

1           Attentional Control Data Collection: A Resource for Efficient Data Reuse

2                   Julia M. Haaf<sup>1</sup>, Madlen Hoffstadt<sup>1</sup>, & Sven Lesche<sup>2</sup>

3                               <sup>1</sup> University of Amsterdam

4                               <sup>2</sup> University of Heidelberg

5                               Version 1, October 2023

6                               Author Note

7           We are indebted to Arte Bischof for her thesis work on the initial SQL data base.

8           This work was supported in part by a Veni grant from the NWO (VI.Veni.201G.019)  
9 and a talent grant by Amsterdam Brain & Cognition (ABC.T09.0921) to J.M.H.

10          The authors made the following contributions. Julia M. Haaf: Conceptualization,  
11 Writing - Original Draft Preparation, Writing - Review & Editing; Madlen Hoffstadt:  
12 Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing; Sven  
13 Lesche: Conceptualization, Writing - Original Draft Preparation, Writing - Review &  
14 Editing.

15          Correspondence concerning this article should be addressed to Julia M. Haaf, Nieuwe  
16 Achtergracht 129B, 1018 WT Amsterdam, The Netherlands.. E-mail: [j.m.haaf@uva.nl](mailto:j.m.haaf@uva.nl)

## Abstract

One or two sentences providing a **basic introduction** to the field, comprehensible to a scientist in any discipline.

Two to three sentences of **more detailed background**, comprehensible to scientists in related disciplines.

One sentence clearly stating the **general problem** being addressed by this particular study.

One sentence summarizing the main result (with the words “**here we show**” or their equivalent).

Two or three sentences explaining what the **main result** reveals in direct comparison to what was thought to be the case previously, or how the main result adds to previous knowledge.

One or two sentences to put the results into a more **general context**.

Two or three sentences to provide a **broader perspective**, readily comprehensible to a scientist in any discipline.

*Keywords:* Open Data, Attentional Control, SQL

Word count: X

## Attentional Control Data Collection: A Resource for Efficient Data Reuse

Making data openly available has been a central demand by reformers since the start of the reproducibility crisis in psychology [REFS]. Fortunately, this demand has lead to a considerable increase in data availability. While only about 25% of data were shared after request in 2006 (Wicherts, Borsboom, Kats, & Molenaar, 2006), publicly sharing data upon publication is now more and more the norm. This cultural shift is also increasingly institutionalized. Universities and funding agencies prioritize open data, and some journals even mandate the publication of data with every published article (Sloman, 2015). In addition, technology like the Open Science Framework (OSF) and other data sharing facilities enable an easy process for researchers, further reducing barriers to share data.

Data sharing serves two goals: 1. To make the scientific process more transparent and enable error and fraud detection, and 2. to make the scientific process more efficient by allowing data reuse for different research projects. Current data sharing efforts, however, seemingly focus overwhelmingly on the first goal [REF Cruewell et al, 2023]. Whenever researchers complying with common data sharing procedures publish an article, they share the corresponding data on the OSF, ideally in a format that allows to redo the exact analyses reported in the article. The OSF repository is linked in the article, and readers may access the data through this link and check whether analysis code and shared data correspond to the results section in the article. This setup, while appropriate for the first goal of data sharing, ignores the second goal of data reuse.

To enable data reuse, data sharing needs to be approached differently. For example, consider a researcher (like the first author of the current paper) might be interested in the Stroop task (Stroop, 1935). The Stroop task is popular in cognitive psychology (MacLeod, 1991), so we may assume that many studies include this or similar tasks in their studies. Instead of running yet another Stroop experiment, the researcher decides to use existing data to explore their research question before designing a more targeted study. First, the

researcher needs to be able to find open Stroop task data. Currently, they could either search for papers on the topic and check whether open data are provided, or search directly via OSF or other data sharing servers. However, neither of these options is very promising as the vast majority of articles in the literature does not provide raw data and data sharing servers are not equipped with sufficient search options. Second, data sets need to be accessed easily and in a general, understandable format ready for reuse. There are data sharing formats that provide this structure [REF], but they are rarely used. Additionally, data are usually shared on the level necessary for the original analysis. In case of the Stroop task, shared data might provide the Stroop effect per participant, but for this new analysis the researcher needs trial-level data. So again, there is yet another barrier for data reuse.

