Introducing the Truth Effect Database (TED) - an open trial-level database promoting FAIR data in truth effect research

Sven Lesche[1] & Annika Stump[2]

[1] Ruprecht-Karls-University Heidelberg

[2] Albert-Ludwigs-University Freiburg

Author Note

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: keyword, more keyword, crazy keyword

Word count: X

Introducing the Truth Effect Database (TED) - an open trial-level database promoting FAIR data in truth effect research

Introduction

The open science movement has emerged in response to growing concerns about the replicability, transparency, and efficiency of scientific research (Ioannidis, 2005; Munafò et al., 2017; Nosek et al., 2015). At its core, open science promotes practices that make research processes and outputs accessible, verifiable, and reusable. Practices such as preregistration, data sharing, and open materials aim to increase the trustworthiness of findings and to foster a more cumulative, collaborative scientific enterprise. These practices, coupled with platforms like the Open Science Framework - OSF (Foster & Deardorff, 2017), have played a pivotal role in advancing openness through tools that support study registration, version control, and public sharing of data and materials. Transparency, once rare, is increasingly becoming the norm (Bauer, 2022).

However, while transparency has improved markedly, efficiency and reusability have lagged behind. Research materials and datasets are often siloed within individual project pages, tailored to host single studies. As a result, even when data are technically open, they are rarely standardized in ways that facilitate reuse, integration, or cumulative analysis. Current open data practices focus primarily on the first two aspects of the FAIR (Findable, Accessible, Interoperable, and Reusable) principles of data sharing (Wilkinson et al., 2016), while neglecting the latter. For example, the OSF is well-suited to documenting and finding individual projects, but its structure does not support the aggregation or harmonization of data across multiple related studies. In addition to structural limitations, researchers' practices like inconsistent file formats, project-specific codebooks, and divergent analytic pipelines continue to hinder interoperability.

This inconsistency not only limits replicability, but also reusability. Should a researcher be interested in estimating an effect size, developing a new computational model or simply re-analyzing existing data instead of collecting an entirely new sample, they have

to comb through OSF, pray for existing codebooks and hope for decipherable raw data. This is both tedious and often unsuccessful (Crüwell et al., 2023; Hardwicke et al., 2021). However, reusability greatly improves scientific efficiency. It saves resources otherwise wasted by planning and collecting an entirely new set of data and helps efficiently allocate public resources as well as participants' and researchers' time.

We argue that to make research outputs genuinely reusable, not just transparent, structure is essential. Reusability requires standardized organization: consistent variable naming, common data schemas, uniform documentation practices, and clear, accessible codebooks. In neuroscience, this principle has been successfully realized through the Brain Imaging Data Structure - BIDS (Gorgolewski et al., 2016), which provides a framework for organizing and annotating neuroimaging data. BIDS has enabled not just more efficient sharing, but also reproducible pipelines, automated analyses, and collaborative efforts at scale (Poldrack et al., 2024).

Structured approaches like BIDS are now gaining traction in cognitive psychology. The Attentional Control Data Collection (ACDC; **Haaf et al., in press**) introduced a trial-level database for data from attentional control experiments, providing standardized variable names, metadata, and analysis-ready formats that streamline cross-study comparisons. Similarly, the FEARBASE project is building a large-scale, open-access repository for fear conditioning studies, adopting a shared structure to ensure comparability and long-term usability (**Lohnsdorf & Ehlers, 2025**). Crucially, there is increasing institutional support for the development of infrastructure enabling data reuse. For example, the German research foundation *DFG* issued an open funding call for projects developing infrastructure for scientific data management (DFG; *Förderprogramm Informationsinfrastrukturen für Forschungsdaten*). These projects spearhead a changing culture of data-reusability but so far represent the exception, rather than the norm.

We believe that there is great potential in the development of structured databases integrating trial level data. Large and structured collections of raw data have several key

applications, extending the general principle of data-reusability. They enable living meta-analyses that can update automatically as new data are added; allow researchers to find relevant raw data based on task characteristics or participant variables; facilitate the straightforward replication of published findings; and enable power analyses using pooled datasets from comparable studies.

Beyond these direct applications, structured datasets also support the development of new methods and models. Trial-level repositories allow for benchmarking existing analytic tools, training machine learning models, and exploring novel statistical approaches.

The present study

In this spirit, we introduce a centralized, trial-level database for the illusory truth effect, drawing on resources and lessons learned from developing ACDC **cite!!**. The truth-effect database (TED) is designed not only with structured organization in mind, but also with an emphasis on ease of use. Particular attention has been given to lowering the barrier for data submission through an intuitive entry-mask, as well as to enabling simple data extraction via an R package. Our aim is to create a living, extensible resource that supports both contributors and users: researchers can add new studies with minimal friction, and others can search, filter, and analyze trial-level data without having to clean or realign disparate datasets. By combining structure with usability, this resource is intended to foster cumulative research on the illusory truth effect and to serve as a model for reusable psychological data infrastructures more broadly.

