**Jump around! An investigation of the relationship between intelligence and the alpha parameter of the Levy-Flight Model**

Sven Lesche1

1 Ruprecht-Karls-University Heidelberg

**Author note**

Code needed to replicate this work can be found at: <https://github.com/SLesche/levy-intelligence>.

Correspondence concerning this article should be addressed to Sven Lesche, Kurfürsten-Anlage 57. E-mail: [sven.lesche@stud.uni-heidelberg.de](mailto:sven.lesche@stud.uni-heidelberg.de)

**Jump around! An investigation of the relationship between intelligence and the alpha parameter of the Levy-Flight Model**

Intelligence is one of the most well researched constructs in psychology, playing a crucial role in predicting success in education, job performance, and overall well-being. Understanding the cognitive processes that underpin intelligence has been the focus of decades of research. One of the most robust findings in this area is the relationship between measures of mental speed and intelligence, with faster cognitive processing often associated with higher intelligence.

In a review of 172 studies, Sheppard and Vernon ([2008](#ref-sheppard2008intelligence)) reported an average correlation of between various mental speed measures and various intelligence measures. Most studies investigating the relationship between intelligence and mental speed use elementary cognitive tasks (ECTs) to obtain a measure of mental speed. ECTs are simple tasks with very low cognitive demands. For example, participants have to respond by key-press in which of two squares a stimulus has appeared. This limits the role of individual differences in strategy and allows researchers to obtain a pure measure of mental speed.

Mental speed is assessed at different levels: behaviorally, through psychophysiological measures, or via cognitive models. On a behavioral level, researchers find robust negative correlations between intelligence and both mean response time and variability of response time ([Doebler & Scheffler, 2016](#ref-doebler2016relationship); [Jensen, 2006](#ref-jensen2006clocking)). More intelligent individuals tend to have lower mean response times and lower variability in response times. Using EEG measures, Schubert et al. ([2017](#ref-schubert2017general)) found that the latency of event-related potentials are able to explain 80% of the variance in intelligence. In cognitive models, evidence accumulation models like the drift-diffusion model (DDM) ([Ratcliff & McKoon, 2008](#ref-ratcliff2008diffusion)) are employed to obtain model parameters of mental speed.

The DDM assumes that participants steadily accumulate evidence during a trial. Once they have accumulated enough evidence pointing towards a response, a decision and subsequent reaction is made. The model attempts to fit the empirical distribution of response times and accuracies by optimizing several different parameters. First, response time is separated into decision time, where the noisy evidence accumulation process takes place, and non-decision time . Non-decision time incorporates processes such as stimulus encoding and motor processes. Decision time is further controlled by three main parameters: (1) the boundary separation describes the distance between decision boundaries and represents the amount of evidence required to reach a decision, (2) the mean drift rate is the average rate at which evidence is accumulated and (3) the mean starting point *z* represents the average point at which the evidence accumulation process begins and can account for an a-priori bias towards one of the two-choice alternatives.

Further parameters of across-trial variability can be added to reflect changes in diffusion model parameters over time. Variability of the starting point *sz* accounts for differences in a-priori bias towards one response alternative across trials ([Laming, 1968](#ref-laming1968information); [Ratcliff & Rouder, 1998](#ref-ratcliff1998modeling)). Variability of drift rate reflects changes in drift rate across trials and allows the model to account for systematic differences between trials, such as error responses being systematically faster than correct responses ([Ratcliff, 1978](#ref-ratcliff1978theory); [Ratcliff & Rouder, 1998](#ref-ratcliff1998modeling)). Finally, variability of non-decision time *st* reflects fluctuations in the duration of encoding or motor-execution processes across trials.

Of these parameters, drift rate has so far shown the most robust correlation to intelligence ([Ratcliff et al., 2011](#ref-ratcliff2011effects); [Schmiedek et al., 2007](#ref-schmiedek2007individual); [Schubert et al., 2015](#ref-schubert2015decomposing)). Boundary separation is often negatively associated with intelligence, but the effect is smaller than for drift rates ([Ratcliff et al., 2011](#ref-ratcliff2011effects); [Schubert et al., 2015](#ref-schubert2015decomposing)).