We think it is necessary to provide a data sharing solution that solves the current problems and enables easy and efficient data reuse. Here, we propose to gather open data sets from a specific research area in an SQL data base. This process requires little to no work in addition to current data sharing policies from the authors of original papers, some work from the lab(s) setting up the data base, and little to no work from the researchers who wish to reuse open data. We describe the process and structure we used to set up a data base of attentional control tasks called the Attentional Control Data Collection (ACDC). The data base includes XXX data sets from XXX publications from tasks like the Stroop, Simon, and flanker tasks. Subsequently, we show how the data can be explored using a Shiny app and accessed using an R-package. In an example analysis, we assess the reliability of the included tasks. This section highlights how an open data base like ACDC can aid meta-analytic efforts as well as methodological innovation.

To provide a little history of the project, the Attentional Control Data Collection was inspired by a collection of open data sets from attentional control tasks by the Perception and Cognition Lab led by J. Rouder ([url](#)). Colleagues provided the first author and Rouder with data sets for their statistical work (Haaf & Rouder, 2017; Rouder, Kumar, & Haaf,

2023). To ensure that data sets were accessible, we gathered them in a github repository. However, there was little structure to the collection, and github repositories are neither stable entities nor are they designed as data storage. Here, we describe how a structured data collection can be achieved and which benefits it provides.

## SQLite Database

One of the most standard ways in computer science for storing data is using a Structured query language (SQL) data base. SQL allows to create, access and manipulate a structured data storage. SQL data bases consist of data tables and relations between these tables. There are many flavors of SQL data bases. Here, we decided to use an SQLite data base, a lightweight solution that allows us to store the entire data base in a single file of moderate size that can be downloaded by researchers for data reuse. In this section we describe the structure of the data base and the data currently included. Researchers who simply want to use ACDC may safely skip this section.

## Database Structure

SQL databases are composed of several data tables consisting of rows and columns. Each row in a data table has a primary key (essentially a row ID) which uniquely identifies it. To connect one table to another table in the database, tables may contain foreign keys which reference a unique row in another data table. In contrast to primary keys, these foreign keys can have duplicate values in the data table, as long as they are unique in their primary table. For instance, a study table may store information about all studies in a database where each row corresponds to a single study. Here the primary key is the study\_id. We can ensure that our database links each study to the publication it was published in by adding a foreign key called publication\_id. This foreign key references the unique identifier of the respective publication in a publication table. Figure 1 illustrates the relationship between study table and publication table. While publication\_id links to a single row in the publication table, it can occur several times in the study\_table as there can be several studies per publication.

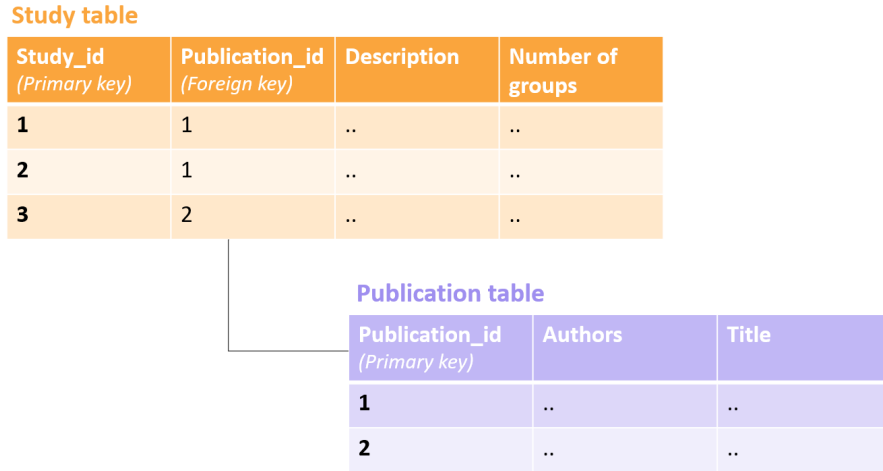


Figure 1. Illustrative example of using foreign and primary keys in a SQL database.

