Meta-Psychology, 2025, vol 9, MP.2021.2908
https://doi.org/10.15626/MP.2021.2908
Article type: Original Article

Open data: Yes
Open materials: Yes
Open and reproducible analysis: Yes
Open reviews and editorial process: Yes
Preregistration: Yes

Edited by: Rickard Carlsson
Reviewed by: Gustav Nilssone, Chelsea Parlett-Pelleriti
Analysis reproduced by: Jens Fust
All supplementary files can be accessed at OSF:
https://doi.org/10.17605/OSF.IO/T6D2Q

The Evolution of Data Sharing Practices in the Psychological Literature

Judith A. Neve[1] and Guillaume A. Rousselet[1]
[1]University of Glasgow

Sharing data has many benefits. However, data sharing rates remain low, for the most part well below 50%. A variety of interventions encouraging data sharing have been proposed. We focus here on editorial policies. Kidwell et al. (2016) assessed the impact of the introduction of badges in *Psychological Science*; Hardwicke, Mathur, et al. (2018) assessed the impact of *Cognition*'s mandatory data sharing policy. Both studies found policies to improve data sharing practices, but only assessed the impact of the policy for up to 25 months after its implementation. We examined the effect of these policies over a longer term by reusing their data and collecting a follow-up sample including articles published up until December 31st, 2019. We fit generalized additive models as these allow for a flexible assessment of the effect of time, in particular to identify non-linear changes in the trend. These models were compared to generalized linear models to examine whether the non-linearity is needed. Descriptive results and the outputs from generalized additive and linear models were coherent with previous findings: following the policies in *Cognition* and *Psychological Science*, data sharing statement rates increased immediately and continued to increase beyond the timeframes examined previously, until reaching close to 100%. In *Clinical Psychological Science*, data sharing statement rates started to increase only two years following the implementation of badges. Reusability rates jumped from close to 0% to around 50% but did not show changes within the pre-policy nor the post-policy timeframes. Journals that did not implement a policy showed no change in data sharing rates or reusability over time. There was variability across journals in the levels of increase, so we suggest future research should examine a larger number of policies to draw conclusions about their efficacy. We also encourage future research to investigate the barriers to data sharing specific to psychology subfields to identify the best interventions to tackle them.

Keywords: journal policies, metascience, open science, reproducibility, scientific publishing

Introduction

Science aims to be transparent and reproducible. However, the scientific cycle suffers from multiple threats, including publication bias (Franco et al., 2014; Ioannidis, 2014; Sterling, 1959) and questionable research practices (see Bakker et al., 2012; John et al., 2012 for a review). This results in a literature riddled with false positives (Forstmeier et al., 2017). A variety of possible solutions have been highlighted, including encouraging or imposing data sharing (Forstmeier et al., 2017; Munafò et al., 2017).

Benefits of Data Sharing

Sharing data has many benefits for both science and individuals. The most obvious benefit is to facilitate the evaluation of published results via analytical reproduction, particularly if analysis code is shared alongside the data (Hardwicke, Mathur, et al., 2018; Obels et al., 2020). Other benefits include to allow meta-analyses to reuse the data directly rather than to rely on summary statistics (Asendorpf et al., 2013; Ioannidis, 2014; Martone et al., 2018), to let other researchers test alternative analyses (e.g. Deschaght et al., 2017), and to allow future studies to have larger and more varied samples (e.g. Nielson et al., 2014; Perrino et al., 2015; see Martone et al., 2018 for a review). Benefits to researchers' careers have also been noted, such as increased citation frequency compared to similar papers that did not share their data (McKiernan et al., 2016; Piwowar and Vision, 2013; see Martone et al., 2018 for a review).

Data Sharing Rates

Despite these benefits, data sharing rates remain low, for the most part below 50% (Table 1). Although these rates increase after contacting authors, they re-

Table 1

Some studies about data retrieval rates.

Reference	Retrieval rates	N	Field or journal
Stodden et al., 2018	26 (12.75%) before contacting authors, 89 (43.63%) after contacting authors	204	*Science* – 2011-2012
Stockemer et al., 2018	13 (8.97%) before contacting authors, 82 (56.55%) after contacting authors	145	Political science journals
Vanpaemel et al., 2015	148 (38%)	394	APA journals

main far from ideal, particularly when considering that *Science* or APA journals state that data must be shared upon request ("Ethical Principles of Psychologists and Code of Conduct," 2017; Hanson et al., 2011; "Science Journals," 2018). Moreover, the proportion of working email addresses decreases with time (Vines et al., 2014), thus showing that needing to contact authors is not a sustainable way of obtaining or sharing data.

Possible Causes of Low Rates

There may be many reasons underpinning these low rates. Martone et al. (2018) provided a list of possible objections to data sharing. Although legal reasons make it impossible to share data containing information that may lead to participant identification, anonymisation has become an ethical standard. There must therefore be additional reasons for the lack of data sharing. These include fear that the data will reveal errors in the original analysis, fear of others using published data to conduct research planned by the collectors, or simply the effort required (Baldwin and Del Re, 2016; Longo and Drazen, 2016; Martone et al., 2018).
However, these concerns only have limited validity: when errors are revealed, this does not damage researchers' reputation (Ebersole et al., 2016). The fear of "data parasites" can be mitigated by only publishing data after completing all relevant analyses. As for the effort associated with sharing data, this can be limited by preparing for data sharing during data collection, analysis, and storing. It therefore appears that most concerns surrounding data sharing can be addressed (Martone et al., 2018).

Possible Interventions for Data Sharing

As addressing these concerns has been insufficient, interventions to improve data sharing have been proposed. These include funding agencies requiring information as to how the data will be shared (e.g. "Final NIH Policy for Data Management and Sharing," 2020), grassroots initiatives such as the Peer Reviewers' Openness Initiative (Morey et al., 2016), or editorial policies (e.g. Eich, 2014; Lindsay, 2017; Sloman, 2015). An increasing number of journals have been implementing policies regarding data sharing. For instance, data sharing has been encouraged in journals such as *Psychological Science*, where badges were implemented to reward open practices in 2014 (Eich, 2014). More drastically, some journals such as *Cognition* have made data sharing a condition of publication (Sloman, 2015).
Research on these policies' impact has operationalised data sharing as the presence of a data sharing statement, as this is a useful and simple to implement proxy, although the presence of a statement in no way guarantees effective data sharing (Hardwicke, Mathur, et al., 2018; Kidwell et al., 2016). For this reason, the reusability of data is a better assessment of effective data sharing, albeit more time consuming and subjective. Reusability can be defined as being constituted of three elements (Hardwicke, Mathur, et al., 2018): actual availability (whether the data can be found, downloaded, and opened), completeness (whether all the data necessary to replicate the results is available), and understandability (whether the data can be made sense of). Taken together, data sharing statements and reusability can provide a useful idea of the state of data sharing.

Previous Findings on the Effect of Journal Policies

Two key studies have examined how the policies implemented in *Psychological Science* (Kidwell et al., 2016) and *Cognition* (Hardwicke, Mathur, et al., 2018) have impacted data sharing statements and reusability. Tables 2-4 summarise their results.

Table 2

Studies about data sharing policies' impact on data sharing statement rates.

Reference	Data sharing statement rates before policy	Trend before policy	Data sharing statement rates after policy	Trend after policy	Journal
Kidwell et al., 2016	2.5%	-	22.8%, peak: 39.4% (first half of 2015)	Monotonic increase every half year	*Psychological Science*
	1% - 4.5%	-	-	-	Comparison journals
Hardwicke, Mathur, et al., 2018	25%	1.04-fold increase every 50 days	78%	1.14-fold increase every 50 days	*Cognition*

Table 3

Kidwell et al. (2016)'s findings on data reusability rates among articles with a data availability statement.

Policy	Rate of each reusability criterion		
	Available	Usable	Complete
No badges	40.5%	21.6%	16.2%
Badges	93.8%	70.3%	70.3%

Table 4

Hardwicke, Mathur, et al. (2018)'s findings on data reusability rates among articles with a data availability statement.

Policy	Reusability rate
No policy	22%
Mandatory data sharing	62%

The Impact of Badges in Psychological Science

Kidwell et al. (2016) considered the entire population of articles published between 1st January 2012 and 31st May 2015 in *Psychological Science* and four comparison journals without policies. Comparing pre- and post-policy data sharing statement rates for each journal, they found that data sharing statement rates had gone up considerably in *Psychological Science* only. They also explored the trend by computing the rates at 6-month intervals, and found that following the policy, there was a positive trend in data sharing statement rates in *Psychological Science*.
Regarding data reusability, they found that articles that had obtained a badge for their shared data had much higher rates of reusability compared to those that had shared their data but did not have a badge (possible causes for not having a badge being either not having had the opportunity, i.e. pre-policy *Psychological Science* or comparison journals, or not having received a badge after the implementation of badges).

The Impact of Cognition's Mandatory Data Sharing Policy

Hardwicke, Mathur, et al. (2018) considered the entire population of articles published in *Cognition* between 1st March 2014 and 31st March 2017. They grouped articles into 50-day bins to compute data sharing proportions and fit a generalised linear model with time and policy as predictors. They found a baseline positive trend, which increased significantly after the implementation of the policy; however, despite the policy stating mandatory data sharing, statements did not reach 100%. The introduction of the policy was also associated with an improvement in data reusability, from 22% to 62%.

Taking These Results Further

These two studies show that policies lead to an increase not only in overall data sharing statement rates, but also in a positive trend over time. Although these results are encouraging, it is important to note that the trend is based on data only for 17 months following the introduction of badges in *Psychological Science* and for 25 months following the introduction of a policy in *Cognition*. It is therefore of interest how these positive trends evolve with time, that is, whether they continue, plateau, or decrease again, and if so, at what point they do so.
Similarly, we can see that policies substantially improve effective data sharing rates. These rates are however far from the ideal 100%. This can be attributed to previous studies computing an overall pre-policy and an overall post-policy rate, thus grouping articles published immediately after the implementation of the policy, which may not have adapted to the policy, with articles published much later, thereby lowering the overall rate. Examining time trends would shed light on data sharing practices and help identify areas that may need more work and interventions in order to promote full transparency.

The Current Study

The current study preregistered three primary goals that aimed to complement previous research. We reused data collected for Hardwicke, Mathur, et al. (2018)'s and Kidwell et al. (2016)'s studies and collected data from articles published in all their examined journals since their cut-off dates and until December 31st, 2019. The first two goals were to expand Hardwicke, Mathur, et al. (2018)'s and Kidwell et al. (2016)'s results regarding the long-term trend in data availability statement rates and data reusability. We assessed whether the previously observed trends in data availability statements continue, or whether they plateau or start decreasing, and examined trends in data reusability. This was done using non-parametric regression in order to capture information in as much depth as possible. The necessity of the non-parametric trend was evaluated by comparing that model to a parametric one.
Thirdly and finally, we assessed whether previous findings regarding policies are specific to journals, or whether they are generalisable. To this aim, we examined the impact of the implementation of badges in 2016 in *Clinical Psychological Science* (Lilienfeld, 2017).

For maximum comparability, this was done by comparing pre- and post-policy point estimates and parametric trends as done in previous research.

Methods

The study's preregistration can be accessed at https://osf.io/6ynur.

Ethics

No ethical issues were identified for this study as it did not involve data from human or animal participants. The study was approved by the Glasgow University School of Psychology ethics committee.

Sample

The sample included the population of empirical articles collected for Hardwicke, Mathur, et al. (2018)'s and Kidwell et al. (2016)'s studies, as well as a follow-up sample of articles published between the cut-off date of their research and December 31st, 2019. The only exclusion criterion was if the article was not empirical.
The journals included were *Psychological Science* (PS), *Clinical Psychological Science* (CPS), *Journal of Personality and Social Psychology* (JPSP), *Journal of Experimental Psychology: Learning, Memory, and Cognition* (JEPLMC), and *Developmental Psychology* (DP), as examined by Kidwell et al. (2016). We used the population of articles published in these journals between 1st January 2012 and 31st May 2015 (Kidwell et al. (2015b)), and an additional sample of the articles published in the same journals between 1st June 2015 and 31st December 2019. *Cognition* was also included in this study, as examined by Hardwicke, Mathur, et al. (2018). We used the population of *Cognition* articles published between 1st March 2014 and 31st March 2017 (Hardwicke, Mathur, et al. (2018b)), as well as a sample of articles published between 1st April 2017 and 31st December 2019.
The population of articles to sample from was established using a database search in Scopus including the journals and their relevant dates, and stratified sampling was then applied to ensure equal sampling from all journals. This was done in R (R Core Team, 2020), using the `sample_n()` function of the tidyverse library (Wickham et al., 2019).
510 articles were selected at random from this population, with 85 articles per journal. Due to Scopus only allowing to filter by year rather than year and month, only 464 of these were included in the current study as the remaining 46 studies were published outside of our sampling timeframe. Combining the current dataset with Hardwicke, Mathur, et al. (2018b)'s and Kidwell et al. (2015b)'s datasets, the sample contained n = 3753 articles. 236 non-empirical articles were excluded. 1 article was excluded due to its publication journal not being included in this study. The remaining sample contained n = 3516 articles. Table 5 provides a breakdown of number of articles per journal, study of origin, and when applicable, pre- or post-policy.