The illusory truth effect

Here a section why we chose the ITE, what it is and why its important Hint at the Innenministerium

Method

The Truth Effect Database (TED)

Resources to build the database, restructure raw data, build the website and integrate the submitted data can be found on Github (https://github.com/SLesche/truth-effect-database).

The database is implemented using Structured Query Language (SQL), specifically *SQLite* (Hipp, 2020). A key advantage of SQLite is its serverless architecture—data is stored in a single, shareable file, allowing researchers to download, interact with, and modify the database locally without the need for a dedicated database server.

As a relational database system, SQLite supports the use of multiple interconnected tables. This structure enables efficient organization of complex data relationships. For instance, each publication may include several studies, which themselves involve multiple experimental conditions. By organizing data across discrete, normalized tables—rather than in a single flat file—the database minimizes redundancy and enhances both storage efficiency and query performance.

The design of our database is illustrated in Figure 1. Broadly, we make use of a table for each part of data related to truth effect experiments. Meta-data concerning the publications themselves as the highest order are stored in a *publication_table* and raw data at the lowest level in the *observation_table*. In addition, we included information on the study, the conditions, the procedure, the statements used, their origin, and additional variables collected in the experiment in respective tables.

We argue that while the use of a relational database adds some complexity, it also introduces an intuitive naming system and structure for variables of interest. Importantly, our goal is that variable names and their table location are the only understanding that users need to have in order to interact with the database. To this end, we have developed tools that require little to no understanding of the database structure or SQL in order to submit and extract data from the database.
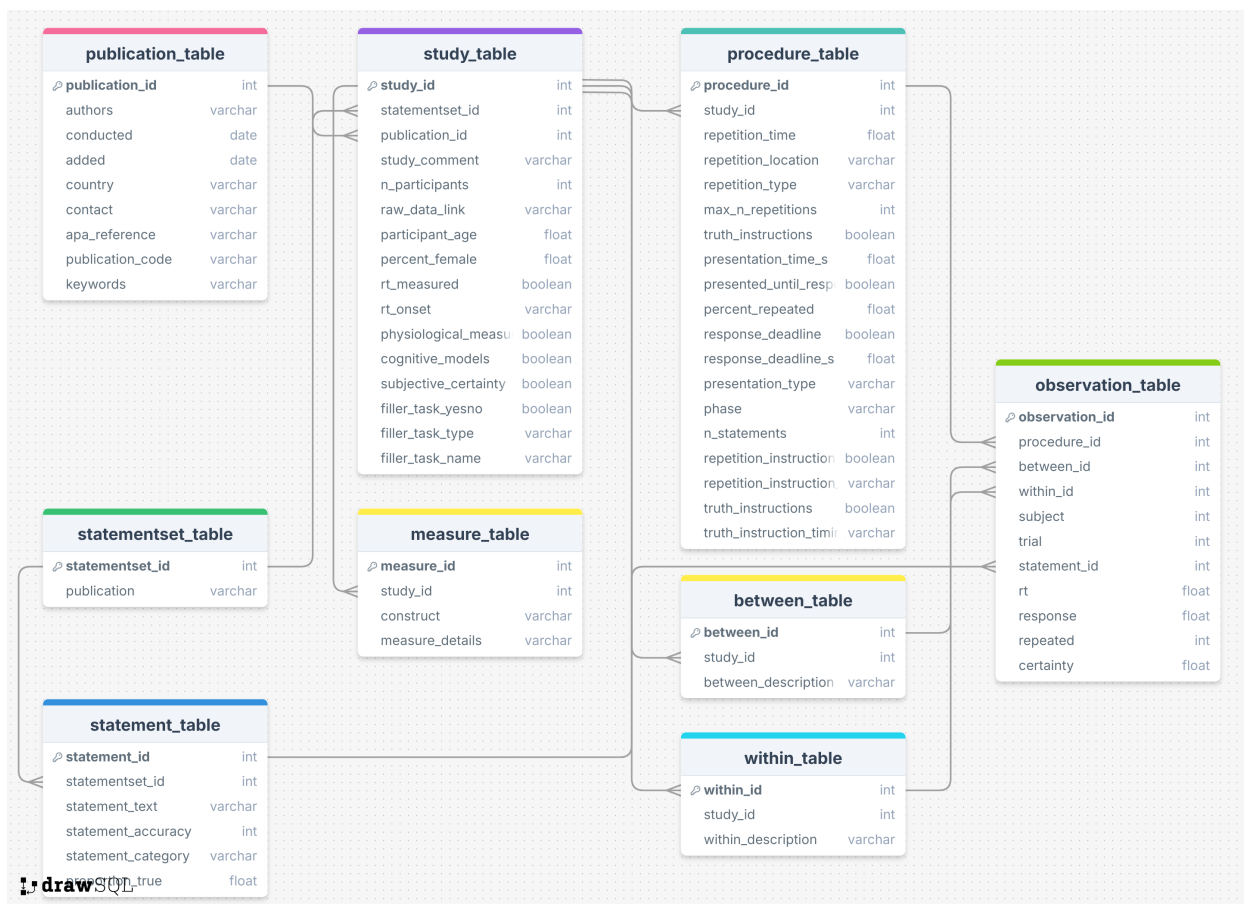
Figure 1