ACDC is adapted to the logic of publications consisting of one or multiple studies which in turn include one or several data sets. The whole structure of ACDC is shown in Figure 2. A *publication table* and a *study table* contain specific information about each publication and study, respectively. Each data set within a study stores trial-level information about a single attentional control task within a certain study. If a between-subject manipulation exists, our data base contains a separate data set for each group and each task. For instance, a study in which two groups (younger and older adults) who completed a Simon and a Stroop task would consist of four data sets (i.e., younger-Stroop, younger-Simon, older-Stroop, older-Simon) in the ACDC data base.

A *data set table* stores information about each data set, such as sample size, the number of within-participant manipulations, and whether a fixation cross was used. The *observation table* hold the trial-level attentional control task data (including response time and accuracy). The task type of each data set (i.e., Stroop, Simon, Flanker, negative priming, or other) and a description of which stimuli were presented in the task are documented in the *task table*.

Note that since the congruency of stimuli, i.e., whether response and stimulus attributes are compatible or incompatible, is part of every attentional control task, it is not

considered a separate within-participant manipulation in this database but is per default included in the observation table. Any additional within-participant manipulations are coded in the within ID column of the observation table. Further information about each condition of each data set (such as the percentage of congruent trials, mean response time, and mean accuracy) are recorded in the *within table*.