Procedure

Initial Sampling

A Boolean string (https://osf.io/tswk2/) was used to search Scopus, restricting the journals, document types, and publication years to the population of interest. The file containing the references of all articles returned by Scopus is available at https://osf.io/myxbg/. The `sample_n()` function of the tidyverse library (Wickham et al., 2019) in R was then used to sample an equal number of articles from each journal. The R script is available at https://osf.io/cuzbd/.

Downloading Articles

Articles could not be downloaded automatically because the selected journals were not open access. After initial sampling, the sample was separated in articles published in APA versus non-APA journals in order to search them in an appropriate database. R was used to write a search code to input into the appropriate database for 50 articles at a time for non-APA journals, and for all selected articles for APA journals. The code is available at https://osf.io/cuzbd/.
This search code was a Boolean string, where the articles' DOIs were separated by an "or" statement. The search code for articles published in APA journals was input in EBSCOhost and articles were downloaded manually from there. The search code for articles published in non-APA journals was input in Scopus and the "download all" function was used to download the open access articles. The remaining articles were flagged by the site and downloaded manually. This procedure was devised during pilot studies (https://osf.io/u3wcx/).

Checking Downloaded Articles

The downloaded PDF files were put in a designated folder in the referencing program Zotero ("Zotero" (2020)), which automatically attempted to retrieve bibliographical information. The collection was then exported in CSV format. Articles sampled from the Scopus database were matched with the downloaded articles based on DOI, using the `full_join()` function in the R tidyverse library (Wickham et al., 2019). We then

Table 5

Sample breakdown.

Sample	Journal						
	PS	*Cognition*	CPS	JEPLMC	JPSP	DP	TOTAL
Current study	75	77	72	75	71	78	448
Previous study	837	591	104	483	419	634	3068
Pre-policy	517	167	118	-	-	-	-
Post-policy	395	501	58	-	-	-	-
TOTAL	912	668	176	558	490	712	3516

examined whether all sampled articles had been downloaded, whether any additional articles had been downloaded, and whether the article titles matched between the two files. When issues appeared, missing articles were downloaded and additional articles were deleted. This procedure was devised during pilot studies. In addition to this, where Zotero did not retrieve sufficient bibliographical information, this was input manually and checked against the target sample. The R script is available at https://osf.io/hm7y6/.

Exclusion of Articles in Overlapping Timeframe

The resulting sample of articles contained publication month information as well as publication year. Articles published at a date included in Kidwell et al.'s (2016) or Hardwicke et al.'s Hardwicke, Mathur, et al. (2018)'s and Kidwell et al. (2016)'s studies were excluded and a CSV file was written containing the bibliographical information of the remaining articles. This file is available at https://osf.io/qhvuz/.

Data Sharing Assessment

All downloaded articles were evaluated using the questionnaire available at https://osf.io/z3gdw/. This questionnaire was formatted as a Google Form and the data was then downloaded as a CSV file. This data file is available at https://osf.io/s8wnf/.

Measures

Bibliographical information was extracted from a combination of the Scopus database and automatic processing of the downloaded articles in Zotero. When not available, the publication date was inserted manually using the issue and volume numbers.
The assessment of open data practices was done using a questionnaire developed by combining questions about data sharing common between Hardwicke, Mathur, et al. (2018a)'s questionnaire and Kidwell et al. (2015a)'s questionnaire. This current questionnaire is available at https://osf.io/z3gdw/.
The questionnaire was checked using pilot studies (supplementary information at https://osf.io/u3wcx/) that aimed to assess coding consistency between the previous studies and the current one, as well as to identify any difficulties in usage. All changes were recorded and the questionnaire was updated regularly on the Open Science Framework for transparency.
The first pilot study resulted in minor changes in wording of the questions and possible answers. Coding congruency between the previous studies and the current one was high (28/30), with discrepancies coming mainly from the assessment of understandability. None of the articles randomly selected from Kidwell et al. (2015b) dataset shared their data. For these two reasons, a second pilot was conducted focusing on articles that had shared their data, and articles that had shared their data but for which understandability was coded as not understandable in the past.
The second pilot study had low congruency between previous studies and the current study (2/20). This was however to be expected as articles were intentionally chosen to be those with problematic data sharing. Moreover, five discrepancies were due to a difference in treatment of articles with a data sharing statement but for which the available file was not data: Hardwicke, Mathur, et al. (2018) coded data as not understandable, whereas the current coder left understandability assessments empty. The questionnaire was therefore adapted to explicitly state the procedure for these cases and past assessments were recoded in R to fit the current procedure. Questions were removed to group similar issues (e.g. broken link and data not at the given link) into a single question to simplify the coding procedure. The use of a pilot sample with numerous issues in data sharing allowed for a thorough check of the questionnaire, thus resulting in the final version.

Variables

Outcomes

The two outcome variables, availability statement and reusability, were treated in the same way during the analysis.

Availability Statement. Availability statement was a binary coded variable, where "0" corresponded to the case where the article did not have a data availability statement, and "1" corresponded to the case where it did.

Reusability. Reusability was a binary coded variable, where "1" was the case where data was available, complete and understandable, and "0" was the case where one or more of these conditions for in-principle reusability was not fulfilled.

Predictors

Predictor variables were time, journal, and policy.

Time. Time was primarily defined by the full publication date of an article due to Kidwell et al. (2015b) only providing this information in their dataset. Unregistered exploratory analyses considered reception date of the article by the journal, which was collected by the present researchers and Hardwicke, Mathur, et al. (2018b).

Journal. Journal was a dummy coded set of five variables corresponding to journals, with *Journal of Experimental Psychology: Learning, Memory, and Cognition* as the baseline. This choice was justified by *Journal of Experimental Psychology: Learning, Memory, and Cognition* having the lowest data sharing rates after *Clinical Psychological Science* (Kidwell et al., 2016). As *Clinical Psychological Science* has introduced a data sharing policy, deviations from a baseline are expected. Dummy coding was chosen as it highlights differences from a selected level of the variable, whereas other coding schemes such as effect or deviation coding examine deviations from an overall mean (Alkharusi, 2012). As we did not expect the overall mean to be representative of all journals, due to the expected variations among journals, it was most appropriate to use this as the baseline.

Policy. Policy was a binary variable coded as "0" if there was no policy in the publication journal at the time of publication of the article, and "1" if there was a policy of any kind.

Analysis Plan

We chose to use Generalised Additive Models (GAMs) and Generalised Linear Models (GLMs). Generalised models were needed as our outcome variables were binary, and the modelling of their probabilities was therefore bound between 0 and 1. The logit link function was used as it is the link function corresponding to binomial data.

GAMs are semi-parametric models. This allows for an exploration of non-linear changes in the trend, such as decreases, plateaus, or accelerated increases (Jones and Almond, 1992; Simpson, 2018). GAMs thus show changes over time in more detail than average increases or changes due to the implementation of policies. However, estimates of non-linear changes are non-parametric and thus cannot be computed, so the additional information regarding the trend needs to be assessed informally using a plot of the trend.

GLMs are parametric models and therefore allow for quantifiable estimations of changes in trends. The measured changes would however only be linear or of a low-order polynomial (if model specified as such; e.g. Zoltowski and Pillow, 2018). This allows for much less flexibility than GAMs in modelling the trend. However, this flexibility might not be needed if the trend continued as observed previously or if changes are minor.

The output of the GAM may suggest higher order polynomials might be more appropriate for the effect of time. When the effective degrees of freedom (i.e. the estimation of the degrees of freedom that would be needed to accurately model the trend in a parametric approach; Hastie and Tibshirani, 1987) for time or an interaction with time were > 1.5 and ≤ 3.5, an additional GLM was fit. This additional GLM included a polynomial term for the effect of time. The order of this polynomial was the rounded effective degrees of freedom. 1.5 was chosen as the lower bound as < 1.5 would be rounded to 1 and this is a linear term which is fit in the GLM described above. 3.5 was chosen as the upper bound as this allows for up to a cubic polynomial. This allows for more flexible estimation of the effect of time while remaining quantifiable, unlike GAMs, and relatively robust, unlike higher degree polynomials (Wang et al., 2017).

The three models can be compared using the Bayesian Information Criterion (BIC; Schwarz, 1978). This criterion compares models' trade-off of goodness-of-fit and complexity without needing them to be nested. The model with the lowest BIC is defined to be the best model for the data. We opted for comparisons using information criteria as the models were not all nested within another, making it inappropriate to compare them using an F-test. We chose to use the BIC as it tends to favour simpler models for explanation, compared to the Akaike Information Criterion (AIC; Akaike, 1974) which is better for models used for prediction (Chakrabarti and Ghosh, 2011).

Analysis was conducted in R, using the tidyverse (Wickham et al., 2019), lubridate (Grolemond and Wickman, 2011), and mgcv (Wood, 2017) libraries. The registered analysis scripts are available at https://osf.io/6ynur/files/. The final analysis scripts are available at https://osf.io/amjdp/. We detail the models used for each research question below.

Research Question 1: How do data sharing statement rates evolve in the long term?

We registered the GAM as:

$$
\begin{aligned}
\text{logit(availability statement)} = {} & \alpha_0 \\
& + \sum_{i=1}^{6} \alpha_i \text{journal} \\
& + \alpha_7 \text{policy} \\
& + \sum_{i=1}^{3} \gamma_i \text{journal} * \text{policy} \\
& + f_1(\text{publication date}) \\
& + \sum_{i=1}^{6} f_{2i}(\text{publication date} \,|\, \text{journal}) \\
& + f_3(\text{publication data} \,|\, \text{policy}),
\end{aligned} \tag{1}
$$

with identifiability constraints $\alpha_1 = f_{21} = 0$; the journals in the journal x policy interaction were only those with post-policy data. f represent the functions of smoothing splines modelling the predictors' effects.

In practice, the model was rank deficient. Rank deficiency may be due to a variety of reasons, such as the scale of the predictors, a sample size that is too small, or collinearity, and causes models that are unstable (O'Brien, 2012). We deviated from the registered procedure to fit a full rank model. We first attempted to resolve rank deficiency by removing policy due to its collinearity with publication date. This model was still rank deficient, so we instead removed the parametric terms due to their collinearity with the smooth interaction terms, yielding the following model:

$$
\begin{aligned}
\text{logit(availability statement)} = {} & f_1(\text{publication date}) \\
& + \sum_{i=1}^{6} f_{2i}(\text{publication date} \,|\, \text{journal}) \\
& + f_3(\text{publication data} \,|\, \text{policy}),
\end{aligned} \tag{2}
$$

with identifiability constraint $f_{21} = 0$. f represent the functions of smoothing splines modelling the predictors' effects.

This final model was of full rank.

We registered the GLM as:

$$
\begin{aligned}
\text{logit(availability statement)} = {} & \alpha_0 \\
& + \sum_{i=1}^{6} \alpha_i \text{journal} \\
& + \alpha_7 \text{policy} \\
& + \sum_{i=1}^{3} \gamma_i \text{journal} * \text{policy} \\
& + \beta_1 \text{publication date} \\
& + \sum_{i=1}^{6} \eta_i \text{journal} * \text{publication date} \\
& + \eta_7 \text{publication data} * \text{policy} \\
& + \eta_{i=8}^{13} \eta_i * \text{journal} * \text{policy} * \text{publication date},
\end{aligned} \tag{3}
$$

with identifiability constraints $\alpha_1 = \gamma_1 = \eta_1 = \gamma_8 = 0$; the journals in the journal x policy interaction were only those with post-policy data.