In both the simple and full DDM, the drift rate is an average rate of evidence accumulation. In each step, this average rate is influenced by noise, resulting in the random walk with a certain drift that characterizes the evidence accumulation process. This noise is assumed to be normally distributed with a mean of 0. The standard deviation of the noise distribution is the “diffusion constant” and is set to 1 in most applications.

Under the assumption that noise is normally distributed, large jumps in accumulated evidence become very unlikely. Recently, Voss et al. ([2019](#ref-voss2019sequential)) introduced an extension of the DDM that modifies the noise distribution. Here, the normal distribution is replaced by a heavy tailed distribution. This increases the likelihood of larger jumps occurring in the evidence accumulation process. The Cauchy distribution is an example of such a heavy-tailed distribution. Both the Cauchy distribution and the normal distribution are specific instances of the family of so-called Levy alpha-stable distributions. Within this family, a stability parameter governs the heaviness in the tails of the distribution. = 2 results in the normal distribution, = 1 results in the Cauchy distribution. In the Levy-flight model, is an additional parameter that can be estimated from the data ([Voss et al., 2019](#ref-voss2019sequential); [Wieschen et al., 2020](#ref-wieschen2020jumping)).

Higher values of lead to more stable evidence accumulation with fewer jumps. Lower values of lead to less stable evidence accumulation, showing more extreme jumps. In behavioral data, higher variability of response times is associated with lower intelligence ([Doebler & Scheffler, 2016](#ref-doebler2016relationship)). Intelligent individuals are more “stable” in their response times. Following these findings, higher values of may be positively associated with intelligence. Wieschen et al. ([2020](#ref-wieschen2020jumping)) refers to this as the “inefficient jumping” hypothesis. The jumps in evidence accumulation more common in heavy-tailed noise distributions are deemed inefficient while completing the response time task. Intelligent individuals should therefore display fewer inefficient jumps and thus have higher values of .

Contrasting this, jumps in evidence accumulation may also be beneficial. In simple response time tasks with high accuracy, erratically accumulating information may be an efficient way to decrease response times. This is termed the “efficient jumping” hypothesis ([Wieschen et al., 2020](#ref-wieschen2020jumping)). The efficient jumping hypothesis claims that more intelligent individuals make use of larger jumps in evidence accumulation to complete the task more quickly. They should therefore display lower values of .

This present work aims to provide a first insight into the question of efficient vs. inefficient jumping. To this end, I applied the Levy-flight model to data from three different response time tasks and investigated the correlation between the parameter and intelligence. Following Schubert et al. ([2016](#ref-schubert2016trait)), I used structural equation modeling to obtain a latent factor for both and intelligence to account for measurement error.

The inefficient jumping hypothesis predicts positive latent correlations between and *g*. More intelligent individuals display smaller heavy-tails. The efficient jumping hypothesis predicts negative correlations between and *g*. More intelligent individuals make efficient use of jumps in evidence accumulation and thus display heavier tails.

# Methods

## Participants

The data I analyse here was published in Schubert et al. ([2017](#ref-schubert2017general)). The original study consisted of three measurement occasions and included EEG measurements in the first and last measurement occasion. The present analysis focuses on the behavioral data collected during the first two measurement occasions. The sample consists of N = 122 (72 female, ) participants. All of the participants had normal or corrected to normal vision, no history of mental illness, provided informed consent, and received 100€ for their participation.

## Materials

I will focus on behavioral data from the Hick, Sternberg, and Posner Tasks. Detailed information can be found in the original publication ([Schubert et al., 2017](#ref-schubert2017general)).

Intelligence was measured using a computer-adapted version of Raven’s Advanced Progressive Matrices (APM) ([Raven, 1981](#ref-raven1981manual)) and the Berlin Intelligence Structure Test (BIS) ([Jäger et al., 1997](#ref-jager1997berliner)).

## Procedure

The first two measurement occasions were spaced approximately four months apart. The Hick, Sternberg, and Posner Tasks were administered at the first measurement occasion in the same order for all participants. The APM and BIS were administered at the second measurement occasion.

## Analysis

### Levy-Flight Model.

I estimated the parameters of the Levy-Flight model using BayesFlow ([Radev et al., 2020](#ref-radev2020bayesflow)). I used accuracy coding, with the upper boundary representing a correct response and the lower boundary representing an incorrect response. Hence, I fixed the start point to 0.5. I allowed drift rate , boundary separation , non-decision time , variability in the non-decision time and to vary between tasks and conditions.