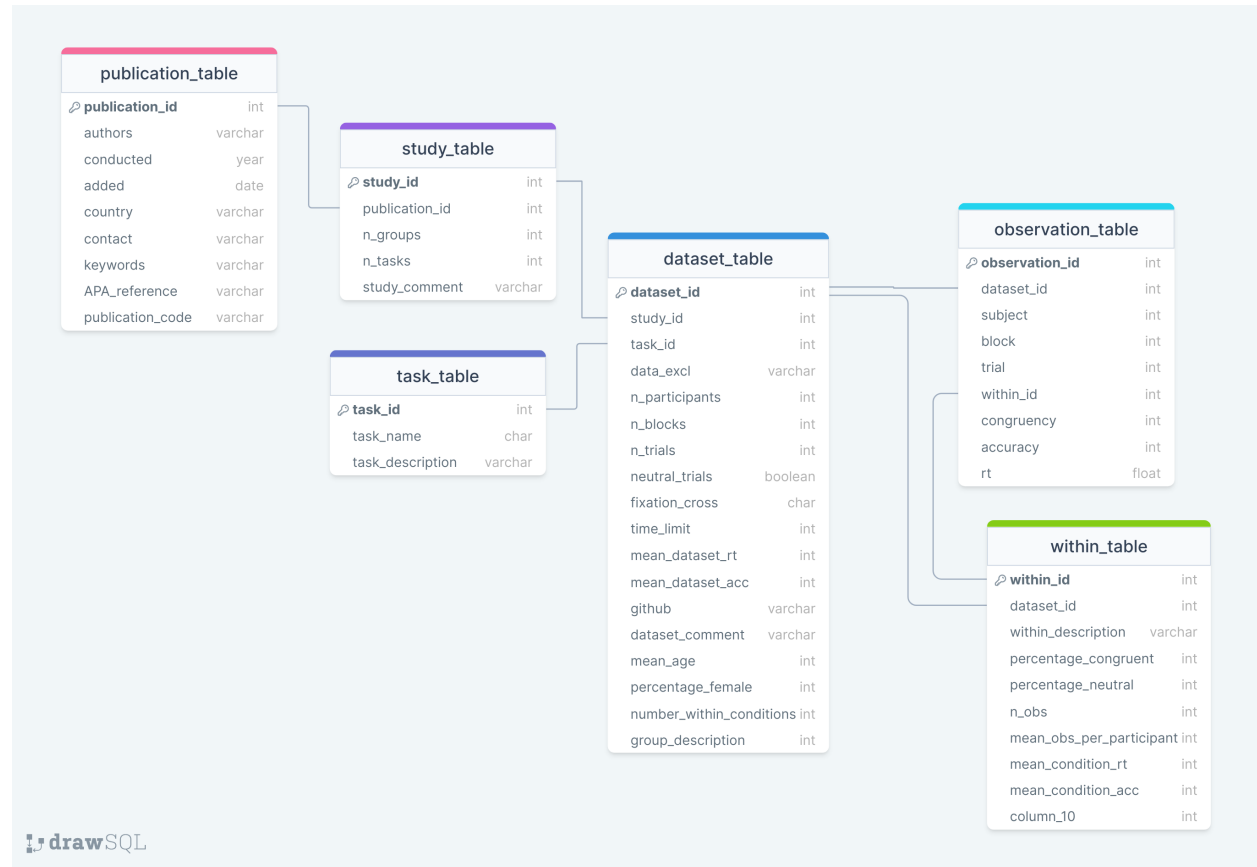


Figure 2. Structure of the ACDC database. Primary keys are indicated by the key symbol. References between data tables are illustrated through lines connecting columns across data tables. This overview includes the data type of each column: integers (int), numbers with decimal places (float), characters (varchar) and logical true/false values (Booleans).

## Included Data

Up to the date of submission of the manuscript, 54 data sets from 11 publications are included in the database. The full list of data sets and references is provided in the Appendix. The current database includes data sets from studies with an experimental as well as a correlational focus. The data contain  $10^6$  observations collected from 7486 participants.

We did not conduct a systematic search for data sets nor attempted to distribute a wider call for open data. Instead, we included the data sets that were already made available to the lab for previous projects, and added data sets from collaborators bit by bit. This approach was chosen to make the project feasible, and to first set up a working data base before large quantities of data are added.

## Contribute

*(TODO: Add a brief description on how researcher can add their own data.)*

## Accessing the Database

One advantage of SQLite databases is that they are simply a file that can be downloaded and locally accessed by anyone. Our database is provided in a github repository.<sup>1</sup> To access the database, researchers can download the file `acdc.db`, and use the SQLite tool of their choice. In addition, we build R-based tools to inspect, access, and select data from ACDC. We introduce these tools, a shiny app and an R package, subsequently.

## Shiny App

The easiest way to inspect the data is using our shiny app provided [here](#). The app allows to inspect all data sets or select data sets with certain specification using the filter box on the left. For example, if a researcher is interested in the flanker task, they may select all flanker data sets for closer inspection. After selection, researchers can choose between an overview of the included data sets, some descriptive statistics, and descriptive plots (see Figure XXX). If they want to further analyze the data, they can download the data via the “get the data tab”, either directly as `csv` file or using the provided R-code.

*(TODO: Add screenshot from the shiny app as figure.)*

---

<sup>1</sup> The newest version can be accessed via <https://github.com/jstbcs/acdc-database/blob/main/acdc.db>, the version at the time of submission can be found [here](#). *(TODO: REPLACE WITH CURRENT COMMIT.)*



Note that all descriptive statistics in the data base are aggregated across congruency conditions. This was a deliberate choice when designing the shiny app. By withholding information about the effect of interest, researches can inspect and select the data based on varying characteristics of the data set (including distributional properties), but remain unbiased as to the most relevant outcome variable. We hope that researchers can then formulate (and perhaps preregister) hypotheses about their reanalysis without much hindsight bias.

*(TODO: Add example descriptive plots.)*

*Julia*

## R-Package

*(SVEN: Add an overview or the R package.)*

*Julia*

```
install.packages("acdcquery") # install package

library(acdcquery) # load package
```

## Queries and Output

*(SVEN: Show and explain an example query with several arguments, selection of the target table, etc.)*

