Data Sharing is Caring – But Most Industrial and Organizational Psychologists Do Not Share Their Data

Ann-Kathrin Torka & Joachim Hüffmeier

TU Dortmund University

Psicostat Meeting, 21st of November 2025

"After research results are published, ***psychologists do not withhold the data*** *on which their conclusions are based* ***from other competent professionals who seek to verify the substantive claims through reanalysis*** *and who intend to use such data only for that purpose, provided that the confidentiality of the participants can be protected and unless legal rights concerning proprietary data preclude their release"* *(APA, 2017).*

Benefits of Data Sharing

- Increased transparency
- Opportunity to reproduce/verify findings (review)
- Decrease of errors/bias in published literature
- Reduction of QRPs (Tenopir et al., 2011; Bosma & Granger, 2022)
- Protection against QRP concerns
- Sustainability: conservation of resources, reuse of datasets, exchange between researchers

Journal Policies

- Journals are gatekeeper (Aguinis et al., 2020a; Kepes et al., 2018)
- TOP Factor: no journal in IOP requires data sharing

Table 1

Overview of the currently realized TOP standard "data transparency" for the journal categories "Business & Management" and "Psychology" (from TOP factor website [https://topfactor.org/])

TOP factor level	Business & Management		Psychology	
	N	%	*N*	%
Level 3: Data must be posted to a trusted repository and reported analyses will be reproduced independently prior to publication.	0	0 %	1	0.42 %
Level 2: Data must be posted to a trusted repository. Exceptions must be identified at article submission.	6	15 %	29	12.08 %
Level 1: Article states whether data are available, and, if so, where to access them.	8	20 %	82	34.17 %
Not Implemented: Journal encourages data sharing or does not provide any information about data transparency.	26	65 %	128	53.33 %
Overall	40	100 %	240	100 %

Note. Some journals are included in "Business & Management" *and* "Psychology."

Journal Policies

- Journals are gatekeeper (Aguinis et al., 2020a; Kepes et al., 2018)
- TOP Factor: no journal in IOP requires data sharing
- approx. 63% of journals in IOP mention *Data Transparency* on their websites (Torka et al., 2023)

Journal Policies

- Journals are gatekeeper (Aguinis et al., 2020a; Kepes et al., 2018)
- TOP Factor: no journal in IOP requires data sharing
- approx. 63% of journals in IOP mention *Data Transparency* on their websites (Torka et al., 2023)

Data Availability Statements (DAS)

- Statement on whether and how data is shared
- **Type 1 Data** (necessary für reproducability) vs. Type 2 Data (collected data that has not been analyzed/reported yet, Schönbrödt et al., 2017)
- Example:

„The Data are available from the authors upon reasonable request“.

Gabelica et al. (2022)

- Field: Biology, Medicine and Health
- Coding of Data Availability Statements
- *N* = 3.416 articles with DAS
- *N* = 1.792 request for data by e-mail
- Analysis of responses and shared data

Shared Data

- approx. 7 % shared their data
- approx. 93 % did not respond or did not share any data
- No difference between articles with and without DAS

Research Questions

1. How many recent articles in IO Psychology contain different levels of DAS?

2. How many authors communicate in their DAS that they will share their data sets (repository, supplement, upon request)?

3. How many data sets that are labeled as “available upon request” will actually be shared?

4. Are data sets more likely to be shared that are labeled as “available upon request” as compared to data sets without DAS?

5. How many shared data sets are prepared in ways that enable others to (re-) analyze them?

6. What are the reasons for not sharing requested data sets?

Method: Sample

Journals

- IOP-journals in the ABS Academic Journal Guide 2021
- Exclusion:
  - Language
  - Focus on theoretical contributions
  - Not primarily IO Psychology (e.g., military or school psychology; Wallrich et al., *In Principal Acceptance*)
- 27 journals

Articles

- Published between January – June 2023
- Exclusion: no data (e.g., editorials, theory papers)
- *N* = 487 articles with datasets

Method: Coding of DAS

Based on Gabelica et al. (2022) and Springer Nature (2023)

- Level 0: No DAS available
- Level 1: Data publicly available in a repository
- Level 2: Data available in a repository with restricted access
- Level 3: Data cannot be shared openly, but are available on request from authors
- Level 4: Data shared with manuscript or Supplementary Information
- Level 5: Data sharing is not applicable
- Level 6: Data cannot/will not be shared openly (e.g., due to data privacy laws or third-party restrictions)