This model was rank deficient. We obtained a model of full rank by removing the policy variable due to collinearity:

$$
\begin{aligned}
\text{logit(availability statement)} = {} & \alpha_0 \\
& + \sum_{i=1}^{6} \alpha_i \text{journal} \\
& + \beta_1 \text{publication date} \\
& + \sum_{i=1}^{6} \eta_i \text{journal} * \text{publication date},
\end{aligned} \tag{4}
$$

with identifiability constraints $\alpha_1 = \eta_1 = 0$.

The polynomial GLM contained additional terms for quadratic and cubic effects of publication date and interactions with these.

Research Question 2: How do effective data sharing rates change with time?

Data reusability was assessed on a subsample, excluding all articles that did not have a statement indicating fully shared data. This differs from the preregistration in that we excluded articles with a statement of partial data availability as these would not fulfil the completeness condition of reusability. The models were registered to be defined as above, with reusability as the outcome instead of availability statement. In practice, due to rank deficiency, we removed parametric predictors and policy from the GAM and policy from the GLMs, leading to the following models for the GAM and the linear GLM, respectively:

$$\text{logit(reusable)} = f_1(\text{publication date}) + \sum_{i=1}^{6} f_{2i}(\text{publication date} \,|\, \text{journal}) \tag{5}$$

with identifiability constraint $f_{21} = 0$. f represent the functions of smoothing splines modelling the predictors' effects.

$$\begin{aligned} \text{logit(reusable)} &= \alpha_0 \\ &+ \sum_{i=1}^{6} \alpha_i \text{journal} \\ &+ \beta_1 \text{publication date} \\ &+ \sum_{i=1}^{6} \eta_i \text{journal} * \text{publication date}, \end{aligned} \tag{6}$$

with identifiability constraints $\alpha_1 = \eta_1 = 0$.
The polynomial GLM included quadratic and cubic terms for publication date.

Research Question 3: What is the impact of the introduction of badges in *Clinical Psychological Science*?

To compare the results from *Clinical Psychological Science* with Kidwell et al. (2016)'s results, pre- and post-policy data sharing statement rates were computed. These rates were compared to the rates in the other comparison journals included in Kidwell et al. (2016)'s study.

The trend was examined by fitting a GLM analogous to that used by Hardwicke, Mathur, et al. (2018):

$$\begin{aligned} \text{logit(availability statement)} &= \beta_0 \\ &+ \beta_1 \text{policy} \\ &+ \beta_2 \text{publication date} \\ &+ \beta_3 \text{publication date} * \text{policy}. \end{aligned} \tag{7}$$

This allowed to compare the slope before the policy to the slope after the policy by comparing β_2 with $\beta_2 + \beta_3$, and assess the significance of this difference via the significance of β_3.

A GAM was also registered to examine the time trend in more detail:

$$\begin{aligned} \text{logit(availability statement)} &= \beta_0 \\ &+ \beta_1 \text{policy} \\ &+ f_1(\text{publication date}) \\ &+ f_2(\text{publication date} \,|\, \text{policy}). \end{aligned} \tag{8}$$

A quadratic polynomial GLM was fit to quantify the trend with more flexibility than the GLM.

Unregistered Exploratory Research

Using Reception Date rather than Publication Date. There can be important delays between the submission of an article and its publication date (Björk and Solomon, 2013). The difference between these dates is particularly relevant here, as we aim to examine the effect of policies affecting article submission guidelines and rules rather than only of time. We fit the models used in Research Question 1 again on a subset of articles that had information regarding reception date, replacing the terms containing publication date by terms containing reception date. The policy variable was defined as whether there was a policy at the time of submission.

The Effect of the 2017 Psychological Science Editorial. We examined the effect of Lindsay (2017)'s editorial in *Psychological Science* that stated data must be shared with reviewers using the following GLM:

$$\begin{aligned} \text{logit(availability statement)} &= \beta_0 \\ &+ \beta_1 \text{badges} \\ &+ \beta_2 \text{mandatory sharing with reviewers} \\ &+ \beta_3 \text{reception data} \\ &+ \beta_4 \text{reception date} * \text{badges} \\ &+ \beta_5 \text{reception data} * \text{mandatory} \\ &\text{sharing with reviewers}, \end{aligned} \tag{9}$$

where "badges" was the possibility of obtaining a badge (0 = no possibility) and "mandatory sharing with reviewers" was a binary variable (0 = not mandatory).

Results

In this section, we examine the results of the three research questions in turn. Each of these sections is split into descriptive and inferential results. Unregistered exploratory analyses are presented last.

Research Question 1: How do data sharing statement rates evolve in the long term?

Descriptive results

Overall, articles published in a journal with a policy had a data sharing statement 42.35% of the time, while articles published in a journal with no policy had a data sharing statement 4.14% of the time. Table 6 provides more detail per journal.

Policies appear to lead to an increase in data sharing statements. As we aim to investigate the long-term effect of policies, pre- and post-policy proportions do not suffice. Figure 1 shows changes in the proportion of data sharing statements in each journal over time.

Regarding the long-term effect of the policies in *Cognition* and *Psychological Science*, *Cognition* approaches

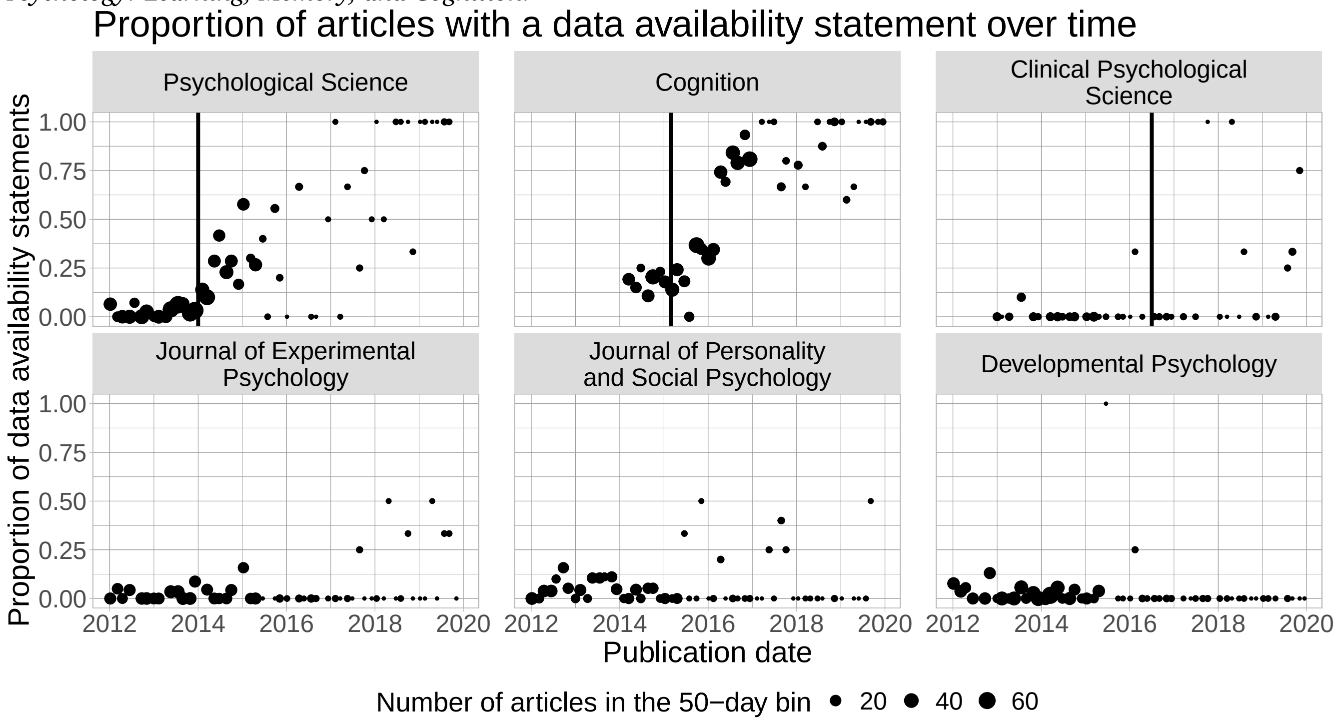
Table 6

Proportion and frequency of data sharing statements per journal and policy.

Policy status	Proportion and frequency of data sharing statements in each journal					
	PS	*Cognition*	CPS	JEPLMC	JPSP	DP
Policy absent	13/517 (2.51%)	30/167 (17.96%)	2/118 (1.69%)	17/558 (3.05%)	27/490 (5.51%)	17/712 (2.39%)
Policy present	118/395 (29.87%)	276/501 (55.09%)	10/58 (17.24%)	-	-	-

Figure 1

Proportion of articles with a data availability statement over time. Bold vertical lines indicate the implementation of a policy. Points show the proportion of articles with a data sharing statement within a 50-day bin. Point size is proportional to the number of articles in the bin. "Journal of Experimental Psychology" refers to Journal of Experimental Psychology: Learning, Memory, and Cognition.



100% sharing despite some dips in the trend in late 2017 and in 2019. Similarly, the increasing trend in *Psychological Science* highlighted by Kidwell et al. (2016) continues, although the variability increases. This is likely due to the reduction in the number of articles per bin, caused by the current study using only a sample of articles rather than the population. The proportion of articles with data sharing statement approaches 100% around mid-2018. *Clinical Psychological Science* also shows an increase in data sharing statements, albeit this happens around two years following the implementation of badges. The three other journals maintain low data sharing rates.

Inferential results

Generalised Additive Model. We fit a GAM to examine non-linear trends. The predictors were smooth functions of time and the interaction of time with policy and with journals, with the baseline journal being *Journal of Experimental Psychology: Learning, Memory, and Cognition*. The deviance explained was of 39.5%, indicating a moderately good fit to the data. The smooth terms are detailed in Table 7. Significance is assessed using a Chi-square test with k' degrees of freedom. k' values represent the maximum degree a polynomial fitting the data could take. They were of 9 for publication date and of 10 for all others. Effective degrees of freedom approximate the degree of a polynomial that would be needed to model the relationship parametrically. Lower effective degrees of freedom indicate larger smoothing

(Wood, 2017). Here, most effective degrees of freedom are much lower than the corresponding k' value, indicating important smoothing.

Publication date significantly impacts the probability of inclusion of a data sharing statement; the effective degrees of freedom are 1, indicating this relationship would be best modelled by a linear trend. *Cognition* and *Psychological Science* have significantly different trends from the baseline. Policy has a very limited impact on the probability of inclusion of a data sharing statement. This may be due to this variable being closely related to publication time, with most of the variation in the data being better explained by time rather than policy. As GAMs do not provide parameter estimates, we examine the highlighted trends using Figure 2.

The probability of inclusion of a data sharing statement increased steeply in *Cognition* during the year following the implementation of a policy and stabilised between 75% and 100%. The increase was less steep in *Psychological Science*, but rates also reached 100%. The changes in *Clinical Psychological Science* are unstable, most likely due to the smaller sample size. The other three journals maintain stable, close to 0 rates.

The GAM allowed for a detailed visual examination of the time trends. We now attempt to quantify these using generalised linear models, which constrain the trend to a certain polynomial degree.

Generalised Linear Model. As time had a strong linear effect in the GAM, we first fit a linear GLM with predictors time, journal, and their interaction. The baseline was *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Parameter estimates are detailed in Table 8. Slope estimates and standard errors are rounded to 4 decimal places as they describe changes for every day. Effects on such a small scale are very limited and therefore small magnitude changes are important.

There is an overall effect of time, where the log odds of inclusion of a data sharing statement increase by 0.0011 every day. Translating this to odds, there is an exp(0.0011) = 1.0011-fold (CI = [1.0005, 1.0017]) increase every day. A 30-day estimate for the odds of inclusion of a data sharing statement was computed by calculating the estimate of the increase in log odds (30*0.0011), then taking its exponent for interpretation on the odds scale. This showed a 1.0344-fold (CI = [1.0152, 1.0535]) increase every 30 days. Journals that have implemented data sharing policies have a positive offset to the log odds of time, indicating a stronger trend than the baseline. This suggests policy may have an effect, although this cannot be formally inferred using the model as policy was not included among the predictors due to collinearity. As our estimates are on a one-day scale, we illustrate the modelled time trend in each journal in Figure 3 to help grasp its magnitude.

Polynomial Generalised Linear Model. We have examined the linear effect of time using a GLM, as suggested by the effective degrees of freedom of publication date in the GAM. As the GAM had some effective degrees of freedom indicating a possible cubic relationship in the interactions of time and journals, we also fit a cubic GLM. To assess its relevance, we compared it to the GLM and the GAM using the BIC (Table 9). It was less appropriate than the GLM for the same predictors as it had a higher BIC. It therefore was discarded.