To train the neural approximator, I employed the following priors:

I simulated 200,000 datasets and then trained the model for 150 epochs. After training, I investigated the computational faithfulness of the neural network by using simulation-based calibration ([Talts et al., 2018](#ref-talts2018validating)). I investigated the bias of the posterior distributions after applying the amortized inference network on newly simulated data. Furthermore, I investigated the ability of the model to recover true generating parameters by comparing the recovered posterior means to the true parameters used in simulation.

### Structural Equation Model.

I used structural equation modelling to evaluate the correlation between a latent first order factor of the parameter over all tasks and a *g* factor over all intelligence measures. In addition to the first order factor over all tasks and conditions, I included task-specific method factors. Intelligence was defined as a first order factor over all sub-tests of the BIS as well as the APM. I then investigated the relationship between the latent factors and *g*.

# Results

Model estimation was done in Python 3.11 with Bayesflow version 1.1.6 ([Radev et al., 2020](#ref-radev2020bayesflow)). Data preparation and analysis was completed in R [Version 4.3.2; R Core Team ([2023](#ref-R-base))][[1]](#footnote-1).

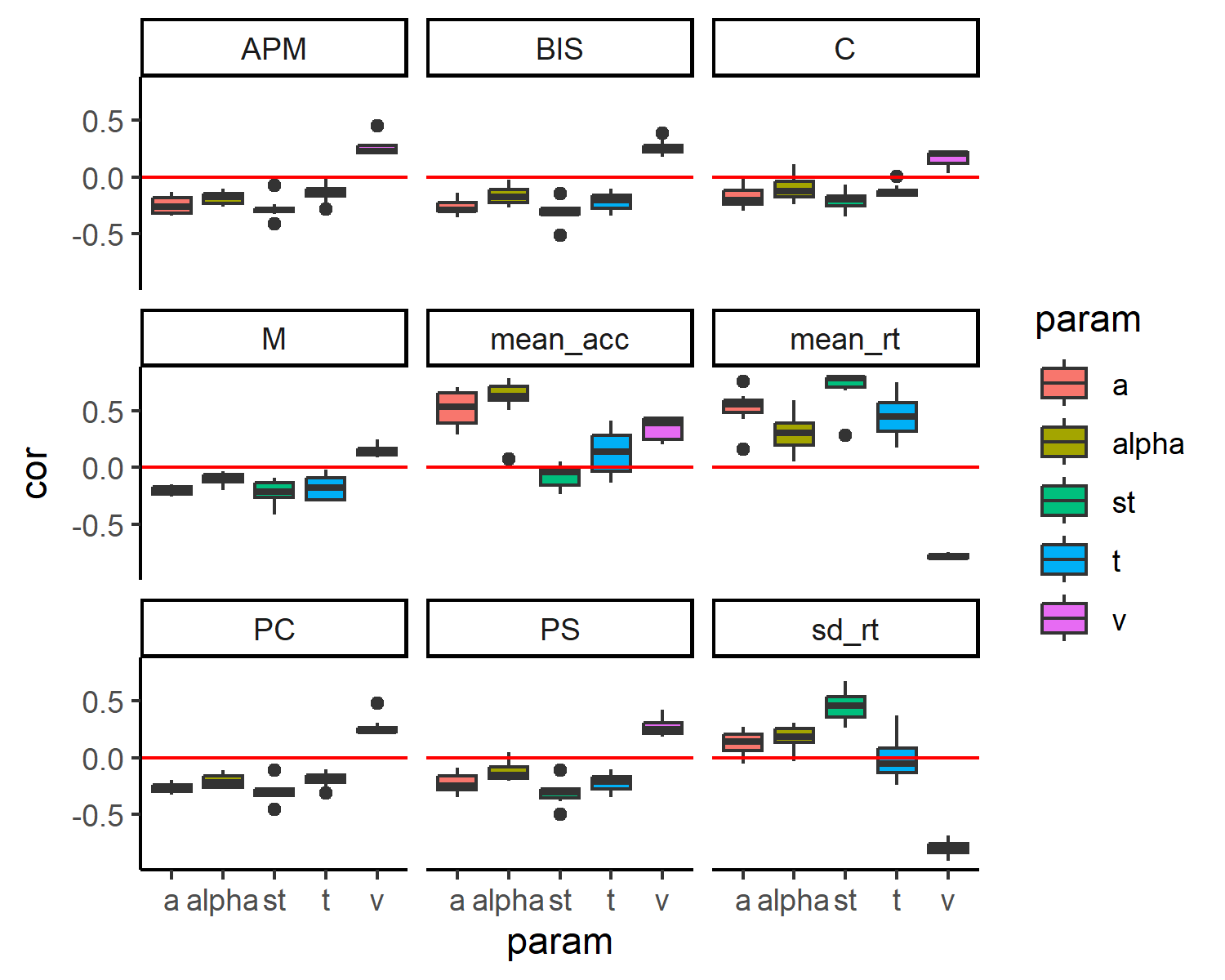
I excluded all response times which deviated more than 3 standard deviations from the mean response time for a given participant in a given condition. This removed 1.33 % of the data in the Hick task, 1.59 % of the data in the Posner task, and 1.68 % of the data in the Sternberg task. No further data-pruning was applied.