*Julia*

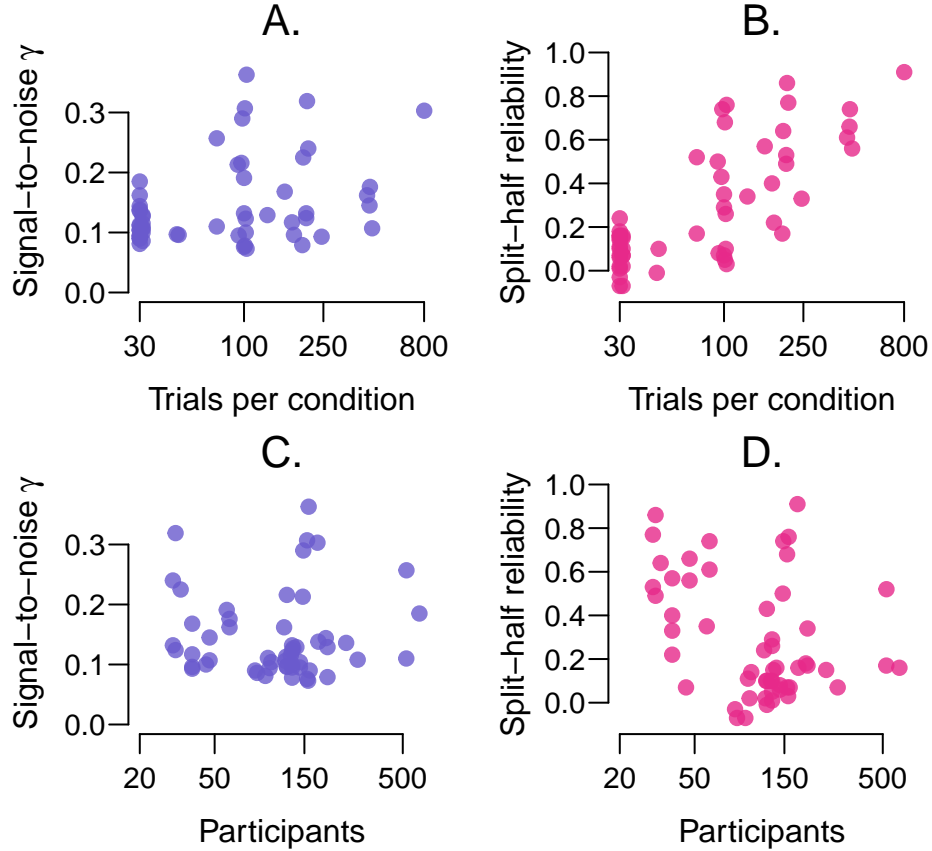
## Reliability of ACDC Tasks

To illustrate how ACDC can be used for reanalysis, and to further assess the reliability of attentional control tasks included, we provide an example data analysis. The reliability of attentional control tasks such as the Stroop task has been critically discussed in recent literature (Draheim, Mashburn, Martin, & Engle, 2019; Hedge, Powell, & Sumner, 2018; Rey-Mermet, Gade, & Oberauer, 2018; Rouder & Haaf, 2019; Rouder et al., 2023). The current analysis can be understood as an extension of to the analysis of 24 data sets conducted by Rouder et al. (2023) (see their Table 2).

Here, we rely on methodological development by Rouder and colleagues (Rouder & Haaf, 2019; Rouder et al., 2023) to survey the reliability of data sets in ACDC. The authors have made suggestions to update the analysis of reliability of individual differences in cognitive tasks based on two simple observation. First, conventional measures of reliability such as the split-half reliability or intra-class correlation are not well suited for cognitive tasks as they are dependent on the number of trials. The reason for this dependency is that observed between-subjects variability of a target effect is a combination of true individual differences and sample noise. If sample noise is large, as is the case in these tasks, reliability can be drastically reduced with a small numbers of trials. Second, unlike with psychometric tests, the number of trials of cognitive tasks is oftentimes chosen arbitrarily, as a matter of preference by the lab conducting the experiment, making it difficult to compare reliability across labs or across tasks (Rouder & Haaf, 2019).

The relationship between common reliability estimates and trial number is illustrated in Figure 3B. The figure shows the split-half reliability coefficients for all ACDC datasets as a function of the (log) number of trials per condition. The correlation between split-half-reliability and log trial number is  $r = 0.72$ , 95% HDI  $[0.59, 0.84]$ ,  $\text{BF}_{10} = 2.52 \times 10^8$ . Figure 3D shows the relationship between split-half reliability coefficients and the (log) number of participants. Theoretically, there should be no relationship. However, many studies with a large number of participants have in turn a low number of trials for reason of practicality. Therefore, empirically we do find a negative relationship between the log number of participants and split-half reliability,  $r = -0.31$ , 95% HDI  $[-0.54, -0.08]$ ,  $\text{BF}_{10} = 6.15$ .