Model Selection. We compared the three models using the BIC (Table 9). The GLM was the best model for the data, accounting for complexity and model fit. This suggests linear trends are sufficient to model data availability statement rates. We conclude time increases the chances of having a data availability statement, particularly in journals that have a policy.

Research Question 2: How do effective data sharing rates change with time?

Among the 510 articles that had a data sharing statement, 496 stated the data was available, 4 stated the data was partly available (2 in *Clinical Psychological Science*, 1 in *Psychological Science*, 1 in *Journal of Personality and Social Psychology*), and 10 stated the data was unavailable (2 in *Clinical Psychological Science*, 8 in *Psychological Science*). All statements of partial or unavailable data were in articles published after the implementation of a policy, apart from the statement in *Journal of Personality and Social Psychology*. In this section, we focus on the subset of articles that had a data sharing statement indicating the data is available.

Descriptive Results

Data was deemed accessible, complete, and understandable in 40.32% of articles that stated their data was shared. Among articles published in the absence of any policy, only 12.38% were reusable. This increased to 47.83% in articles published in the presence of a policy. Table 10 provides more detail per journal.

Table 11 provides more detail for each reusability criterion per policy status.

Policies are associated with higher reusability rates, albeit these are far from perfect. We investigate the change of rate in more detail by looking at time trends (Figure 4).

Visual examination of Figure 4 suggests time does not substantially change reusability rates. Policies increase these rates, as the *Cognition* and the *Psychological Science* panels show post-policy rates scattering around 50%, while the pre-policy rates in these two journals and the rates in the remaining three journals are close

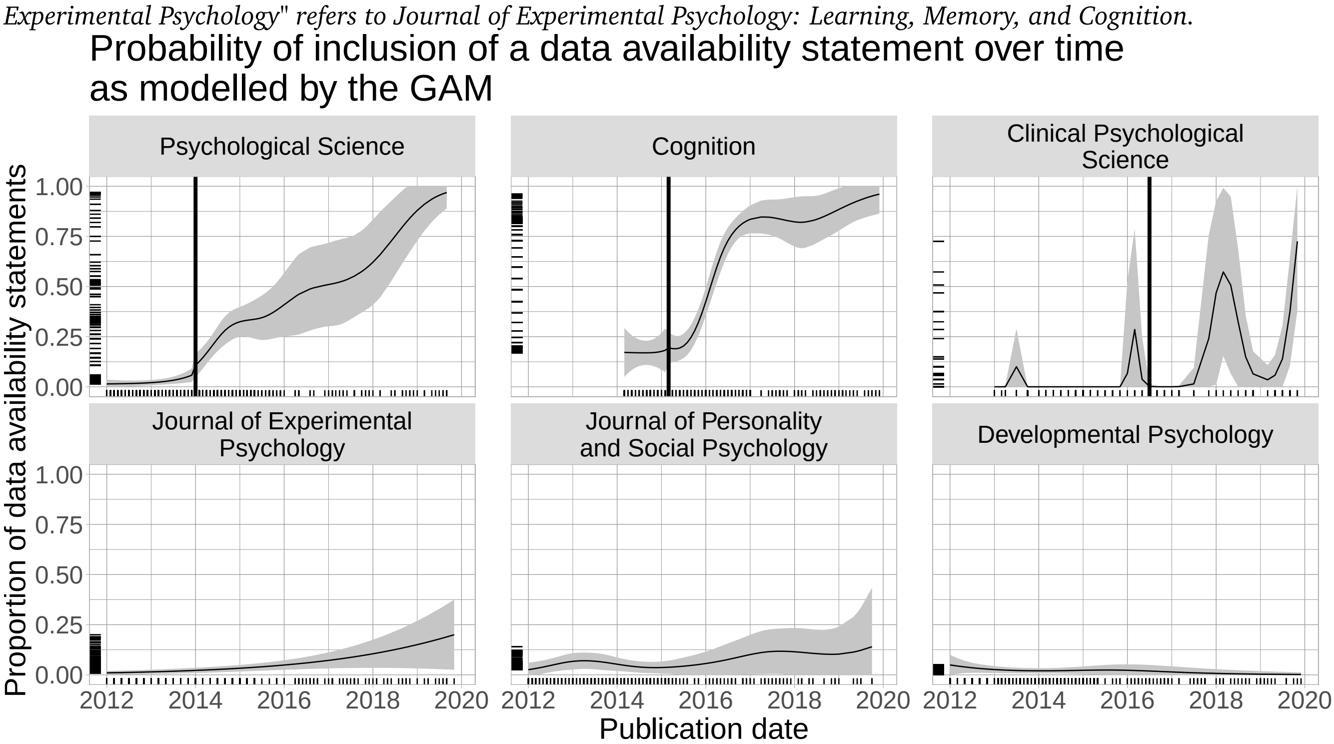
Table 7

Effective degrees of freedom and statistical significance of the smooth terms in the GAM modelling data availability statement.

Term	Effective degrees of freedom	Chi-square	p-value
Publication date	1.00	13.13	0.0003
Publication date offset for PS	3.84	17.43	0.0019
Publication date offset for *Cognition*	4.77	40.93	0.0000
Publication date offset for CPS	8.64	13.72	0.1286
Publication date offset for JPSP	5.33	11.21	0.0942
Publication date offset for DP	3.50	9.58	0.0508
Publication date offset when there is a policy	4.98	3.97	0.5517

Figure 2

Probability of inclusion of a data availability statement over time as modelled by the GAM. Bold vertical lines represent the implementation of a policy. Solid lines represent the trend estimate for the journal. Grey shading represents the confidence bands. The rugs on the x-axis and y-axis show the concentration of articles at the given time. "Journal of Experimental Psychology" refers to Journal of Experimental Psychology: Learning, Memory, and Cognition.



to 0. *Clinical Psychological Science* does not have enough observations to make conjectures.

Inferential Results

Generalised Additive Model. We first fit a GAM to examine the non-linear trends. The predictors were smooth functions of time and time given journal, with *Journal of Experimental Psychology: Learning, Memory, and Cognition* as the baseline journal. The deviance explained was of 20%, indicating a bad fit to the data. The interactions of time and journals had effective degrees of freedom between 2 and 4.31 for a k' of 10, indicating a large amount of smoothing, and p-values between 0.1980 and 0.9453. The main effect of time had effective degrees of freedom 8.58 for a k' of 9, showing a lack of smoothing for that term, and a p-value of 0.0902. As GAMs do not provide parameter estimates, we examine the modelled relationships in Figure 5.

Table 8

Parameter estimates and statistical significance of the GLM modelling data availability statements.

Parameter	Estimate (log odds)	Standard error	z value	p-value
Intercept (JEPLMC)	-21.93	5.17	-4.24	0.0000
Publication date slope (JEPLMC)	0.0011	0.0003	3.63	0.0003
Intercept offset for PS	-21.60	6.50	-3.32	0.0009
Intercept offset for *Cognition*	-33.77	7.16	-4.72	0.0000
Intercept offset for CPS	-14.16	10.38	-1.37	0.1723
Intercept offset for JPSP	12.83	7.02	1.83	0.0676
Intercept offset for DP	29.33	10.25	2.86	0.0042
Slope offset for PS	0.0014	0.0004	3.67	0.0002
Slope offset for *Cognition*	0.0022	0.0004	5.09	0.0000
Slope offset for CPS	0.0008	0.0006	1.37	0.1699
Slope offset for JPSP	-0.0007	0.0004	-1.74	0.0818
Slope offset for DP	-0.0018	0.0006	-2.87	0.0041

Figure 3

Probability of inclusion of a data availability statement over time as modelled by the GLM. Bold vertical lines represent the implementation of a policy. Solid lines represent the trend estimate for the journal. Grey shading represents the confidence bands. The rugs on the x-axis and y-axis show the concentration of articles at the given time. "Journal of Experimental Psychology" refers to Journal of Experimental Psychology: Learning, Memory, and Cognition.

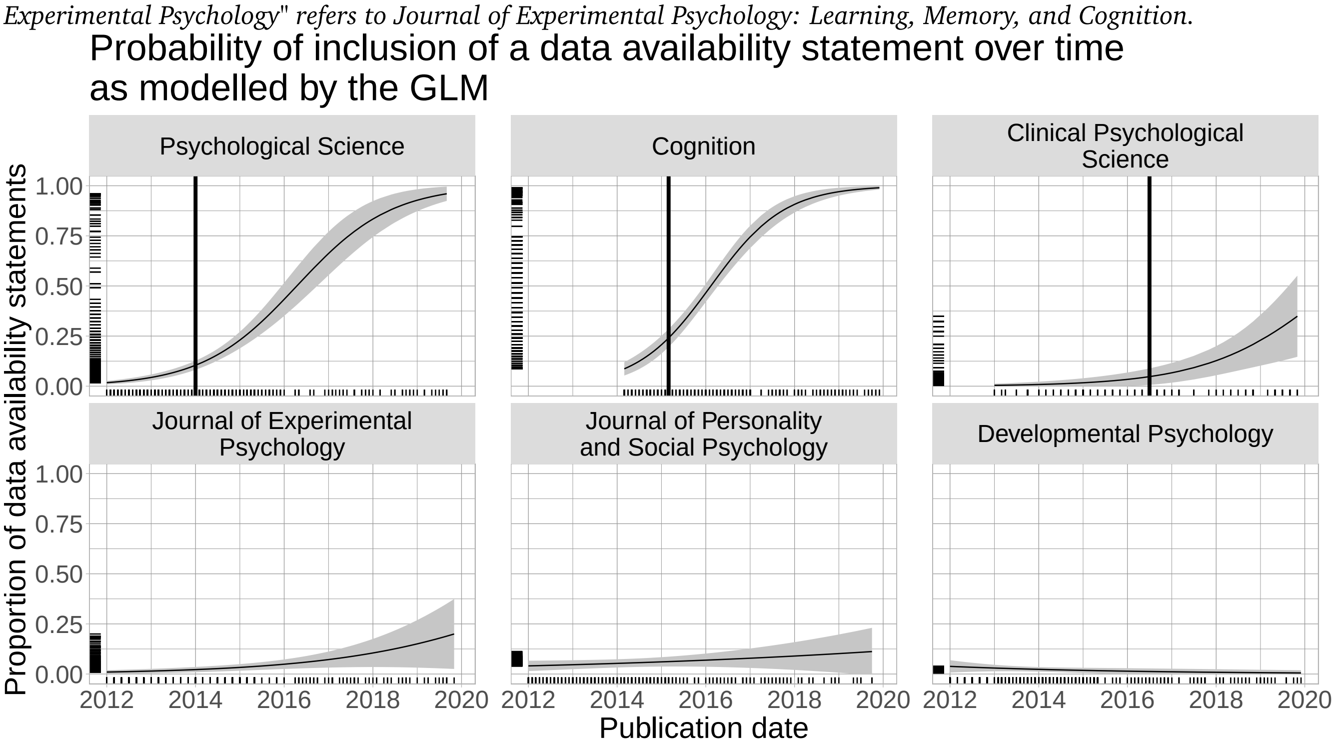


Table 9

BIC of the models for data availability statements.

Model	BIC
GAM	2031.897
GLM	1955.678

Figure 5 highlights large amounts of uncertainty and no clear trend. The GAM overfits the data and hinders interpretability. We therefore fit a GLM to constrain the trends linearly and improve interpretability.

Generalised Linear Model. The GLM had time, jour-

Table 10

Proportion and frequency of reusable data per journal and policy.