Overview of TED Structure

Data Submission

To ease the process of submitting data to the database, we built a website (https://slesche.github.io/truth-effect-database/) that guides users through the submission process and checks submitted information for inconsistencies or errors. This website is designed to minimize both the effort of submitting new data to the database as well as identify potential errors in the submission process.

Data Extraction

We provide the R package *acdcquery* (Lesche, 2025) to simplify the process of extracting data. This package provides functions to facilitate connection to the database and extraction of data. Users can define filter arguments using `add_argument()` and

request specific variables from any table in the database using `query_db()`.

For example, users may request all available trial-level data from studies with greater than 200 participants from the test phase using the code below:

```r
library(acdcquery)


conn <- connect_to_db("path/to/ted.db")


arguments <- list() %>%
  add_argument(
    conn,
    "n_participants",
    "greater",
    200
  ) %>%
  add_argument(
    conn,
    "phase",
    "equal",
    "test"
  )


results <- query_db(
  conn,
  arguments,
  target_vars = "default", # Returns all variables in target_table
  target_table = "observation table"
)
```

The R package will automatically construct the necessary SQL query and return the filtered data with the selected columns to the user. A detailed user guide on using the R package to interact with our database can be found on the databases GitHub (https://github.com/SLesche/truth-effect-database).

Data Selection

We included only data from studies specifically focused on the repetition-based illusory truth effect. To qualify for inclusion, studies had to involve participants making truth judgments about visual or auditory stimuli, with a subset of those stimuli having been presented previously to allow for repetition effects. Our data selection process was based in part on the work of Henderson et al. (2022), who identified 17 publications with accessible raw data. Additionally, we contacted colleagues directly to request openly available datasets, further expanding the scope of included studies.

Data Analysis

To illustrate the analytical potential of TED, we conducted a multilevel model predicting truth judgments, incorporating both fixed and random effects of repetition at the statement, subject, and procedure (i.e., study) levels. The scale and structure of TED enable us to estimate the variance in the repetition effect simultaneously across all three levels. This approach allows us to determine whether the size of the truth effect varies more substantially across individuals, across experimental procedures, or across the specific statements used.

Results

```
## [1] "Querying through observation table. Query times may be longer."
```
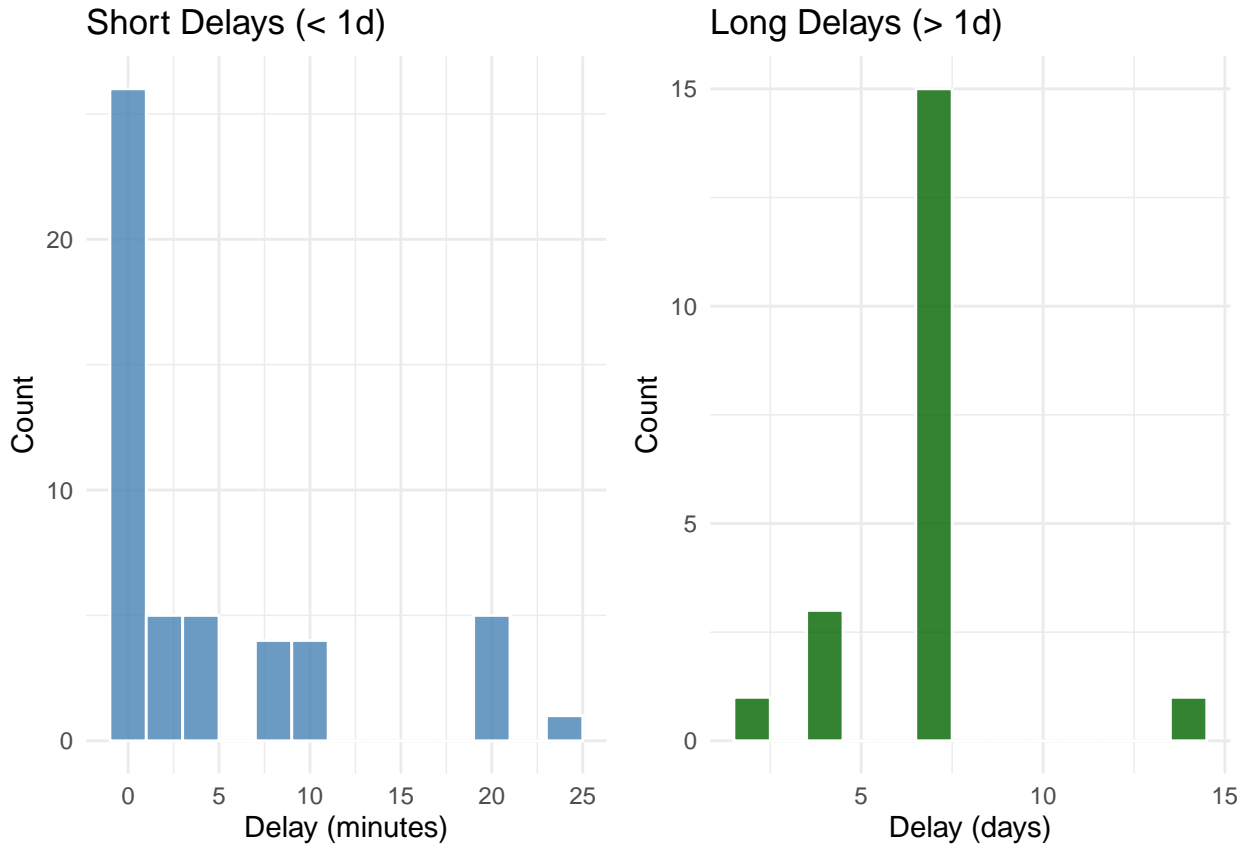
All analysis were conducted using R [Version 4.2.2; R Core Team (2022)][1].