To counter the lack of comparability of conventional measures of reliability, Rouder et al. (2023) proposed a measure of reliability that is derived from hierarchical modeling of the nested data structure. This modeling approach allows to separately estimate true individual variability and sample noise. As a result, we may determine how much true variability there is relative to variability due to sample noise. This ratio, termed  $\gamma^2$ , can be expressed as



*Figure 3.* Reliability estimates as a function of number of trials and participants. A. The signal-to-noise ratio is independent of the (log) number of trials per condition but has a small positive correlation. B. The split-half reliability is highly correlated with the (log) number of trials per condition. C. The signal-to-noise ratio is independent of the (log) number of participants. D. The split-half reliability has a slight negative relationship with (log) number of participants due to the fact that large-participants studies commonly use a smaller number of trials per task.

$$\gamma^2 = \frac{\sigma_{\text{true}}^2}{\sigma_{\text{noise}}^2},$$

where the numerator is the true variance of individuals' effects and the denominator is the variance due to sample noise. For the purpose of individual differences research,  $\sigma_{\text{true}}^2$  represents the signal and  $\sigma_{\text{noise}}^2$  represents noise, therefore, Rouder and colleagues labeled  $\gamma^2$  as the signal-to-noise ratio.

Figure 4A. shows the ratio of signal and noise standard deviations,  $\gamma$  for all data sets

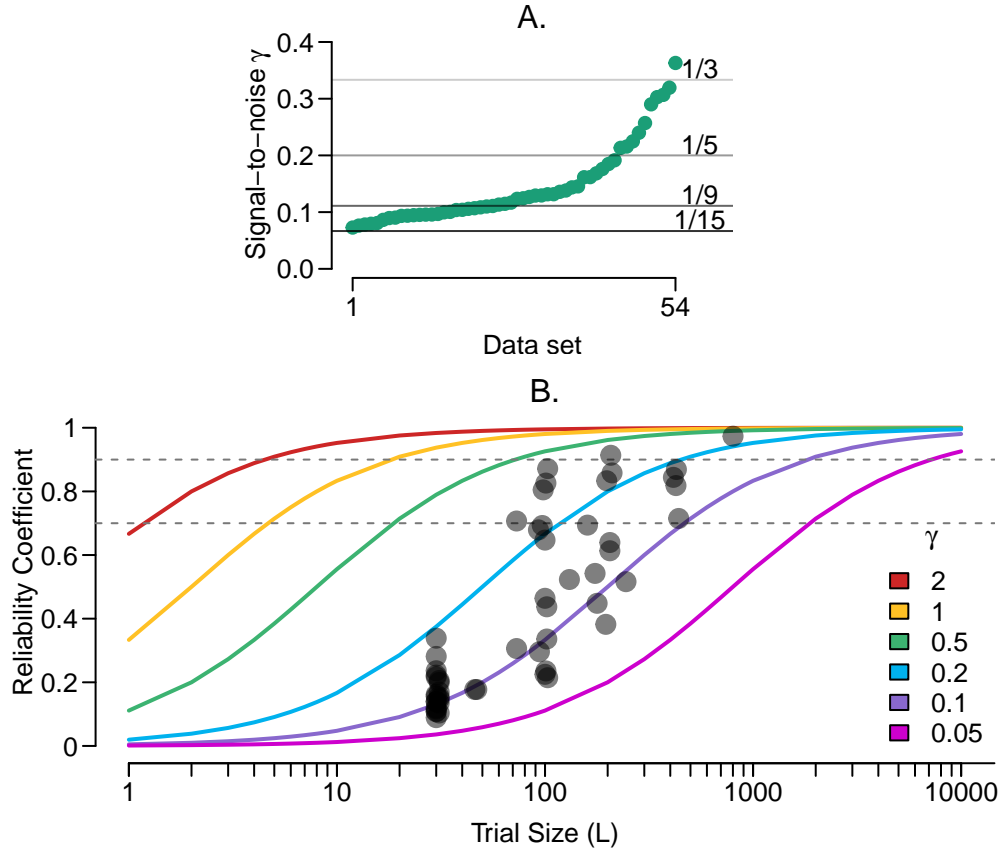


Figure 4. Signal-to-noise ratio and reliability. A. Signal-to-noise ratio  $\gamma$  for all data sets ordered from smallest to largest. B. Reliability as a function of number of trials for all data sets in ACDC. The lines represent different signal-to-noise ratios. For a fixed  $\gamma$ , the relationship between the number of trials per condition and the reliability is deterministic.

ordered from smallest to largest. The median of  $\gamma$  is  $Med = 0.12$  corresponding to a ratio of roughly 1 to 8.5. That is, for every unit of signal there are 8.5 units of noise.