Method: E-Mail Phase

- Articles with DAS “available upon request”
- Articles without DAS

- Standardized & personalized e-Mail

- E-Mails in April 2024
- Reminder after 2 months, June 2024

Method: Coding of Data Sets

- Level 0: The data set is not available in the stated repository
- Level 1: No data in the dataset contain clearly understandable labels
- Level 2: Some data in the dataset contain clearly understandable labels (i.e., it is for instance clear from the labels which variables are the predictor, criterion, and potential mediator and/or moderator variables or covariates)
- Level 3: All data in the dataset contain clearly understandable labels
- Level 4: The data set contains a codebook.

Results: DAS

RQ1:

How many articles contain the different levels of DAS?

- 286 articles (58.73 %)

RQ2:

How many authors communicate in their DAS that they will share their data (DAS Level 1-4)?

- 229 articles (80.07 %)

Table 2

Frequencies of Different DAS Levels

	Level	*n*	%
Level 0	No DAS available	201	41.27 %
Level 1	Data publicly available in a repository	43	8.83 %
Level 2	Data available in a repository with restricted access	3	0.62 %
Level 3	Data cannot be shared openly, but are available on request from authors	141	28.95 %
Level 4	Data shared with manuscript or Supplementary Information	24	4.93 %
Level 5	Data sharing is not applicable	21	4.31 %
Level 6	Data cannot/will not be shared openly (e.g., due to data privacy laws or third-party restrictions)	36	7.39 %
Levels 1 and 3*		1	0.21 %
Levels 1 and 6		3	0.62 %
Levels 3 and 6*		12	2.46 %
Levels 4 and 5		1	0.21 %
Levels 4 and 6		1	0.21 %
Total		487	100 %

Note. Some articles included more than one data set (e.g., multi study papers). Therefore, we assigned two different DAS Levels. The percentages do not add up to 100 % due to rounding. *As these articles include at least one study with a data set (i.e., the coding includes DAS Level 3), they were included in our data request.

Results

RQ1:

How many
different le

- 286 art

RQ2:

How many
DAS that t
Level 1-4)?