.

In the current version of the manuscript, we included 48 studies from 21 publications, spanning 9983 participants contributing 525844 trials. A complete list of the included publications can be found in Table **table with refs, studies, trials here**.

Sample composition ranged from 29 to 949 participants. On average, studies included 212.17 participants ($\mu_{age} = 32.49, \sigma_{age} = 7.20$). An overview of the rating scale usage for truth judgments and the use of a filler task over all included studies can be found in Figure 2.

On average, studies employed 66.49 ($SD = 41.51$) statements per participant in the

---

[1] We, furthermore, used the R-packages *acdcquery* (Version 1.0.1; Lesche, 2025), *brms* (Bürkner, 2017, 2018, 2021), *knitr* (Version 1.42; Xie, 2015), *lme4* (Version 1.1.32; Bates et al., 2015), *papaja* (Version 0.1.1; Aust & Barth, 2022), and *tidyverse* (Version 2.0.0; Wickham et al., 2019).

Breakdown of Study Variables
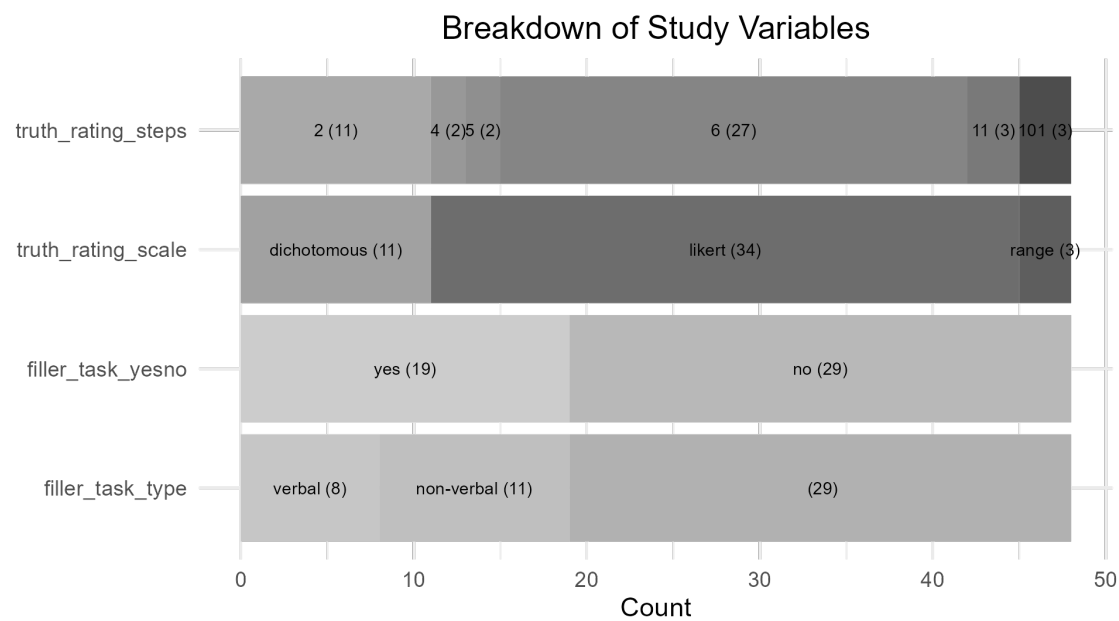
Figure 2

Overview of Study-related variables in TED

judgment session and in 85.71 % of procedure settings exactly 50% of statements were repeated. Of 70 judgment phases, 71.43 % were conducted on the same day as the exposure phase. The average delay between exposure and judgment phase if both were conducted on the same day was 4.80 minutes. The average delay between exposure and judgment phase, given the judgment phase was conducted at least one day after the exposure phase, was 6.65 days. An overview of additional variables pertaining to the procedure of the included studies can be found in Figure 3.