## Levy Flight Model estimation

The inference net showed no sign of bias in posterior estimation (see Figure 3). It also displayed acceptable recovery of true generating parameters (see Figure 4).

## Relationship of and g

On a bivariate level, displayed consistently negative correlations with all cognitive performance measures (see Figure 1). Averaged over tasks and conditions, showed a correlation with performance in the APM of -0.18 and -0.16 in the BIS.

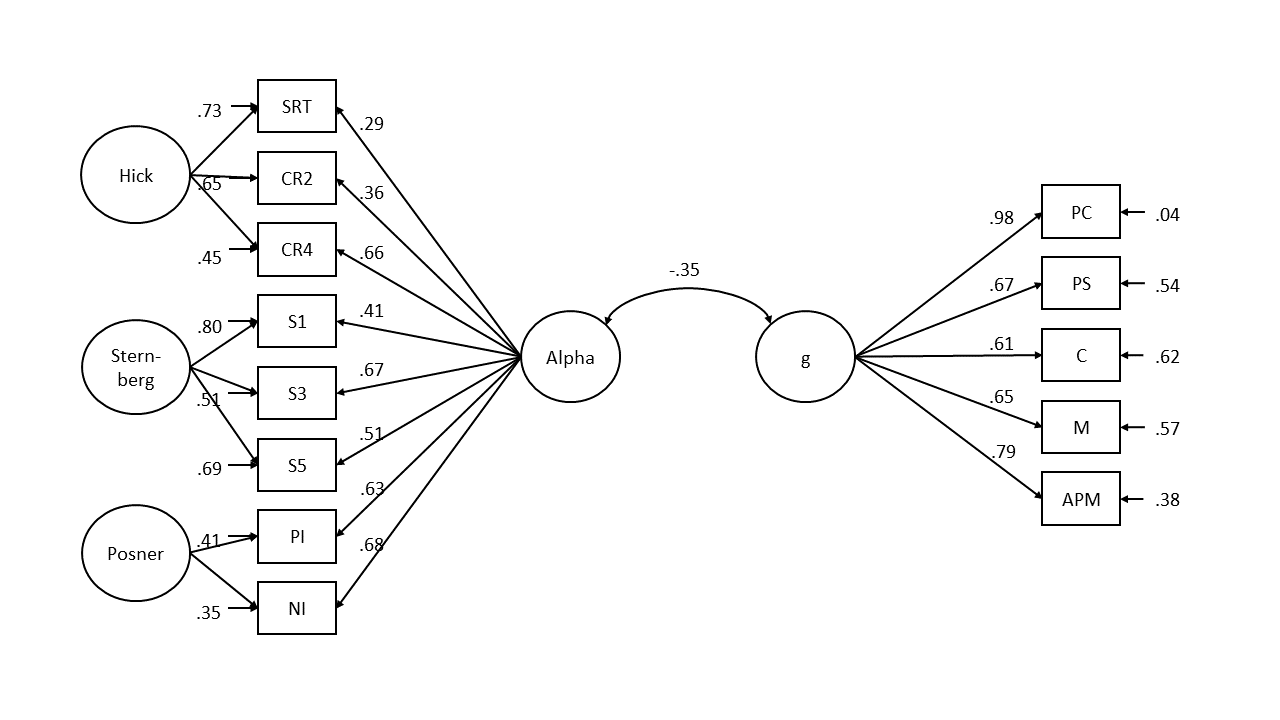


*Figure 1*

Bivariate Correlation between parameters and cognitive performance

## SEM

The model showed acceptable fit = 133.39, p < .001, CFI = 0.9, RMSEA = 0.08 and can be seen in Figure 2. On a latent level and *g* correlated to .