Using the ACDC datasets, we can also illustrate the independence of  $\gamma$

## Discussion

## Appendix

Table 1  
*All datasets currently in ACDC.*

ID	Task name	I	K	$\gamma$	$\rho$	$\rho^*$
Chetverikov et al. (2017)						
1	flanker	58	100	0.191	0.35	0.51
Enkavi et al. (2019)						
2	stroop	523	73	0.257	0.52	0.68
3	simon	522	73	0.110	0.17	0.29
Hedge et al. (2018)						
4	stroop	47	439	0.107	0.56	0.72
5	flanker	47	426	0.145	0.66	0.80
6	stroop	60	428	0.176	0.74	0.85
7	flanker	60	414	0.162	0.61	0.75
Kucina et al. (2023)						
8	other	31	205	0.124	0.49	0.65
9	other	30	205	0.132	0.53	0.68
10	other	31	207	0.319	0.86	0.92
11	flanker	33	198	0.225	0.64	0.77
12	other	30	210	0.240	0.77	0.87
Löffler, Frischkorn, Hagemann, Sadus, and Schubert (2022)						
13	stroop	148	98	0.290	0.74	0.85
14	flanker	147	93	0.213	0.50	0.67
15	negative priming	142	94	0.095	0.08	0.13
Ebersole et al. (2016)						
16	stroop	160	30	0.090	0.07	0.13
17	stroop	288	31	0.108	0.07	0.13
18	stroop	136	31	0.129	0.16	0.27
19	stroop	195	30	0.144	0.18	0.30
20	stroop	93	30	0.081	-0.07	-0.13
21	stroop	117	30	0.162	0.24	0.38
22	stroop	127	31	0.115	0.10	0.18
23	stroop	142	30	0.104	0.06	0.12
24	stroop	131	31	0.127	0.15	0.25
25	stroop	82	30	0.090	-0.03	-0.05
26	stroop	98	30	0.094	0.02	0.04
27	stroop	119	31	0.106	0.02	0.03
28	stroop	100	30	0.104	0.14	0.24
29	stroop	120	30	0.113	0.10	0.18
30	stroop	613	30	0.185	0.16	0.28
31	stroop	96	30	0.111	0.11	0.20
32	stroop	178	30	0.138	0.16	0.26
33	stroop	84	31	0.086	-0.07	-0.13
34	stroop	45	31	0.100	0.07	0.12
35	stroop	250	30	0.136	0.15	0.25

Table 1 continued

ID	Task name	I	K	$\gamma$	$\rho$	$\rho^*$
36	stroop	129	30	0.095	0.01	0.02
Pratte, Rouder, Morey, and Feng (2010)						
37	simon	38	245	0.093	0.33	0.48
38	stroop	38	160	0.168	0.57	0.72
39	simon	38	174	0.117	0.40	0.56
40	stroop	38	178	0.096	0.22	0.34
Rey-Mermet et al. (2018)						
41	stroop	129	102	0.123	0.26	0.41
42	stroop	129	100	0.132	0.29	0.45
43	flanker	129	102	0.100	0.10	0.17
44	flanker	129	101	0.078	0.05	0.09
45	stroop	155	101	0.307	0.68	0.81
46	stroop	155	100	0.076	0.07	0.12
47	flanker	158	103	0.363	0.76	0.86
48	flanker	157	103	0.073	0.03	0.05
Stahl et al. (2014)						
49	stroop	199	131	0.129	0.34	0.50
50	flanker	199	196	0.079	0.17	0.29
Tang, Bugg, Snijder, Conway, and Braver (2023)						
51	stroop	176	801	0.303	0.91	0.95
Von Bastian, Souza, and Gade (2016)						
52	stroop	121	47	0.096	0.10	0.18
53	simon	121	97	0.216	0.43	0.60
54	flanker	121	46	0.097	-0.01	-0.01

*Note.*  $\gamma$  = signal-to-noise ratio;  $I$  = number of participants;  $K$  = number of trials per condition;  $\rho$  = split-half reliability coefficient;  $\rho^*$  = Spearman-Brown corrected reliability coefficient.