Policy status	Proportion and frequency of reusable data in each journal					
	PS	*Cognition*	CPS	JEPLMC	JPSP	DP
Policy absent	2/13 (15.38%)	2/30 (6.67%)	0/2 (0.00%)	5/17 (29.41%)	3/26 (11.54%)	1/17 (5.88%)
Policy present	55/109 (50.46%)	131/276 (47.46%)	1/6 (16.67%)	-	-	-

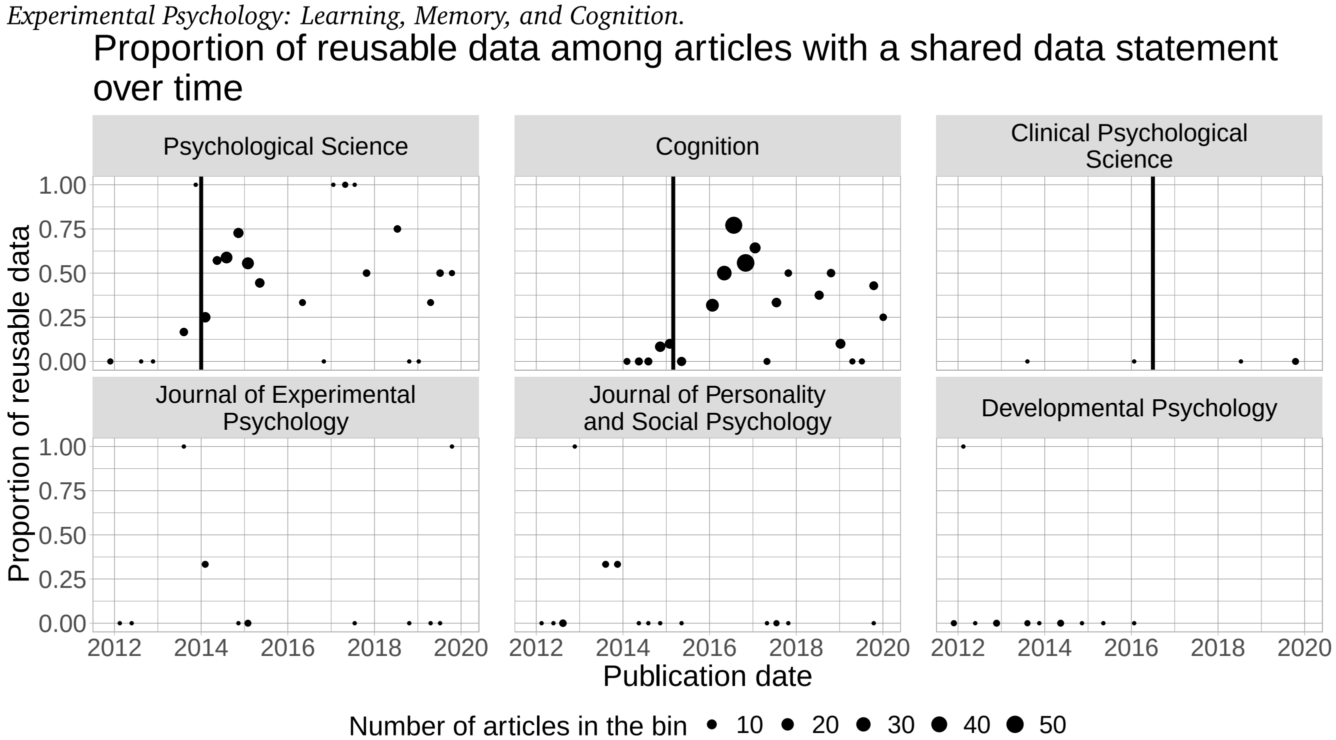
Table 11

Proportion and frequency of articles with a shared data statement that meet each reusability criterion

Policy status	Proportion and frequency of each data reusability criterion		
	Downloaded data out of articles with a shared data statement	Complete data out of downloaded data	Understandable data out of complete data
Policy absent	67/105 (63.81%)	19/67 (28.36%)	13/19 (68.42%)
Policy present	368/391 (94.12%)	232/368 (63.04%)	187/232 (80.60%)

Figure 4

Proportion of reusable data among articles with a shared data statement over time. Bold vertical lines indicate the implementation of a policy. Points show the proportion of articles with a data sharing statement within a 90-day bin. Point size is proportional to the number of articles in the bin. "Journal of Experimental Psychology" refers to Journal of Experimental Psychology: Learning, Memory, and Cognition.



nal, and their interaction as predictors. It showed slope parameter estimates lesser than 0.001, all standard errors were larger than their corresponding estimate, and p-values were between 0.372 and 0.932. This shows very uncertain differences in reusability across journals or time. It is possible this model was too constrained to identify the trend, so we next fit a polynomial GLM to improve the flexibility.

Polynomial Generalised Linear Model. The polynomial GLM was a cubic polynomial. Some fitted values were numerically adjusted by R, indicating issues in model fit as it contained values out of bounds. The BIC of this model was higher than that of the GAM and the GLM (Table 12). We therefore discarded this model as its complexity did not improve its fit.

Model Selection. We compared the three models using the BIC (Table 12). The GAM was the best model for the data, accounting for complexity and model fit.

Figure 5

Probability of data reusability over time as modelled by the GAM. Vertical lines represent the implementation of a policy. Solid lines represent the trend estimate for the journal. Grey shading represents the confidence bands. The rugs on the x-axis and y-axis show the concentration of articles at the given time. "Journal of Experimental Psychology" refers to Journal of Experimental Psychology: Learning, Memory, and Cognition.

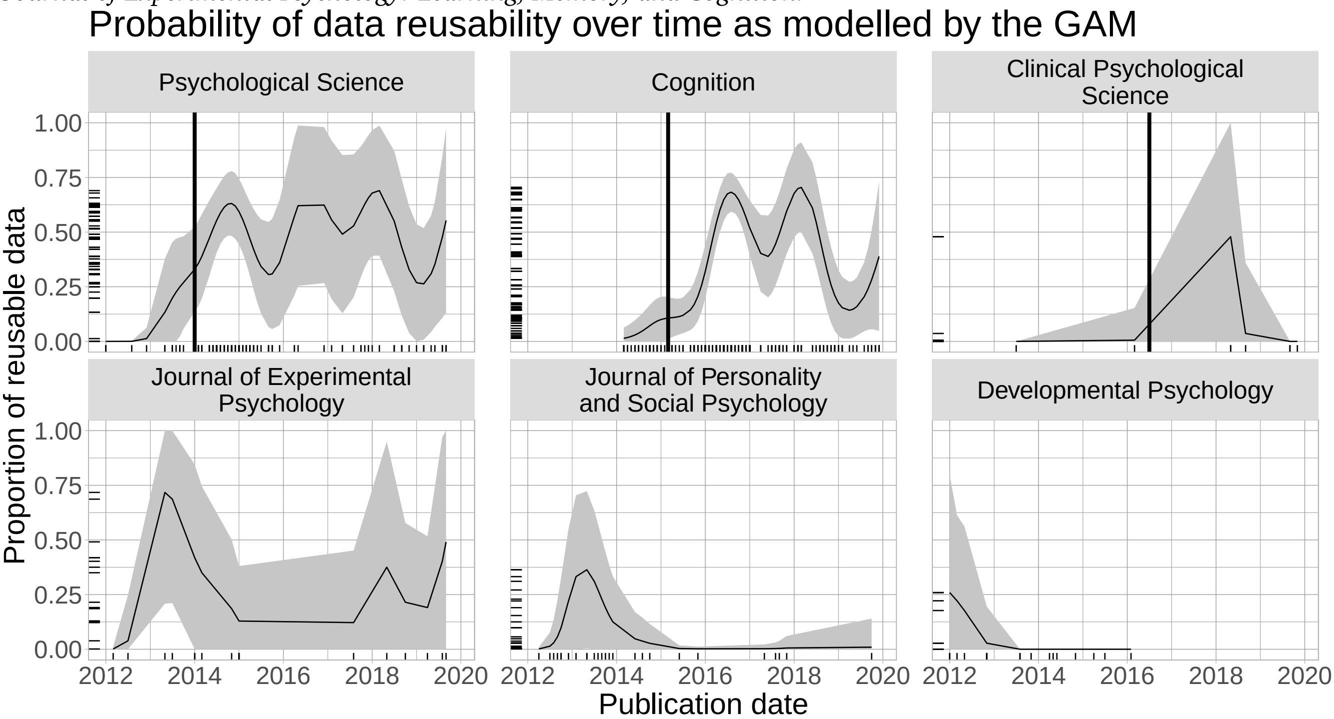


Table 12

BIC of the models for data reusability.

Model	BIC
GAM	689.6847
GLM	704.4952
Polynomial GLM	706.9807

This suggests linear trends are not sufficient to model data availability statements; however, the GAM does not highlight clear trends regarding data reusability. The effects of time on data reusability are very uncertain.

Research Question 3: What is the impact of the introduction of badges in *Clinical Psychological Science*?

In this section, we focus on the subset of articles that were published in *Clinical Psychological Science*.

Descriptive Results

Before badges were implemented in *Clinical Psychological Science*, 1.69% of articles had a data availability statement. This increased to 17.24% over the three and a half years after the implementation of badges. Figure 6 shows the evolution over time. Most data sharing statements are over a year after the implementation of badges, suggesting a delayed effect or no effect of policy.

Figure 6

Data availability statements in Clinical Psychological Science over time. Each point corresponds to an article. The bold vertical line represents the implementation of badges.

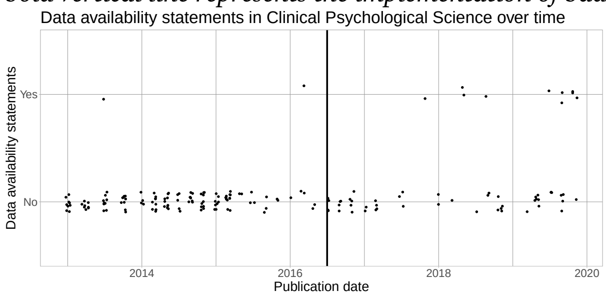
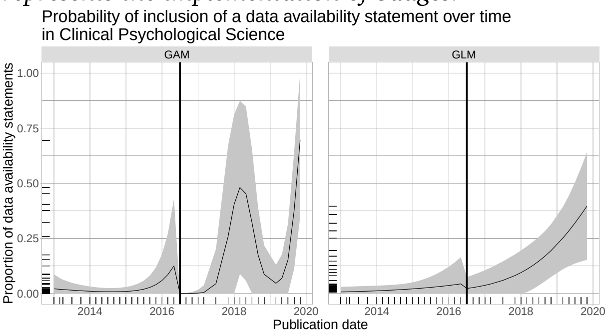


Figure 7

Probability of a data availability statement over time in Clinical Psychological Science. The shaded areas correspond to the confidence interval. The bold vertical line represents the implementation of badges.



Inferential Results

Generalised Linear Model. The trend was first examined by fitting a GLM analogous to that used by Hardwicke et al. (2018). Parameter estimates are detailed in Table 13. The effects of time and policy on the presence of data availability statements in *Clinical Psychological Science* were very uncertain.

Generalised Additive Model. We then fit a GAM to examine the time trend in more detail. The deviance explained was 42.2%, showing a moderately good fit for the data. Smooth terms are detailed in Table 14. k' was 9 for publication date and 10 for the interaction of publication date and policy. The effective degrees of freedom show a large amount of smoothing.

There was no significant effect of time or policy. The effective degrees of freedom of publication date suggest this effect could be modelled by a quadratic polynomial.

Polynomial Generalised Linear Model. Due to the GAM's interaction term's effective degrees of freedom, we fit a quadratic GLM to the data. We compared the model to the GAM and the GLM using the BIC to assess whether it yielded an improvement in fit (Table 15). The BIC was higher than that of the GAM or GLM, suggesting the model to be overly complex for the data. It was therefore discarded.

Unregistered Exploratory Results

The Effects of Time and Policy when Considering Reception Date

It may be more reasonable to operationalise time as reception date rather than full publication date in order to adequately compare journals. Bias in favour of journals with shorter delays may have been introduced as their articles are more recent and thus may reflect a different approach to open science.

First, we check if there is a difference between journals in the delay between submission and publication. The median delay was of 400 days (IQR = 226.75) and varied across journals ($F(5, 1032) = 8.93$, $p < 0.001$). Median delays for each journal are presented in Figure 8.

We fit the GAM and the GLMs examining data sharing statement rates again, defining policy and time using the reception date.

The GAM was a sum of smoothing functions for reception date, reception date given the journal, and reception date given policy status. It had deviance explained 32.1%. This is lower than the 39.5% of the original GAM, indicating a worse fit, although this may be due to the reduction in sample size. The smooth terms are detailed in Table 16.

The effective degrees of freedom for all offsets due to journals were 2, suggesting quadratic effects of time and its interaction with journals. We thus fit a quadratic GLM to the data. The quadratic GLM had BIC 1088.64 compared to the GAM having BIC 1054.05, showing the quadratic GLM to be overly complex without a sizeable improvement in fit. We thus discarded it and fit a linear GLM to examine simpler trends. Parameter estimates are detailed in Table 17.