Detailed information on the statements presented is available for 45 out of 48 studies. Data on the accuracy of a statement is available for 83.58 % of trials, the detailed statement text is available for 69.90 % of trials, and response times are available for 24.31 % of trials.

Multi-level Modelling

To illustrate the benefits of our large collection of trial-level data, we applied multilevel models predicting truth judgments with repetition as a fixed effect and random
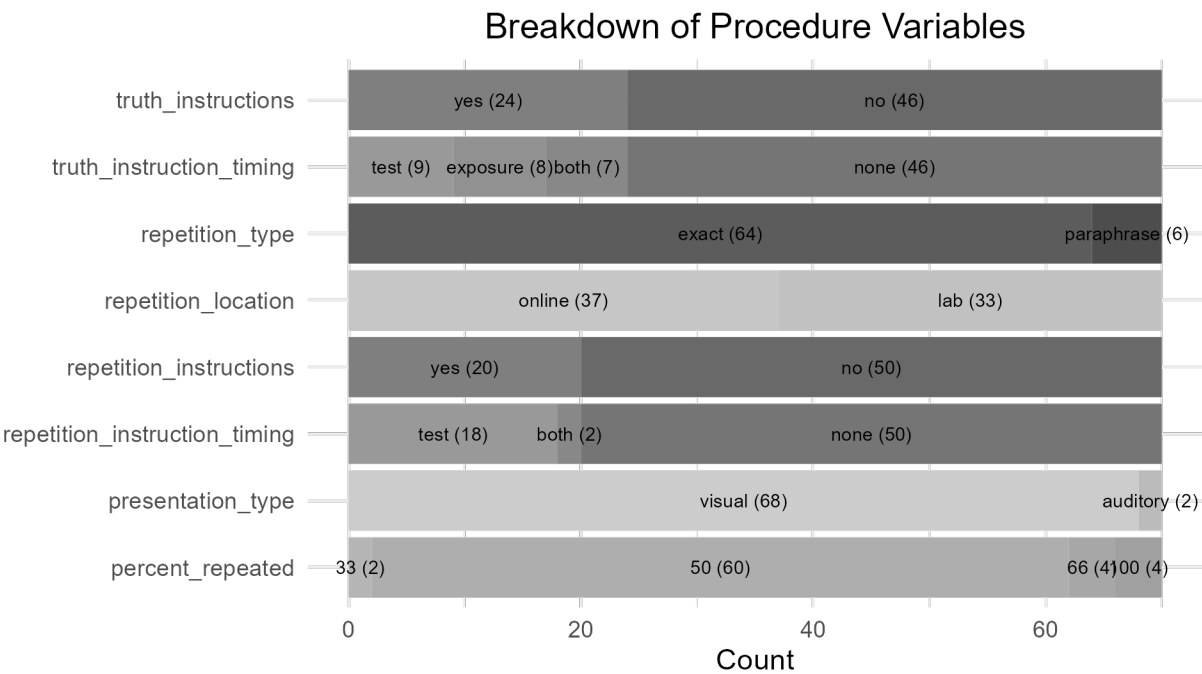
Figure 3

Overview of Procedure-related variables in TED

intercepts and slopes at the subject, statement, and procedure levels.

We analyzed the dichotomous and Likert-type response formats separately due to differences in their scale characteristics. Dichotomous responses (e.g., true/false) require logistic models, whereas Likert-type responses (e.g., 1–5 ratings) allow for linear models. All responses were maximum-normalized to the range 0-1 with one representing the maximum possible response indicating a "true" judgment.

We fit five models per dataset. A baseline model with only random intercepts. A full model including random intercepts and random slopes for repetition across all grouping factors. And three reduced models, each excluding the random slope of repetition for one grouping factor (subject, statement, or procedure).

Dichotomous Truth Judgments

The analysis was based on 112399 trials nested within 1576 subjects, 997 statements, and 14 procedures.

Table 1

	npar	AIC	BIC	logLik	deviance	Chisq
model_check_dichotomous	5.00	131,915.86	131,964.01	-65,952.93	131,905.86	NA
model_nosubject_re_dichotomous	9.00	131,743.99	131,830.66	-65,862.99	131,725.99	179.87
model_noprocedure_re_dichotomous	9.00	130,994.35	131,081.02	-65,488.17	130,976.35	749.64
model_nostatement_re_dichotomous	9.00	130,937.61	131,024.28	-65,459.80	130,919.61	56.74
model_full_re_dichotomous	11.00	130,934.97	131,040.90	-65,456.48	130,912.97	6.64

All models that included a random effect of repetition at any level outperformed the baseline model with only random intercepts. Detailed model comparisons, based on information criteria, are presented in Table 1.