- 229 art

Table 2

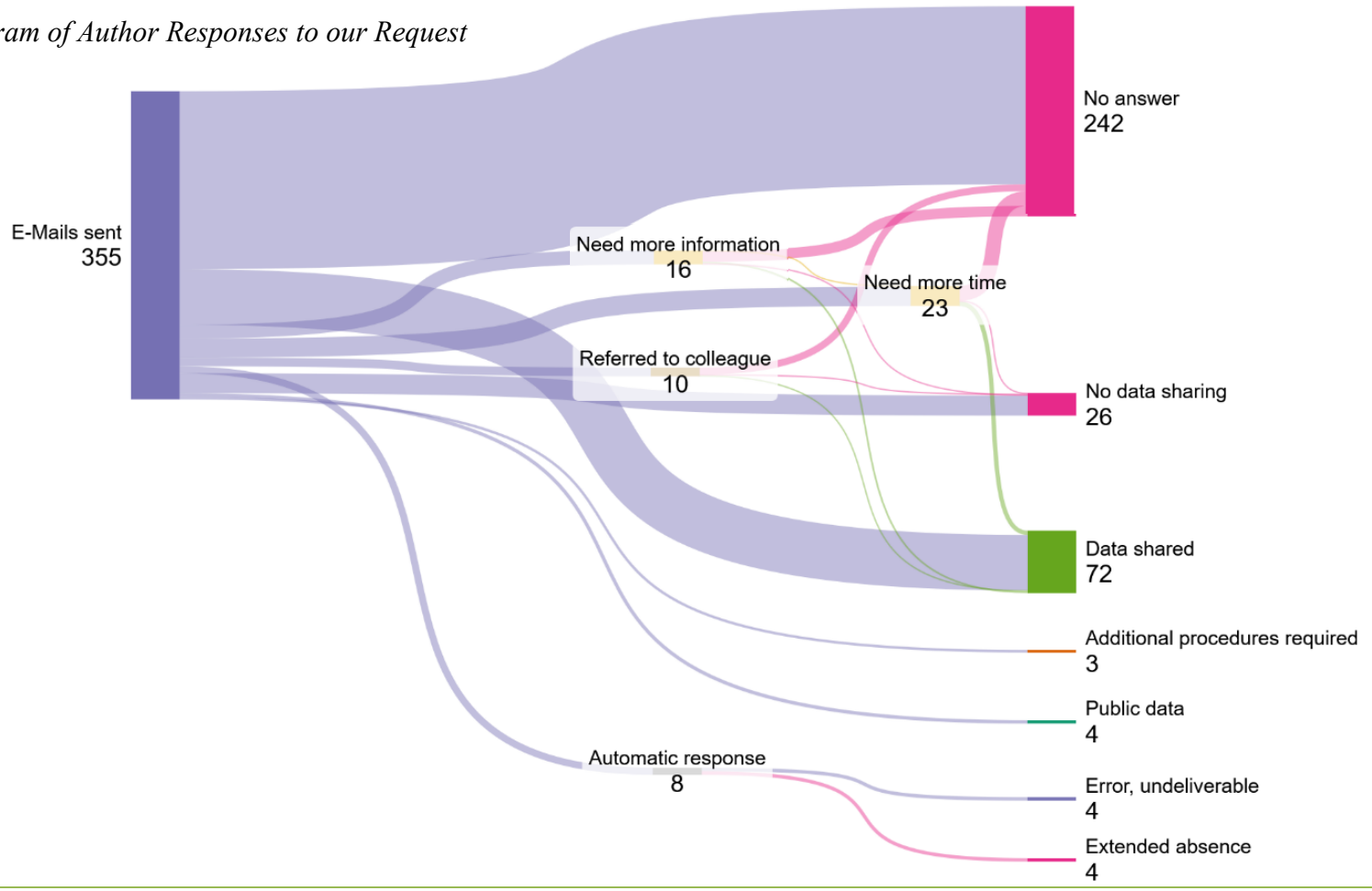
Frequencies of Different DAS Levels

	Level	*n*	%
Level 0	No DAS available	201	41.27 %
Level 1	Data publicly available in a repository	43	8.83 %
Level 2	Data available in a repository with restricted access	3	0.62 %
Level 3	Data cannot be shared openly, but are available on request from authors	141	28.95 %
Level 4	Data shared with manuscript or Supplementary Information	24	4.93 %
Level 5	Data sharing is not applicable	21	4.31 %
Level 6	Data cannot/will not be shared openly (e.g., due to data privacy laws or third-party restrictions)	36	7.39 %

Results: Data Sharing

Figure 2

Sankey Diagram of Author Responses to our Request



Results: Data Sharing

- 355 articles contacted
- 72 articles (20.28 %) shared their data

RQ3:

How many data sets that are labeled as "**available upon request**" will actually be shared?

- 154 articles with DAS Level 3
- 44 articles (28.57 %) shared their data

RQ 4:

Are data sets more likely to be shared that are labeled as "available upon request" as compared to data sets **without DAS**?