The log odds (slope offsets) are negative for all journals, indicating the model predicts a lower rate of increase of availability statement rates over time in all journals compared to *Journal of Experimental Psychology: Learning, Memory, and Cognition*. In *Developmental Psychology* and *Journal of Personality and Social Psychology*, the size of the slope offset is larger than the size of the baseline slope and hence cancels it out, suggesting a decrease in data sharing statements over time. We illustrate the GLM and the GAM in Figure 9 to compare them and examine the trend modelled by the GAM.

The GAM may overfit the data in *Psychological Science* as shown by the variations in the curve. The GAM includes policy as a parameter, allowing to highlight the effect of policy in *Cognition* and *Clinical Psychological Science*, but not in *Psychological Science* due to a lack of pre-policy data. The current models and the original ones that consider publication date (Figure 2, Figure 3) show similar trends, in which journals that have implemented a policy have increasing rates of data sharing statements.

The Effect of the 2017 Editorial in Psychological Science

The data sharing statement rate in *Psychological Science* approaches 100%. This happens around a year after the implementation of mandatory data sharing with reviewers. As the trend is best modelled linearly (Table

Table 13

Parameter estimates and statistical significance for the GLM modelling data availability statements in Clinical Psychological Science

Parameter	Estimate	Standard error	z value	p-value
Intercept	-29.36	37.80	-0.78	0.437
Publication date	0.0016	0.0023	0.67	0.501
Policy	-21.09	43.48	-0.49	0.628
Publication date x Policy	0.0012	0.0026	0.46	0.646

Table 14

Effective degrees of freedom and significance of the smooth terms in the GAM modelling data availability statements in Clinical Psychological Science

Term	Effective degrees of freedom	Chi-square	p-value
Publication date	1.86	1.49	0.602
Publication date x policy	4.07	6.53	0.187

Figure 8

Distribution of time between reception date and full publication date in each journal. "Journal of Experimental Psychology" refers to Journal of Experimental Psychology: Learning, Memory, and Cognition.

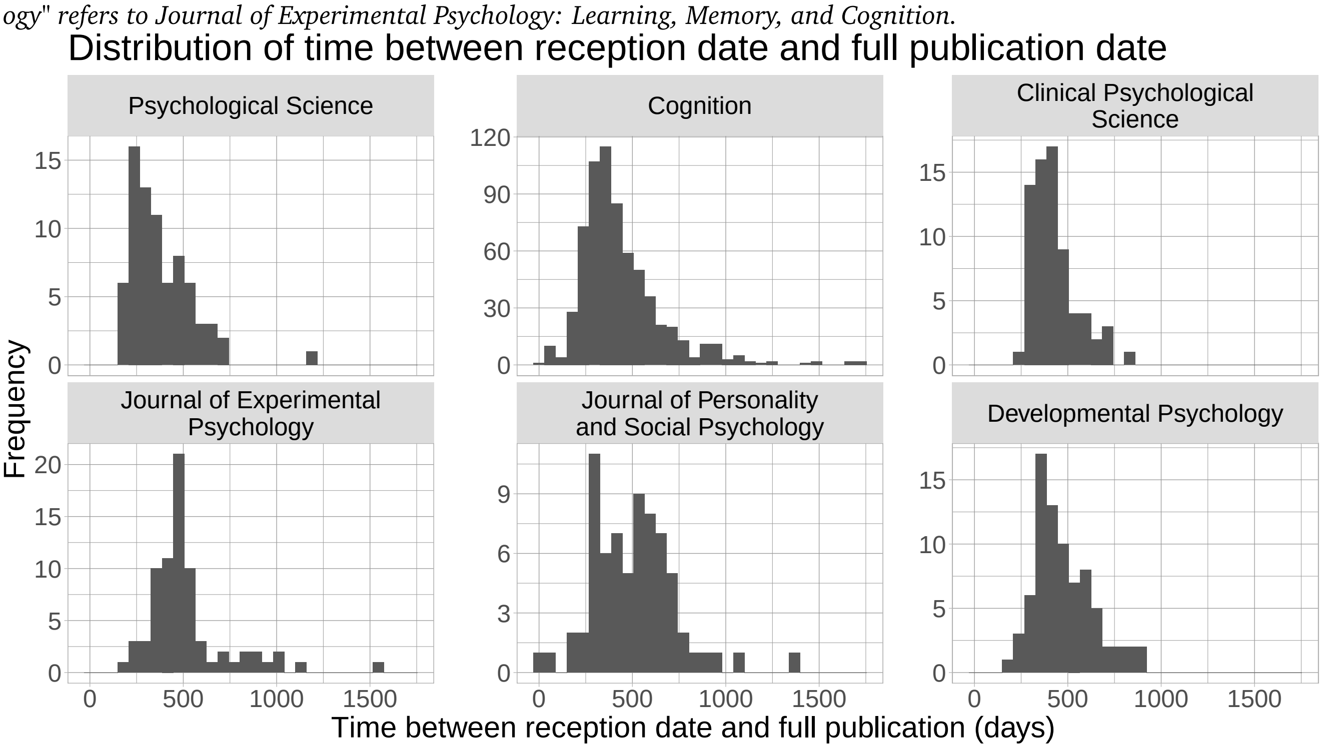


Table 15

BIC of the models for data reusability.

Model	BIC
GAM	86.5
GLM	86.15
Polynomial GLM	93.96

9), we examine the effect of this second policy using a GLM. We only used observations from *Psychological Science* that had information regarding the reception date, yielding a sample size of n = 75. The parameter estimates were highly uncertain, likely due to the very small sample size. Figure 10 illustrates the trend.

Due to the small number of observations, this es-

Table 16

Effective degrees of freedom and significance of the smooth terms in the GAM modelling data availability statements.

Term	Effective degrees of freedom	Chi-square	p-value
Reception date	1.00	6.40	0.0114
Reception date offset for PS	2.00	5.99	0.0500
Reception date offset for *Cognition*	2.00	16.19	0.0003
Reception date offset for CPS	2.00	3.65	0.1616
Reception date offset for JPSP	2.00	6.03	0.0492
Reception date offset for DP	2.00	6.44	0.0399
Reception date offset when there is a policy	6.81	38.52	0.0000

Table 17

Parameter estimates and significance of the GLM modelling data availability statements.

Parameter	Estimate (log odds)	Standard error	z value	p-value
Intercept	-58.53	22.43	-2.61	0.0091
Reception date slope	0.0033	0.0013	2.53	0.0114
Intercept offset for PS	28.72	24.48	1.17	0.2408
Intercept offset for *Cognition*	12.72	22.80	0.56	0.5768
Intercept offset for CPS	22.43	26.64	0.84	0.3997
Intercept offset for JPSP	68.74	28.14	2.44	0.0146
Intercept offset for DP	134.10	54.46	2.46	0.0138
Slope offset for PS	-0.0015	0.0014	-1.06	0.2916
Slope offset for *Cognition*	-0.0005	0.0013	-0.38	0.7046
Slope offset for CPS	-0.0013	0.0015	-0.839	0.4014
Slope offset for JPSP	-0.0040	0.0017	-2.436	0.0149
Slope offset for DP	-0.0081	0.0033	-2.446	0.0144

timate is crude. We however notice a change in the intercept following each policy, going from 0 to 50% after implementing badges, and from 50% to 100% after making sharing data with reviewers mandatory. Notably, there is only one article that does not state whether its data is shared after the second policy.

Discussion

Our study aimed to investigate the long-term impact of the implementation of data sharing policies in journals, examine the direct impact of the implementation of badges in *Clinical Psychological Science*, and identify trends in the reusability of shared data. Following the introduction of a data sharing policy in a journal, there was a substantial increase in data availability statements over time (Figure 2). Reusability was subject to a level change following policies (Figure 4). We examine these findings in the light of the literature and assess shortcomings of this study.

The Effects of Data Sharing Policies on Data Availability Statement Rates

The Long-term Effects of the Mandatory Data Sharing Policy in Cognition

Hardwicke, Mathur, et al. (2018) investigated the effect of the mandatory data sharing policy in *Cognition*. They found the policy to lead to a large increase in data sharing statements. Data sharing statement rates did not reach 100%, albeit they showed an increasing trend over time. These findings were limited by the fact they only considered articles published within two years following the implementation of the policy. We expanded on this by collecting a follow-up sample and showed the trend continued to increase to reach 100% of articles containing a data sharing statement (Figure 3). This suggests the mandatory data sharing policy was effective, but subject to an adjustment period. As this is a case study, we cannot generalise these results to all mandatory data sharing policies. We encourage future research to examine other similar policies (e.g. Marks, 2020) in order to draw conclusions about their overall effectiveness.

Figure 9

Probability that a given article contains a data availability statement. Bold vertical lines represent the implementation of a policy. Solid lines represent the trend estimate for the journal. Grey shading represents the confidence bands. The rugs on the x-axis and y-axis show the concentration of articles at the given time. "Journal of Experimental Psychology" refers to Journal of Experimental Psychology: Learning, Memory, and Cognition.

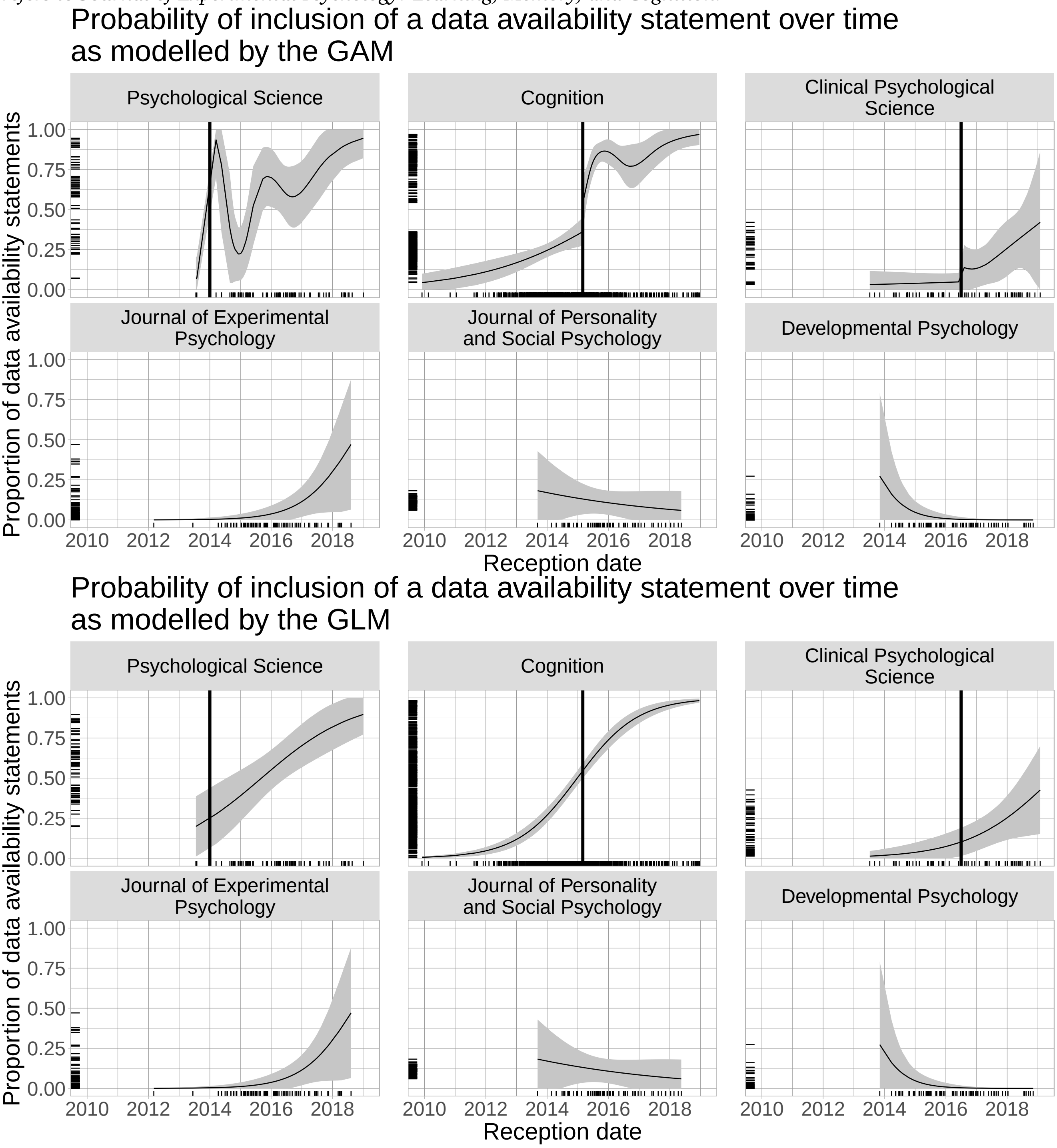
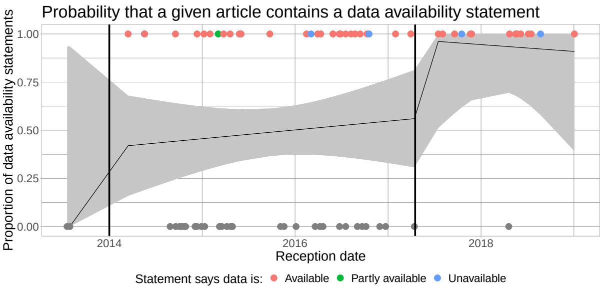


Figure 10

Probability that a given article contains a data availability statement. Bold vertical lines represent the implementation of a new policy. Solid lines represent the trend estimate. Grey shading represents the confidence bands. The points represent the observations, where 0 means no statement and 1 means there is a statement.