*Figure 2*

Structural equation model of the relationship between and intelligence

# Discussion

In this work, I applied a Levy-Flight Model ([Voss et al., 2019](#ref-voss2019sequential)) to data from three different tasks ([Schubert et al., 2017](#ref-schubert2017general)) in order to investigate the relationship between the stability parameter and intelligence. Bivariate correlations between and measures of mental speed were negative with an average correlation of -0.12. Following Schubert et al. ([2016](#ref-schubert2016trait)), we also investigated these results using SEM to obtain a latent trait measure for . The latent correlation between and intelligence was .

This negative correlation provides support for the efficient jumping hypothesis ([Wieschen et al., 2020](#ref-wieschen2020jumping)). More intelligent individuals display less stable evidence accumulation processes. This indicates that they make use of this instability in easy response time tasks in order to optimize speed and accuracy.

These results stand in contrast to earlier findings showing a negative correlation between variance in the response times and intelligence ([Doebler & Scheffler, 2016](#ref-doebler2016relationship)). More intelligent individuals display less variability. These differing results suggest that is not a measure of variability, despite the intuitive similarity between variability and instability. Additionally, shows stronger average correlations with mean response time 0.30 than with the standard deviation of response time 0.18. These correlations suggest that stability in the evidence accumulation process leads to higher average response times and slightly higher variability in the response times.

Future research should focus on investigating more difficult response time tasks. The relationship between and intelligence may change in tasks that require more higher-order processing. I expect that the functional benefits of instability in the evidence accumulation process are only present in simple tasks and will revert in more complex tasks ([Wieschen et al., 2020](#ref-wieschen2020jumping)).

Nonetheless, this work provides first insights into the functionality of stability in evidence accumulation processes. My results suggest that more intelligent individuals display less stability in order to improve the response time by allowing bigger jumps in evidence accumulation processes. Future work will focus on investigating this relationship in more complex tasks, where I expect the functionality of instability to decline.

# References

Aust, F., & Barth, M. (2023). *papaja: Prepare reproducible APA journal articles with R Markdown*. <https://github.com/crsh/papaja>

Doebler, P., & Scheffler, B. (2016). The relationship of choice reaction time variability and intelligence: A meta-analysis. *Learning and Individual Differences*, *52*, 157–166.

Jäger, A. O., Süß, H.-M., & Beauducel, A. (1997). *Berliner Intelligenzstruktur-Test: BIS-Test*.

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. Elsevier.

Laming, D. R. J. (1968). *Information theory of choice-reaction times.*

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>

Radev, S. T., Mertens, U. K., Voss, A., Ardizzone, L., & Köthe, U. (2020). BayesFlow: Learning complex stochastic models with invertible neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, *33*(4), 1452–1466.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*(5), 347–356.

Ratcliff, R., Thapar, A., & McKoon, G. (2011). Effects of aging and IQ on item and associative memory. *Journal of Experimental Psychology: General*, *140*(3), 464.

Raven, J. (1981). Manual for Raven’s progressive matrices and Mill Hill vocabulary scales. *(No Title)*.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>

Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H.-M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General*, *136*(3), 414.

Schubert, A.-L., Frischkorn, G. T., Hagemann, D., & Voss, A. (2016). Trait characteristics of diffusion model parameters. *Journal of Intelligence*, *4*(3), 7.

Schubert, A.-L., Hagemann, D., & Frischkorn, G. T. (2017). Is general intelligence little more than the speed of higher-order processing? *Journal of Experimental Psychology: General*, *146*(10), 1498–1512. <https://doi.org/10.1037/xge0000325.supp>

Schubert, A.-L., Hagemann, D., Voss, A., Schankin, A., & Bergmann, K. (2015). Decomposing the relationship between mental speed and mental abilities. *Intelligence*, *51*, 28–46.