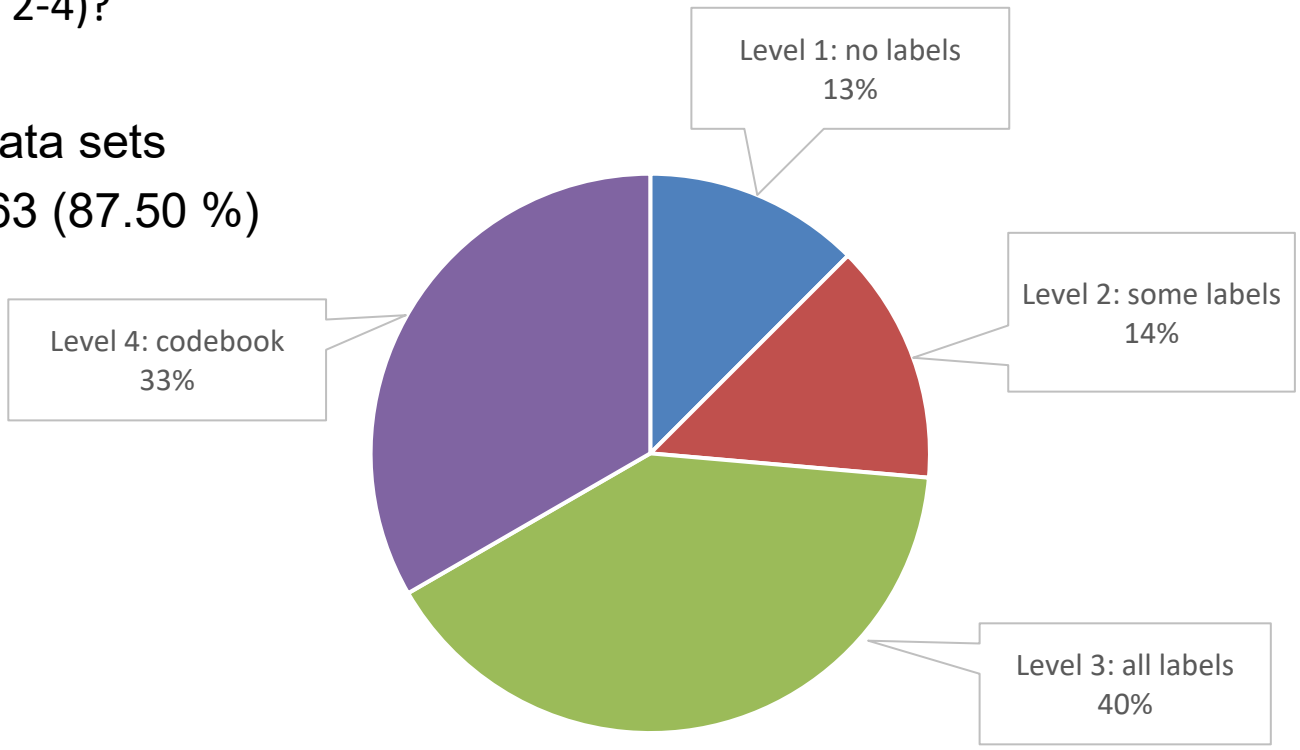
- 201 articles without DAS
- 28 articles (13.93 %) shared their data

$$\chi^2 (1) = 11.56, p < .001$$

Results: Data Sets

RQ 5: How many shared data sets are prepared in ways that enable others to (re-) analyze them (Levels 2-4)?

- *N* = 72 shared data sets
- Levels 2-4: *n* = 63 (87.50 %)