The Long-term Effects of the Implementation of Badges in Psychological Science

Mandatory data sharing is not the only way to increase data sharing statement rates. Kidwell et al. (2016) investigated the effect of badges to reward data sharing in *Psychological Science*. They found an increasing trend in data sharing statement rates following the introduction of badges, going from 12.8% in the articles published within 6 months after the implementation of badges to 39.4% in the articles published between 12 and 18 months after the implementation of badges. We expanded on their results by collecting a follow-up sample of articles. We showed data sharing statement rates continued to increase until they were close to 100% of articles containing a data sharing statement (Figure 3). It is important to note this might not be only the result of badges, as a second policy regarding data sharing was introduced in 2017 (Lindsay, 2017). This policy made it mandatory to share data with reviewers or provide an explanation as to why the data cannot be shared. All but one sampled articles submitted to *Psychological Science* after the implementation of this policy contain a data sharing statement. Although the sample size for which this was examined was too small to draw conclusions on the certainty of the effect of the second policy, we fit a GLM which showed a change in the trend: the estimate of the proportion of shared articles jumped from a little over 50% directly before the introduction of the policy to a little under 100% directly after (Figure 8). We encourage future research to verify these findings using a larger sample.

The Effects of the Implementation of Badges in Clinical Psychological Science

Our findings regarding the effect of badges in *Psychological Science* are confounded by the implementation of a second policy. Moreover, examining only *Psychological Science* is a case study which may not generalise to other similar policies. *Clinical Psychological Science*, which has implemented badges more recently, was examined here, allowing to compare the effect of badges in the two journals. There was a much weaker increase in *Clinical Psychological Science* than in *Psychological Science*. This suggests similar policies may have different effects; this is further supported by Rowhani-Farid and Barnett, 2018's findings that the implementation of badges in *Biostatistics* was also much less successful than in *Psychological Science*, albeit more successful than the implementation of badges in *Clinical Psychological Science*. We now examine possible reasons for which badges may have been less effective in *Clinical Psychological Science* than in *Psychological Science*.
There are various barriers to data sharing, such as legal barriers regarding anonymisation of data or preparing the data for sharing being too time-consuming (Houtkoop et al., 2018). These barriers vary across data types: for instance, anonymisation is easier to achieve for one-off online surveys than for longitudinal studies or physiological data (Audette et al., 2020; Choudhury et al., 2014). Predominant types of data and data sensitivity vary across psychological fields; therefore, while *Psychological Science*'s scope encompasses all fields in psychology and thus reflects general trends across psychology, *Clinical Psychological Science* focuses on one area only, for which data sharing may be more difficult to achieve due to higher sensitivity of the data (Tackett et al., 2017). More generally, this may highlight the differences in the prevalence of open practices across fields (for an overview in clinical psychology, see Tackett et al., 2017). We encourage further research to examine these differences and the reasons behind them in order to identify suitable interventions to increase open data.

Conclusions on the Effect of Policies on Data Availability Statement Rates

Our results shed light on the efficacy of policies in promoting data sharing statements. We have shown policies can lead to an increase in data sharing, but the effects are not immediate and are best described by an accelerating trend over time. Moreover, this does not work equally in all journals, as illustrated by the weak trend in *Clinical Psychological Science*. Further research is needed to examine a larger number of policies to compare incentivising (here, through badges) and mandating (for all readers or for reviewers) data sharing. This should also expand on how fields differ in existing policies and their efficacy, and what barriers are predominant in each field, in order to best tackle them and increase open data.

The Effects of Data Sharing Policies on Data Reusability

Data availability statements do not guarantee data is open in practice. Hardwicke, Mathur, et al. (2018) and Kidwell et al. (2016) have shown data to be reusable in at most 70% of articles with a data sharing statement published after the implementation of a policy. Among articles with a data sharing statement published in the absence of any policy, reusability rates were at most 22%. We expanded on these findings by examining reusability rates over time. We found reusability rates to be higher after the implementation of a policy, strengthening Kidwell et al. (2016)'s and Hardwicke, Mathur, et al. (2018)'s findings. However, no trend over time was identified: reusability rates seemed to stabilise around 50% after the implementation of a policy, regardless of the nature of the policy (Figure 4). It is important to note that our results are framed in the context of considering only articles with a shared data statement. Post-policy reusability rates remain constant, but not null. Therefore, as the proportion of articles with a shared data statement increases, the proportion of articles with reusable data when considering all published articles increases too. As data sharing statement rates approach 100%, however, stable reusability rates raise concerns regarding the efficacy of policies, as effective data sharing rates are far from perfect.

Limitations in Assessing Reusability

Reusability is however complex to assess. We examine limitations in its assessment in order to better understand the concerns raised above. We considered three criteria to define reusability: availability, completeness, and understandability. Availability of the data increased to almost 100% after the introduction of policies, while completeness and understandability increased much less. It is important to note that the latter two criteria are subjective. Completeness can be defined as all the data necessary to reproduce the analyses or all the data necessary to reproduce the analyses and sample descriptions. Understandability may depend on the amount of time spent examining the published article and investigating the data. Our pilot studies showed understandability raised many discrepancies, illustrating the subjectivity of the criterion. It is possible a larger proportion of articles would have been deemed reusable had more time been allocated to assessing the reusability of data. This weakness is shared with Kidwell et al. (2016)'s and Hardwicke, Mathur, et al. (2018)'s studies. We suggest the best way to examine data reusability is through attempting to use the data rather than through visual inspection. Additionally, further research may focus on determining consistent and detailed criteria for completeness and reusability. For instance, a list of acceptable file types might be established, as certain files cannot be opened without specialist software. Similarly, to standardise assessments of completeness and understandability, it would be helpful to determine whether variables (e.g. gender, age) that are not used in the analyses but are presented in the article are required for data completeness and if a codebook is necessary for all variables (e.g. if a "treatment" variable is coded 0 and 1, we may assume 0 to be a control condition due to conventions). In this article, highly stringent criteria were used: all mentioned variables were required for completeness, and data was only considered to be understandable if coding schemes were explicitly stated, regardless of conventions.

The limitations of the reusability assessment hinder conclusions on the true reusability rates, but not on their evolution. As all additional articles sampled for the current study were coded by the same researcher, the criteria used to assess completeness and understandability were constant. The lack of change over time therefore remains concerning: it is clear reusability rates are stable at a rate below 100%.

Conclusions on the Effect of Policies on Data Reusability

Our findings on data reusability highlight the limitations of the current policies. Although the intention to point readers towards data has increased, the efficacy of this is very limited; more guidance and support in sharing data could improve these rates further (Houtkoop et al., 2018; see Obels et al., 2020 for more detailed suggestions). Journals could also increase verification that data is shared efficiently before publishing an article or awarding a badge. Further research is needed to identify more accurate values of reusability rates.

Further Limitations

Overall, this study has extended previous findings by showing continuing increases in data sharing statements and a level change in data reusability. However, inference on the effect of policies was hindered by the operationalisation of time and the models used. We now examine these limitations.

The Operationalisation of Time

Reusability and data sharing statement rates were primarily examined using publication date rather than submission date. This may impact our results as studies published directly after the implementation will not have been subject to the policy during the publication process. Our examination of the trends using reception date was hindered by the much lower sample sizes as this excluded Kidwell et al., 2015b's dataset; we therefore encourage future research to complete their dataset with reception date information as well as include a larger number of articles published in the follow-up timeframe.

Unbalanced Sampling Over Time

This study reused previous data containing the population of articles published within the same time period and extended the results by sampling articles from a follow-up time period. This has led to highly unbalanced data over time. It is possible that, were the trends observed in the original datasets not continued, the drop in sampling may have not allowed us to identify this change in trend. Considering the models and plots, this appears to be primarily an issue in *Clinical Psychological Science*, where a change of trend was expected due to its implementation of badges, but closer examinations were inconclusive. This may be because the models were biased due to the large amount of pre-policy data, in which there was very little application of open science practices. Future research may consider rebalancing the data using either weighting or a larger dataset.

Determining the Effects of Policies

The preregistered models accounted both for policy and time. However, we have not been able to examine the trend using both these predictors due to collinearity issues. Moreover, causality could not be drawn from the preregistered nor the current models. Causal links between policies and increases in data sharing statements cannot be established from the models we fit. This is partly due to policy being removed from the model predictors due to collinearity, thus not allowing to formally infer changes in the trend observed in our graphical and descriptive estimates. More importantly, this study is observational in nature and it is possible that policies have led to a population shift, which could be the true cause of the rate changes (Hardwicke, Mathur, et al., 2018; Kidwell et al., 2016). The mandatory data sharing policy in *Cognition* may have deterred researchers from publishing in the journal, while the possibility of earning badges in *Psychological Science* and *Clinical Psychological Science* may have attracted researchers favourable to open science and data sharing. Although we have not been able to investigate these possibilities, they have been deemed unlikely by prior research (Hardwicke, Mathur, et al., 2018; Kidwell et al., 2016). To shed light on the causal effects of policies and potential mediating factors, further research is needed to investigate the attitudes towards policies and motivations of researchers publishing in journals with policies.

Conclusion

Our study is valuable as it is the first to explore the possibility of non-linear time trends in data sharing statements and time trends in data reusability. Our results are limited due to the operationalisation of time and reusability as well as by the small number of policies examined. Despite this, by examining follow-up samples, we have strengthened prior findings showing policies to lead to increases in data sharing. In addition to the research questions of interest, we have outlined the possible impact of the secondary policy in *Psychological Science*, suggesting making data sharing mandatory to some degree is much more effective in improving data sharing rates than incentivising it. The latter results are very exploratory in nature, were conducted on a very small sample; they are however promising, and further investigations are encouraged to examine a larger number of policies. This would help to identify the best interventions to increase efficient data sharing and move towards open science.

Author Contact

Correspondence concerning this article should be addressed to Judith Neve. E-mail: judith.nevedm@gmail.com. https://orcid.org/0000-0002-1168-0182.
Guillaume Rousselet orcID: https://orcid.org/0000-0003-0006-8729.

Conflict of Interest and Funding

There was no conflict of interest. The research was not funded as it was part of the first author's education.

Author Contributions

Judith A. Neve designed the project and measures, collected data, analysed and interpreted the data, and wrote the report. Guillaume A. Rousselet supervised the project, provided guidance at all stages and ensured it was methodologically sound, coded the pilot data, and edited the report.