Sheppard, L. D., & Vernon, P. A. (2008). Intelligence and speed of information-processing: A review of 50 years of research. *Personality and Individual Differences*, *44*(3), 535–551.

Talts, S., Betancourt, M., Simpson, D., Vehtari, A., & Gelman, A. (2018). Validating Bayesian inference algorithms with simulation-based calibration. *arXiv Preprint arXiv:1804.06788*.

Voss, A., Lerche, V., Mertens, U., & Voss, J. (2019). Sequential sampling models with variable boundaries and non-normal noise: A comparison of six models. *Psychonomic Bulletin & Review*, *26*, 813–832.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. <https://doi.org/10.21105/joss.01686>

Wieschen, E. M., Voss, A., & Radev, S. (2020). Jumping to conclusion? A lévy flight model of decision making. *The Quantitative Methods for Psychology*, *16*(2), 120–132.

# Appendix

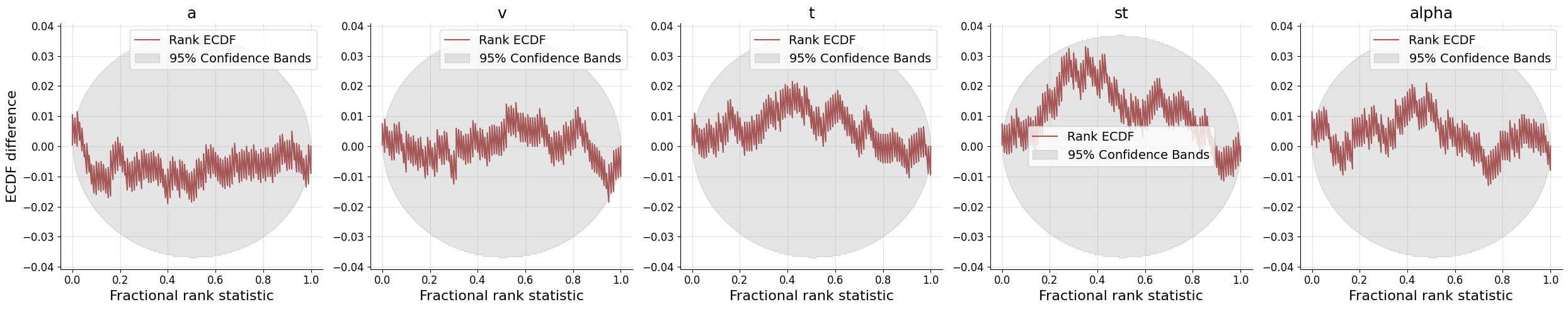


Figure 3: Empirical cumulative density functions of rank statistics

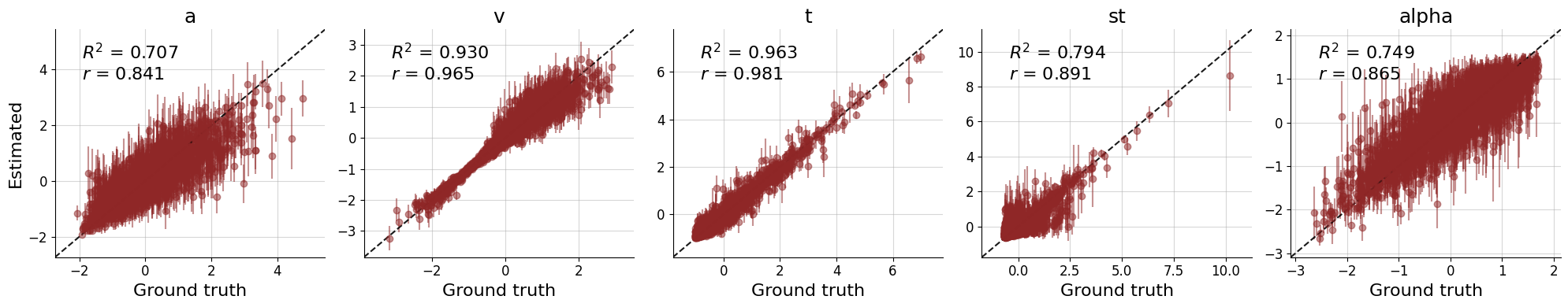


Figure 4: Recovery of true model parameters

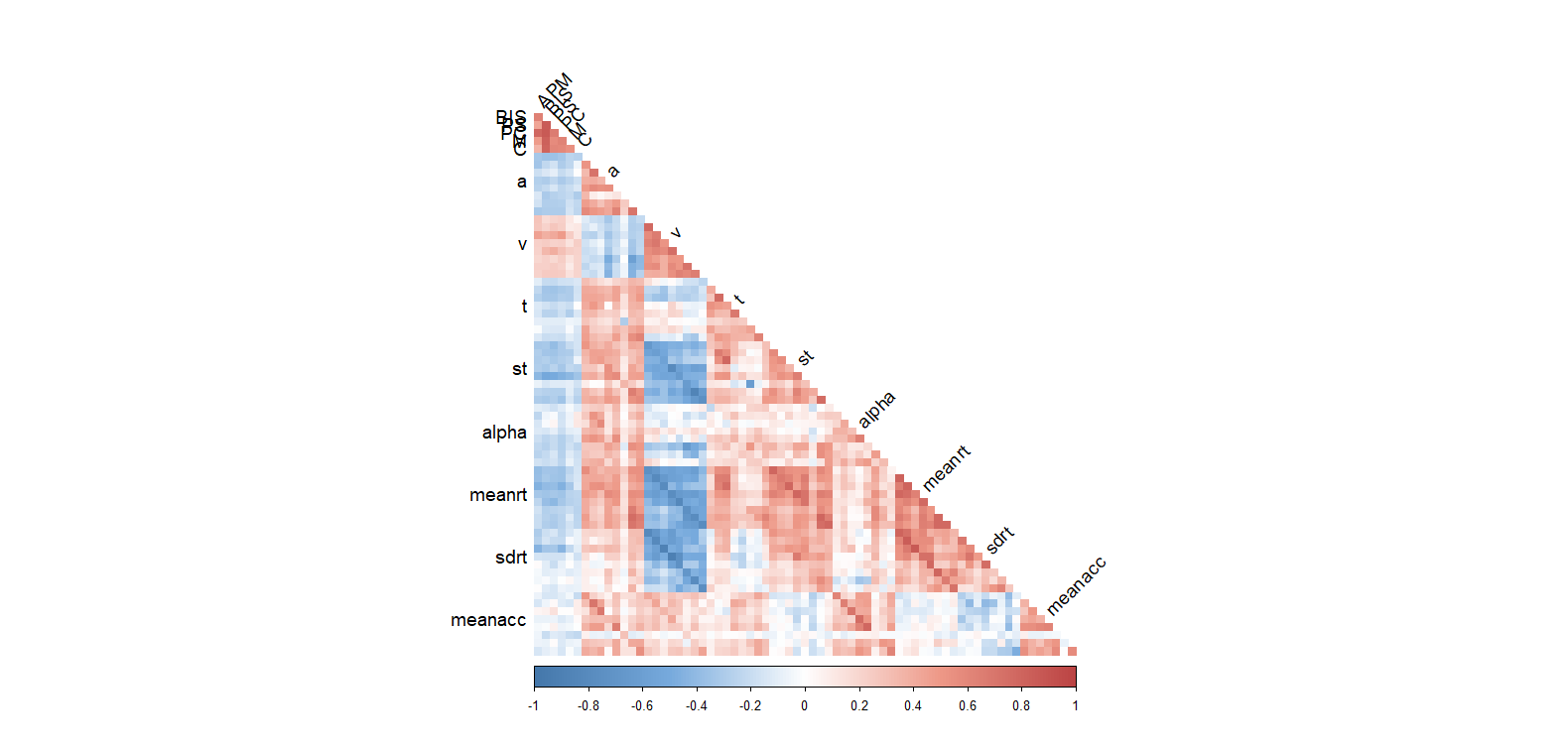


Figure 5: Correlation matrix of behavioral data and model parameters. Task and condition labels are omitted to improve readability

1. We, furthermore, used the R-packages *lavaan* (Version 0.6.18; [Rosseel, 2012](#ref-R-lavaan)), *papaja* (Version 0.1.2; [Aust & Barth, 2023](#ref-R-papaja)), and *tidyverse* (Version 2.0.0; [Wickham et al., 2019](#ref-R-tidyverse)). [↑](#footnote-ref-1)