## References

- Chetverikov, A., Iamschinina, P., Begler, A., Ivanchei, I., Filippova, M., & Kuvaldina, M. (2017). Blame everyone: Error-related devaluation in eriksen flanker task. *Acta Psychologica*, 180, 155–159.
- Draheim, C., Mashburn, C. A., Martin, J. D., & Engle, R. W. (2019). Reaction time in differential and developmental research: A review and commentary on the problems and alternatives. *Psychological Bulletin*, 145(5), 508.
- Ebersole, C. R., Atherton, O. E., Belanger, A. L., Skulborstad, H. M., Allen, J. M., Banks, J. B., et al.others. (2016). Many labs 3: Evaluating participant pool quality across the academic semester via replication. *Journal of Experimental Social Psychology*, 67, 68–82.
- Enkavi, A. Z., Eisenberg, I. W., Bissett, P. G., Mazza, G. L., MacKinnon, D. P., Marsch, L. A., & Poldrack, R. A. (2019). Large-scale analysis of test–retest reliabilities of self-regulation measures. *Proceedings of the National Academy of Sciences*, 116(12), 5472–5477.
- Haaf, J. M., & Rouder, J. N. (2017). Developing constraint in Bayesian mixed models. *Psychological Methods*, 22(4), 779–798.
- Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, 50, 1166–1186.
- Kucina, T., Wells, L., Lewis, I., Salas, K. de, Kohl, A., Palmer, M. A., ... Heathcote, A. (2023). Calibration of cognitive tests to address the reliability paradox for decision-conflict tasks. *Nature Communications*, 14(1), 2234.
- Löffler, C., Frischkorn, G. T., Hagemann, D., Sadus, K., & Schubert, A.-L. (2022). The common factor of executive functions measures nothing but speed of information uptake.
- MacLeod, C. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109, 163–203.
- Pratte, M. S., Rouder, J. N., Morey, R. D., & Feng, C. (2010). Exploring the differences in

distributional properties between stroop and simon effects using delta plots. *Attention, Perception, & Psychophysics*, 72, 2013–2025.

Rey-Mermet, A., Gade, M., & Oberauer, K. (2018). Should we stop thinking about inhibition? Searching for individual and age differences in inhibition ability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(4), 501.

Rouder, J. N., & Haaf, J. M. (2019). A psychometrics of individual differences in experimental tasks. *Psychonomic Bulletin and Review*, 26(2), 452–467. Retrieved from <https://doi.org/10.3758/s13423-018-1558-y>

Rouder, J. N., Kumar, A., & Haaf, J. M. (2023). *Why many studies of individual differences with inhibition tasks may not localize correlations*. *Psychonomic Bulletin and Review*. Retrieved from <https://doi.org/10.3758/s13423-023-02293-3>

Sloman, S. A. (2015). Opening editorial: The changing face of cognition. *Cognition*, 135, 1–3.

Stahl, C., Voss, A., Schmitz, F., Nuszbaum, M., Tüscher, O., Lieb, K., & Klauer, K. C. (2014). Behavioral components of impulsivity. *Journal of Experimental Psychology: General*, 143(2), 850.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662.

Tang, R., Bugg, J. M., Snijder, J.-P., Conway, A. R., & Braver, T. S. (2023). The dual mechanisms of cognitive control (DMCC) project: Validation of an online behavioural task battery. *Quarterly Journal of Experimental Psychology*, 76(7), 1457–1480.

Von Bastian, C. C., Souza, A. S., & Gade, M. (2016). No evidence for bilingual cognitive advantages: A test of four hypotheses. *Journal of Experimental Psychology: General*, 145(2), 246.

Wicherts, J. M., Borsboom, D., Kats, J., & Molenaar, D. (2006). The poor availability of psychological research data for reanalysis. *American Psychologist*, 61(7), 726–728.

Retrieved from <http://wicherts.socsci.uva.nl/datasharing.pdf>