The model incorporating random effects at all hierarchical levels provided the best overall fit **here, fit indices!**. Compared to the model excluding random effects of statements, it demonstrated a slightly lower AIC and a marginally higher BIC. This is attributable to the minimal variance observed in the repetition effect at the statement level. Notably, the variance of the random effect of repetition was highest at the subject level, followed by the procedure level.

Scale Truth Judgments

The analysis was based on 335462 trials nested within 6378 subjects, 1663 statements, and 47 procedures.

We employed the "bobyqa" optimizer, as some models initially failed to converge, and estimated the models using full maximum likelihood. An overview of the information criteria for all evaluated models is presented in Table 2.

The model including random effects at all hierarchical levels provided the best overall fit **here, fit indices!!**, exhibiting improved AIC and BIC values compared to the

Table 2

	npar	AIC	BIC	logLik	deviance	Chisq	Df	
model_check_scale	6.00	82,220.60	82,284.94	-41,104.30	82,208.60	NA	NA	
model_nosubject_re_scale	10.00	76,088.85	76,196.08	-38,034.42	76,068.85	6,139.76	4.00	
model_noprocedure_re_scale	10.00	69,974.52	70,081.76	-34,977.26	69,954.52	6,114.32	0.00	
model_nostatement_re_scale	10.00	69,205.50	69,312.73	-34,592.75	69,185.50	769.02	0.00	
model_full_re_scale	12.00	68,957.01	69,085.69	-34,466.51	68,933.01	252.49	2.00	

model that excluded random effects of statements. As in the dichotomous analysis, the variance of the repetition effect was near zero at the statement level, while being most pronounced at the subject level.

To further explore the influence of temporal delay between the exposure and judgment phases on interindividual variability in the repetition effect, we stratified the data. Specifically, we re-estimated the multilevel model (with random effects at all levels) separately for judgments made on the same day as the exposure phase and for those made on a subsequent day.

The results of this model can be found in table **table!**

Fit indices and effect sizes of these models. Variance measures here, too.

Model Sameday

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']
## Formula: response ~ repeated + (1 + repeated | subject) + (1 + repeated |
##     statement_id) + (1 + repeated | procedure_id)
##    Data: scale_data %>% filter(repetition_time <= 180)
##      AIC       BIC    logLik  deviance  df.resid
##  58089.80  58215.03 -29032.90  58065.80    251760
```

```
## Random effects:

##   Groups         Name         Std.Dev. Corr

##   subject        (Intercept) 0.08765

##                  repeated     0.11473  -0.20

##   statement_id (Intercept) 0.14378

##                  repeated     0.02567  -0.39

##   procedure_id (Intercept) 0.04136

##                  repeated     0.07676  -0.28

##   Residual                   0.26055

## Number of obs: 251772, groups:

## subject, 5783; statement_id, 1331; procedure_id, 33

## Fixed Effects:

## (Intercept)      repeated

##     0.50909        0.09176

## optimizer (nloptwrap) convergence code: 0 (OK) ; 0 optimizer warnings; 1 lme4 warning
```

Model Later

```
## Linear mixed model fit by maximum likelihood  ['lmerMod']