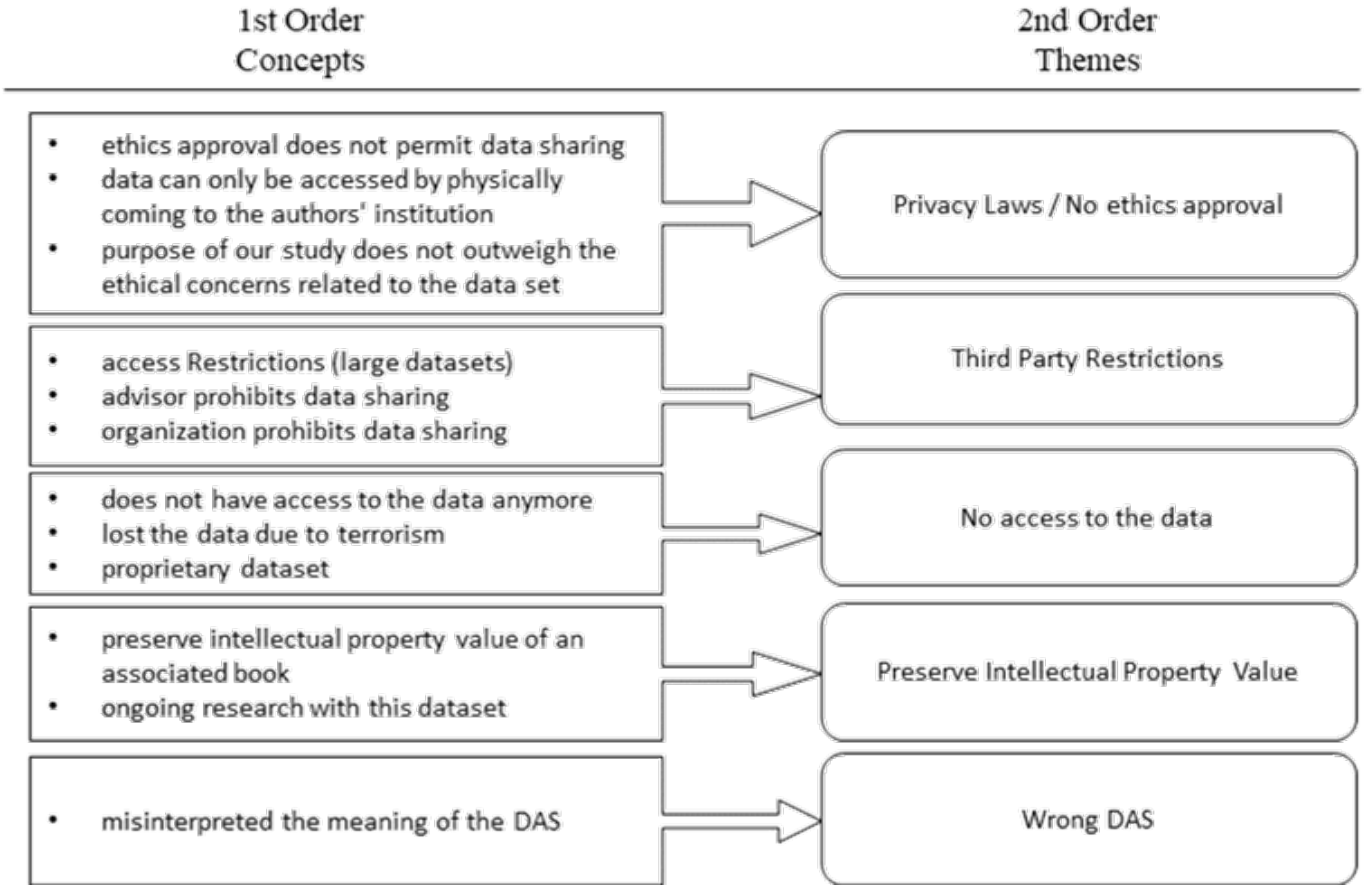
Results: Reasons against Data Sharing

RQ 6: What are the reasons for not sharing requested data sets?

- *N* = 27 authors explicitly did **not** want to share their data
- Gioia method (Gioia et al., 2012)

Results: Reasons against Data Sharing

RQ 6: What are the reasons for not sharing requested data sets?



Explorative Coding

Study Characteristics	Articles *without* shared Datasets *N* = 117	Articles *with* shared Datasets *N* = 60
Mean journal impact factor *M* (*SD*)	4.57 (2.47)	5.58 (2.71)
Mean size of author group *M* (*SD*)	3.82 (1.93)	3.98 (4.74)
Country Author Group	Homogeneous *n* = 72 International *n* = 44	Homogeneous *n* = 29 International *n* = 30

Explorative Coding

Study Type	Articles *without* shared Datasets *N* = 117	Articles *with* shared Datasets *N* = 60
survey	9	1
secondary data analysis	0	1
experiment: in the field	1	2
correlational field study: cross-sectional design	13	8
correlational field study: longitudinal design	26	12
experiment: online	7	9
experiment: in the lab/classroom	6	6
meta-analysis	4	11
review	11	3
qualitative research	12	1
analyses of archival data	7	2
observational study	0	1
test/scale development	12	3
method development beyond test/scale development	2	0

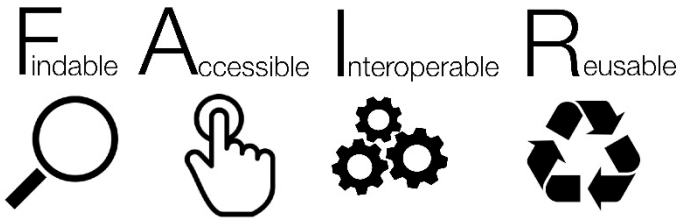
Discussion

- About half of all articles have a DAS
- DAS „upon request“ most common (approx. 50%)
- Intention (DAS) increases actual data sharing
- BUT: even with DAS „upon request“, approx. 70% do not share data
- Shared data largely enable (re)analyses

Implications

- Journals: DAS easy method
  - DAS probably not sufficient
  - Incentives for data sharing (e.g., Badges)

- Authors: Consider Data Sharing
  - Consent form
  - Anonymization of sensitive data
  - Flexible instead of strict data sharing
  - FAIR data sharing



Thank you for your attention! ☺

Questions?

ann-kathrin.torka@tu-dortmund.de