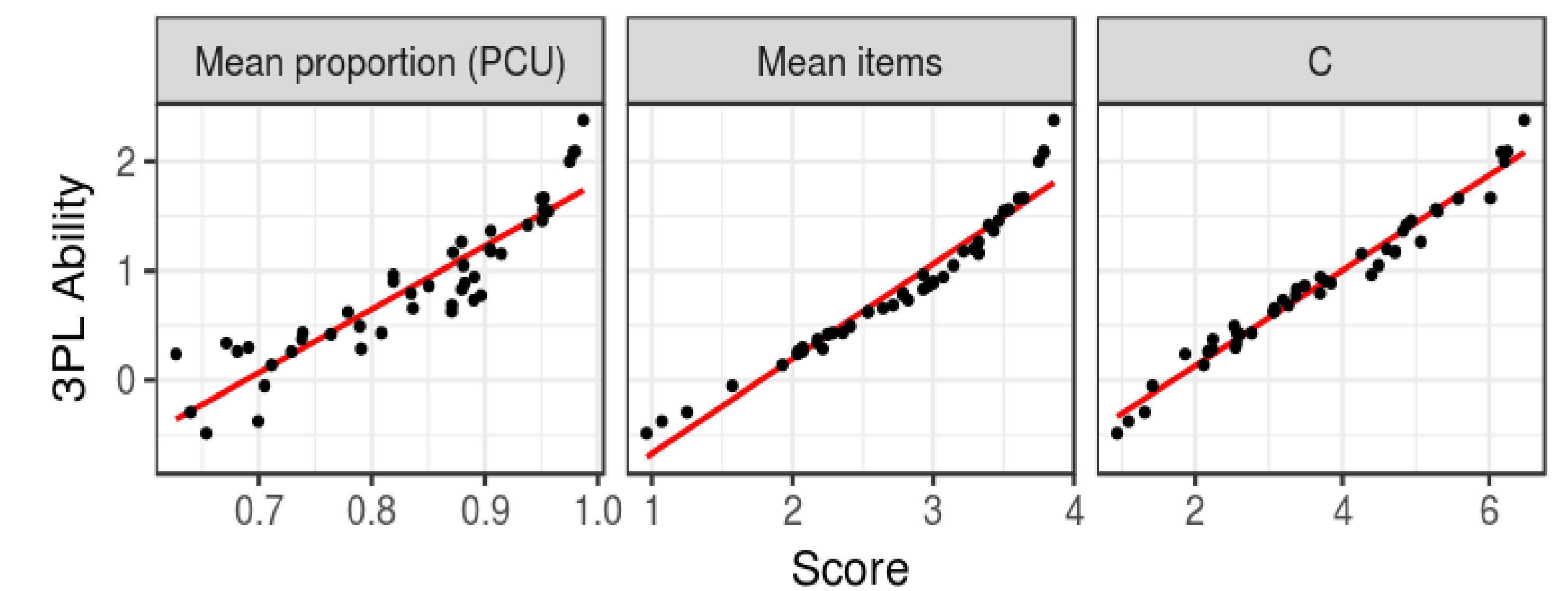
Estimate	95% CI (N = 46)
$C = 3.7$ chunks	3.2 – 4.2 chunks
$P_C = 94.5\%$	93.2 – 96.1 %
$P_O = 9.1\%$	3.9 – 15.0 %

For individuals (panels), the model has a good fit (slopes) to actual recall (dots) and C is identifiable (black horizontal 95% CI):

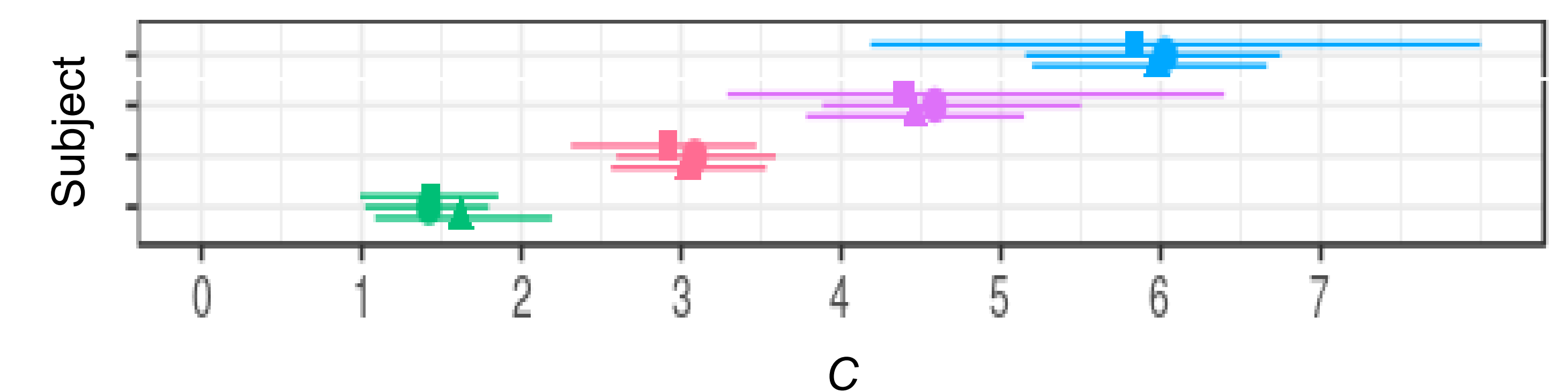


Results: validity and robustness

Compared to conventional scoring methods, C s **scales linearly with the underlying ability** ($r = .987$), as estimated using a three-parameter logistic model (3PL):



Contrary to conventional indices, C **remains stable when some N_{items} are omitted**. Here are estimates of C based on $N_{items} = 1-7, 3-7,$ and $1-5$ for each of four example subjects (colors):



Discussion

The primary strength of the performance discontinuity score of C is the **low model complexity high interpretability**. It requires super- and supra- C trials and the non-zero P_O slope would be inaccurate for large N_{items} , so the model has a **restricted domain of applicability**. Within this domain, it outperforms classical scoring of Complex Span tasks.