## Formula: response ~ repeated + (1 + repeated | subject) + (1 + repeated |

##     statement_id) + (1 + repeated | procedure_id)

##     Data: scale_data %>% filter(repetition_time > 180)

##       AIC       BIC     logLik  deviance  df.resid

## 10121.914 10233.932 -5048.957 10097.914      83678

## Random effects:

##   Groups         Name         Std.Dev. Corr

##   subject        (Intercept) 0.07323

##                  repeated     0.07048  0.04

##   statement_id (Intercept) 0.12034
```

```
##              repeated    0.02854  -0.21
##   procedure_id (Intercept) 0.03184
##              repeated    0.04206  -0.16
##   Residual                0.24810
## Number of obs: 83690, groups:
## subject, 1246; statement_id, 804; procedure_id, 14
## Fixed Effects:
## (Intercept)     repeated
##     0.50578       0.06057
```

The variance of the random effect of repetition was higher in the data containing truth judgments from the same day as the exposure phase (number 1) vs. (number 2).

Discussion

In this paper, we introduce the Truth-Effect Database (TED), a large-scale, open-access repository of trial-level data on truth judgments. TED is designed to support cumulative, reproducible research on the truth effect by standardizing and centralizing data from a wide range of studies. With contributions from over 10,000 participants, our database offers a comprehensive resource for researchers interested in understanding how repetition and related factors influence truth judgments.

Our goal in developing TED was to address key challenges in open research data: the fragmentation of data across studies, the lack of standard formats for sharing experimental data, and the resulting lack of interoperability and reusability **citations here**. To meet these challenges, we provide a structured SQL-based schema that organizes data at multiple levels—including publications, studies, participants, and individual trials. This structure allows for fine-grained analysis while maintaining consistency across diverse datasets.

We also provide tools to facilitate access and use of the database. In particular, the R package `acdcquery` **cite!!** allows users to extract data with flexible filtering criteria,

supporting both simple and complex queries. Additionally, we have developed a web-based data entry interface that automatically checks for consistency and flags incorrect variable names or entries (https://slesche.github.io/truth-effect-database/). This helps maintain high standards for incoming data and ensures that TED remains a reliable resource for the community, while simplifying the submission process and lowering the barrier to entry.

To illustrate the analytical potential of TED, we conducted multilevel models estimating random effects at the subject, statement, and publication levels to investigate where variance in the truth effect originates. This approach is only feasible with a dataset of TED's size and granularity.

In both the Likert-scale and dichotomous truth judgment data, we found that the random slopes of repetition on the statement level showed very little variance. This suggests that specific properties of individual statements (e.g., their semantic content or plausibility) have limited influence on the magnitude of the truth effect. However, statement-level variance in random intercepts was relatively large. This is consistent with expectations: some statements are objectively easier to judge as true or false, while others are more ambiguous, leading to baseline differences in overall truth ratings.

We also observed substantial variance in the publication-level random effect of repetition. This is unsurprising, as studies varied considerably in design features such as the number of statements, delay between exposure and judgment, or whether participants were warned about repetition. Nevertheless, the publication-level variance in the effect of repetition was smaller than at the subject level.

Indeed, the largest variance in the truth effect emerged at the subject level, indicating substantial individual differences in susceptibility to the repetition effect. This finding aligns with growing interest in understanding the psychological traits or cognitive mechanisms that moderate the truth effect at the individual level. Some studies **include studies here**.

However, efforts to predict individual differences in the truth effect have so far

produced mixed results **studies**.

Thus, it remains an open question whether the variability in the truth effect reflects reliable, trait-like differences. TED provides a promising platform for revisiting the question of reliable individual differences in the truth effect in future work.

Lastly, we explored whether the magnitude of variance in the repetition effect differed depending on the delay between exposure and judgment—a key procedural variable in truth effect research. Based on prior studies **include study**, we expected variability in the truth effect to be greater when the exposure and judgment occur on the same day, as repetition-based fluency is likely more salient in short-delay conditions.

also the overall size of truth effect here? In line with this expectation, we found that the variance in the random effect of repetition was indeed larger when judgments were made on the same day as exposure, compared to when judgments were delayed to a later day. **Interpretation**

However, these exploratory results should be interpreted cautiously. We did not test the delay-related differences for significance, and the analysis did not control for potential confounding factors such as sample composition, procedural variation, or other study-level covariates. **Something, in time with more data, this will become more representative. But it already points in one direction**

Taken together, these initial analyses demonstrate TED's utility for modeling complex interactions and variance structures that would be difficult to assess using single-study data. The open, extensible nature of TED ensures that these questions can be revisited and refined as the database continues to grow. We see this as a first step toward more nuanced investigations of the truth effect grounded in large-scale, reproducible data.

Future research

TED offers a unique foundation for advancing research through open, cumulative, and reproducible science. As the database grows in scope and depth, we anticipate several promising directions for future research and methodological development that TED is

particularly well-positioned to support.

One immediate application lies in enabling living meta-analyses of truth effect data that can evolve as new studies are contributed. Closely related, TED can support customized power analyses based on empirical variance from real experimental data. Researchers designing new studies can use TED's trial-level data to simulate expected effects based on previously collected data with similar study characteristics, improving the precision and efficiency of future study planning.

TED's reproducibility potential lies in its ability to support both the replication and reanalysis of existing studies using openly available, trial-level data. Researchers can directly replicate earlier truth effect experiments by matching procedural details drawn from the database, or they can reanalyze existing datasets to test new hypotheses, apply alternative statistical models, or verify published results. This facilitates both direct and conceptual replications and provides a foundation for transparent, data-driven reassessment of prior findings.

Beyond replication, TED also enables the development of formal cognitive models of the truth effect. With detailed trial-level data and metadata on procedural variables, researchers can test and compare models such as drift diffusion models, multinomial processing tree models, or signal detection accounts of truth judgments. The scale and granularity of TED make it possible to fit hierarchical models across subjects, statements, or studies, and to split data into independent training and testing sets—supporting robust model evaluation and cross-validation. This opens the door for developing and benchmarking cognitive theories of truth judgments with strong empirical support.

Importantly, these efforts are supported by TED's intentional focus on long-term sustainability. The database is built with minimal reliance on proprietary infrastructure or paid software. All tools and data are hosted on open platforms, primarily GitHub, and contributions are made through transparent, version-controlled workflows. This lightweight architecture was chosen to promote longevity and ensure that TED remains accessible and

modifiable regardless of funding cycles or institutional changes.

To ensure proper attribution and foster community investment, users of TED data are expected to cite all original publications corresponding to the datasets they use. In return, contributions to TED not only extend the utility of existing research but also amplify the visibility of individual studies, creating a mutually reinforcing incentive structure for both data sharing and reuse.

As TED continues to grow, we encourage researchers not only to contribute data but also to build on TED's infrastructure itself. TED itself is an extension of the Attentional Control Data Collection (ACDC) **cite!!** project, improving on data submission and ease-of-use and can as a blueprint and jumping-off point for future database development. All aspects of the project—including the SQL schema, submission interface, R extraction package, and the interactive overview website—are open source and fully forkable. Researchers are invited to reuse or adapt these components to develop similar infrastructures for other domains of psychological or behavioral science. In doing so, TED can help establish a broader culture of open, scalable, and sustainable data practices in experimental research.

Taken together, these features position TED as a platform for cumulative science: one that supports rigorous replication, sophisticated cognitive modeling, empirical synthesis, and infrastructure reusability while lowering the barriers to collaboration and long-term research development.

Conclusion

In this paper, we introduced the Truth-Effect Database (TED), a structured, large-scale resource for cumulative research on truth judgments. TED includes not only standardized, trial-level data but also open tools for user interaction, including a submission interface, a data exploration website, and an R package for flexible data extraction.

We illustrated the utility of TED through initial multilevel analyses, which highlighted substantial subject-level variance in the truth effect. This finding points to the

need for further theoretical and empirical work on individual differences in susceptibility to repetition.

Beyond this first demonstration, TED provides the foundation for a wide range of future research. These include living meta-analyses that evolve as new data are added, simulation-based power analyses informed by real-world variance, rigorous replication and reanalysis of existing studies, and the development of formal cognitive models. As an open and extensible infrastructure, TED also serves as a blueprint for sustainable, community-driven database development in psychological science.

References

Aust, F., & Barth, M. (2022). *papaja: Prepare reproducible APA journal articles with R Markdown.* https://github.com/crsh/papaja

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bauer, P. J. (2022). *Psychological science stepping up a level* (No. 2; Vol. 33, pp. 179–183). SAGE Publications Sage CA: Los Angeles, CA.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*(1), 395–411. https://doi.org/10.32614/RJ-2018-017

Bürkner, P.-C. (2021). Bayesian item response modeling in R with brms and Stan. *Journal of Statistical Software*, *100*(5), 1–54. https://doi.org/10.18637/jss.v100.i05

Crüwell, S., Apthorp, D., Baker, B. J., Colling, L., Elson, M., Geiger, S. J., Lobentanzer, S., Monéger, J., Patterson, A., Schwarzkopf, D. S., et al. (2023). What's in a badge? A computational reproducibility investigation of the open data badge policy in one issue of Psychological Science. *Psychological Science*, *34*(4), 512–522.

Foster, E. D., & Deardorff, A. (2017). Open Science Framework (OSF). *Journal of the Medical Library Association*, *105*(2). https://doi.org/10.5195/jmla.2017.88

Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., Flandin, G., Ghosh, S. S., Glatard, T., Halchenko, Y. O., et al. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, *3*(1), 1–9.

Hardwicke, T. E., Bohn, M., MacDonald, K., Hembacher, E., Nuijten, M. B., Peloquin, B. N., DeMayo, B. E., Long, B., Yoon, E. J., & Frank, M. C. (2021). Analytic reproducibility in articles receiving open data badges at the journal Psychological

Science: An observational study. *Royal Society Open Science*, *8*(1), 201494.

Henderson, E. L., Westwood, S. J., & Simons, D. J. (2022). A reproducible systematic map of research on the illusory truth effect. *Psychonomic Bulletin & Review*, *29*(3), 1065–1088. https://doi.org/10.3758/s13423-021-01995-w

Hipp, R. D. (2020). *SQLite* (Version 3.31.1).

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), e124.

Lesche, S. (2025). *Acdcquery: Query the attentional control data collection.* https://github.com/SLesche/acdc-query

Munafò, M. R., Nosek, B. A., Bishop, D. V., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 0021.

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., et al. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425.

Poldrack, R. A., Markiewicz, C. J., Appelhoff, S., Ashar, Y. K., Auer, T., Baillet, S., Bansal, S., Beltrachini, L., Benar, C. G., Bertazzoli, G., et al. (2024). The past, present, and future of the brain imaging data structure (BIDS). *Imaging Neuroscience*, *2*, 1–19.

R Core Team. (2022). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

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

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., et al. (2016). The

FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data, 3*(1), 1–9.

Xie, Y. (2015). *Dynamic documents with R and knitr* (2nd ed.). Chapman; Hall/CRC. https://yihui.org/knitr/

Appendix

Talking about appendices