Open Science Practices

This article earned the Preregistration+, Open Data, Open Code and the Open Materials badge for preregistering the analysis before data collection, and for making the data and materials openly available. It has been verified that the analysis reproduced the results presented in the article. The entire editorial process, including the open reviews, is published in the online supplement.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Alkharusi, H. (2012). Categorical Variables in Regression Analysis: A Comparison of Dummy and Effect Coding. *International Journal of Education*, *4*, 202–210. https://doi.org/10.5296/ije.v4i2.1962

Asendorpf, J. B., Conner, M., Fruyt, F. D., Houwer, J. D., Denissen, J. J. A., Fiedler, K., Fiedler, S., Funder, D. C., Kliegl, R., Nosek, B. A., Perugini, M., Roberts, B. W., Schmitt, M., van Aken, M. A. G., Weber, H., & Wicherts, J. M. (2013). Recommendations for Increasing Replicability in Psychology. *European Journal of Personality*, *27*(2), 108–119. https://doi.org/10.1002/per.1919

Audette, L. M., Hammond, M. S., & Rochester, N. K. (2020). Methodological Issues With Coding Participants in Anonymous Psychological Longitudinal Studies. *Educational and Psychological Measurement*, *80*(1), 163–185. https://doi.org/10.1177/0013164419843576

Bakker, M., van Dijk, A., & Wicherts, J. M. (2012). The Rules of the Game Called Psychological Science. *Perspectives on Psychological Science*, *7*(6), 543–554. https://doi.org/10.1177/1745691612459060

Baldwin, S. A., & Del Re, A. C. (2016). Open Access Meta-Analysis for Psychotherapy Research. *Journal of Counseling Psychology*, *63*(3), 249–260. https://doi.org/10.1037/cou0000091

Björk, B.-C., & Solomon, D. (2013). The publishing delay in scholarly peer-reviewed journals. *Journal of Informetrics*, *7*(4), 914–923. https://doi.org/10.1016/j.joi.2013.09.001

Chakrabarti, A., & Ghosh, J. K. (2011). AIC, BIC and Recent Advances in Model Selection. In P. S. Bandyopadhyay & M. R. Forster (Eds.), *Philosophy of Statistics* (pp. 583–605, Vol. 7). North-Holland. https://doi.org/10.1016/B978-0-444-51862-0.50018-6

Choudhury, S., Fishman, J. R., McGowan, M. L., & Juengst, E. T. (2014). Big data, open science and the brain: Lessons learned from genomics. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00239

Deschaght, P., Vintém, A. P., Logghe, M., Conde, M., Felix, D., Mensink, R., Gonçalves, J., Audiens, J., Bruynooghe, Y., Figueiredo, R., Ramos, D., Tanghe, R., Teixeira, D., Van de Ven, L., Stortelers, C., & Dombrecht, B. (2017). Large Diversity of Functional Nanobodies from a Camelid Immune Library Revealed by an Alternative Analysis of Next-Generation Sequencing Data. *Frontiers in Immunology*, *8*. https://doi.org/10.3389/fimmu.2017.00420

Ebersole, C. R., Axt, J. R., & Nosek, B. A. (2016). Scientists' Reputations Are Based on Getting It Right, Not Being Right. *Plos Biology*, *14*(5), e1002460. https://doi.org/10.1371/journal.pbio.1002460

Eich, E. (2014). Business Not as Usual. *Psychological Science*, *25*(1), 3–6. https://doi.org/10.1177/0956797613512465

Ethical principles of psychologists and code of conduct. (2017). https://www.apa.org. Retrieved December 5, 2020, from https://www.apa.org/ethics/code

Final NIH Policy for Data Management and Sharing. (2020). Retrieved December 9, 2020, from https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html

Forstmeier, W., Wagenmakers, E.-J., & Parker, T. H. (2017). Detecting and avoiding likely false-positive findings – a practical guide. *Biological Reviews*, *92*(4), 1941–1968. https://doi.org/10.1111/brv.12315

Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science*, *345*(6203), 1502–1505. https://doi.org/10.1126/science.1255484

Grolemond, G., & Wickman, H. (2011). Dates and Times Made Easy with lubridate. *Journal of Statistical Software*, *40*(3), 1–25.

Hanson, B., Sugden, A., & Alberts, B. (2011). Making Data Maximally Available. *Science*, *331*(6018), 649–649. https://doi.org/10.1126/science.1203354

Hardwicke, T. E., Mathur, M., MacDonald, K., Nilsonne, G., Banks, G. C., Kidwell, M. C., Hofelich Mohr, A., Clayton, E., Yoon, E. J., & Tessler, M. H. (2018a). Materials. Retrieved March 18, 2021, from https://osf.io/k2mdr/

Hardwicke, T. E., Mathur, M., MacDonald, K., Nilsonne, G., Banks, G. C., Kidwell, M. C., Hofelich Mohr, A., Clayton, E., Yoon, E. J., & Tessler, M. H. (2018b). Data. Retrieved March 18, 2021, from https://osf.io/6s8b3/

Hardwicke, T. E., Mathur, M. B., MacDonald, K., Nilsonne, G., Banks, G. C., Kidwell, M. C., Hofelich Mohr, A., Clayton, E., Yoon, E. J., Tessler, M. H., Lenne, R. L., Altman, S., Long, B., & Frank, M. C. (2018). Data availability, reusability, and analytic reproducibility: Evaluating the impact of a mandatory open data policy at the journal Cognition. *Royal Society Open Science*, *5*(8), 180448. https://doi.org/10.1098/rsos.180448

Hastie, T., & Tibshirani, R. (1987). Generalized Additive Models: Some Applications. *Journal of the American Statistical Association*, *82*(398), 371–386. https://doi.org/10.2307/2289439

Houtkoop, B. L., Chambers, C., Macleod, M., Bishop, D. V. M., Nichols, T. E., & Wagenmakers, E.-J. (2018). Data Sharing in Psychology: A Survey on Barriers and Preconditions. *Advances in Methods and Practices in Psychological Science*, *1*(1), 70–85. https://doi.org/10.1177/2515245917751886

Ioannidis, J. P. A. (2014). How to Make More Published Research True. *Plos Medicine*, *11*(10), e1001747. https://doi.org/10.1371/journal.pmed.1001747

John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological science*, *23*(5), 524–532.

Jones, K., & Almond, S. (1992). Moving Out of the Linear Rut: The Possibilities of Generalized Additive Models. *Transactions (Institute of British Geographers)*, *17*, 434–447. https://doi.org/10.2307/622709

Kidwell, M. C., Lazarevic, L., Baranski, E., Hardwicke, T. E., Piechowski, S., Falkenberg, L.-S., Kennett, C., Slowik, A., Sonnleitner, C., & Hess-Holden, C. L. (2015a). Coding the Articles: Materials. Retrieved March 18, 2021, from https://osf.io/8kt4b/

Kidwell, M. C., Lazarevic, L., Baranski, E., Hardwicke, T. E., Piechowski, S., Falkenberg, L.-S., Kennett, C., Slowik, A., Sonnleitner, C., & Hess-Holden, C. L. (2015b). Data. Retrieved March 18, 2021, from https://osf.io/u6g7t/

Kidwell, M. C., Lazarević, L. B., Baranski, E., Hardwicke, T. E., Piechowski, S., Falkenberg, L.-S., Kennett, C., Slowik, A., Sonnleitner, C., Hess-Holden, C., Errington, T. M., Fiedler, S., & Nosek, B. A. (2016). Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency. *PLOS Biology*, *14*(5), e1002456. https://doi.org/10.1371/journal.pbio.1002456

Lilienfeld, S. O. (2017). Clinical Psychological Science: Then and Now. *Clinical Psychological Science*, *5*(1), 3–13. https://doi.org/10.1177/2167702616673363

Lindsay, D. S. (2017). Sharing Data and Materials in Psychological Science: *Psychological Science*. https://doi.org/10.1177/0956797617704015

Longo, D. L., & Drazen, J. M. (2016). Data Sharing. *New England Journal of Medicine*, *374*(3), 276–277. https://doi.org/10.1056/NEJMe1516564

Marks, D. F. (2020). Increasing the transparency, openness and replicability of psychological research: Mandatory data sharing for empirical studies in the Journal of Health Psychology. *Journal of Health Psychology*, *25*(6), 729–732. https://doi.org/10.1177/1359105320916858

Martone, M. E., Garcia-Castro, A., & VandenBos, G. R. (2018). Data Sharing in Psychology. *The American psychologist*, *73*(2), 111–125. https://doi.org/10.1037/amp0000242

McKiernan, E. C., Bourne, P. E., Brown, C. T., Buck, S., Kenall, A., Lin, J., McDougall, D., Nosek, B. A., Ram, K., Soderberg, C. K., Spies, J. R., Thaney, K., Updegrove, A., Woo, K. H., & Yarkoni, T. (2016). How open science helps researchers succeed. *Elife*, *5*, e16800. https://doi.org/10.7554/eLife.16800

Morey, R. D., Chambers, C. D., Etchells, P. J., Harris, C. R., Hoekstra, R., Lakens, D., Lewandowsky, S., Morey, C. C., Newman, D. P., Schönbrodt, F. D., Vanpaemel, W., Wagenmakers, E.-J., & Zwaan, R. A. (2016). The Peer Reviewers' Openness Initiative: Incentivizing open research practices through peer review. *Royal Society Open Science*, *3*(1), 150547. https://doi.org/10.1098/rsos.150547

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., du Sert, N. P., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 0021. https://doi.org/10.1038/s41562-016-0021

Nielson, J. L., Guandique, C. F., Liu, A. W., Burke, D. A., Lash, A. T., Moseanko, R., Hawbecker, S., Strand, S. C., Zdunowski, S., Irvine, K.-A., Brock, J. H., Nout-Lomas, Y. S., Gensel, J. C., Anderson, K. D., Segal, M. R., Rosenzweig, E. S., Magnuson, D. S. K., Whittemore, S. R., McTigue, D. M., ... Ferguson, A. R. (2014). Development of a Database for Translational Spinal Cord Injury Research. *Journal of Neurotrauma*, *31*(21), 1789–1799. https://doi.org/10.1089/neu.2014.3399

Obels, P., Lakens, D., Coles, N. A., Gottfried, J., & Green, S. A. (2020). Analysis of Open Data and Computational Reproducibility in Registered Reports in Psychology. *Advances in Methods and Practices in Psychological Science*, 2515245920918872. https://doi.org/10.1177/2515245920918872

O'Brien, R. M. (2012). Visualizing Rank Deficient Models: A Row Equation Geometry of Rank Deficient Matrices and Constrained-Regression. *PLoS ONE*, *7*(6). https://doi.org/10.1371/journal.pone.0038923

Perrino, T., Beardslee, W., Bernal, G., Brincks, A., Cruden, G., Howe, G., Murry, V., Pantin, H., Prado, G., Sandler, I., & Brown, C. H. (2015). Toward Scientific Equity for the Prevention of Depression and Depressive Symptoms in Vulnerable Youth. *Prevention Science*, *16*(5), 642–651. https://doi.org/10.1007/s11121-014-0518-7

Piwowar, H. A., & Vision, T. J. (2013). Data reuse and the open data citation advantage. *Peerj*, *1*, e175. https://doi.org/10.7717/peerj.175

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. Vienna, Austria.

Rowhani-Farid, A., & Barnett, A. G. (2018). Badges for sharing data and code at Biostatistics: An observational study. (7:90). https://doi.org/10.12688/f1000research.13477.2

Schwarz, G. (1978). Estimating the Dimension of a Model. *Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

Science Journals: Editorial policies. (2018). Science | AAAS. Retrieved December 5, 2020, from https://www.sciencemag.org/authors/science-journals-editorial-policies

Simpson, G. L. (2018). Modelling Palaeoecological Time Series Using Generalised Additive Models. *Frontiers in Ecology and Evolution*, *6*, 149. https://doi.org/10.3389/fevo.2018.00149

Sloman, S. A. (2015). Opening editorial: The changing face of Cognition. *Cognition*, *135*, 1–3. https://doi.org/10.1016/j.cognition.2014.11.004

Sterling, T. D. (1959). Publication decisions and their possible effects on inferences drawn from tests of significance—or vice versa. *Journal of the American statistical association*, *54*(285), 30–34.

Stockemer, D., Koehler, S., & Lentz, T. (2018). Data Access, Transparency, and Replication: New Insights from the Political Behavior Literature. *PS: Political Science & Politics*, *51*(4), 799–803. https://doi.org/10.1017/S1049096518000926

Stodden, V., Seiler, J., & Ma, Z. (2018). An empirical analysis of journal policy effectiveness for computational reproducibility. *Proceedings of the National Academy of Sciences*, *115*(11), 2584–2589. https://doi.org/10.1073/pnas.1708290115

Tackett, J. L., Lilienfeld, S. O., Patrick, C. J., Johnson, S. L., Krueger, R. F., Miller, J. D., Oltmanns, T. F., & Shrout, P. E. (2017). It's Time to Broaden the Replicability Conversation: Thoughts for and From Clinical Psychological Science. *Perspectives on Psychological Science*, *12*(5), 742–756. https://doi.org/10.1177/1745691617690042

Vanpaemel, W., Vermorgen, M., Deriemaecker, L., & Storms, G. (2015). Are We Wasting a Good Crisis? The Availability of Psychological Research Data after the Storm. *Collabra: Psychology*, *1*(1), Art. 3. https://doi.org/10.1525/collabra.13

Vines, T. H., Albert, A. Y. K., Andrew, R. L., Débarre, F., Bock, D. G., Franklin, M. T., Gilbert, K. J., Moore, J.-S., Renaut, S., & Rennison, D. J. (2014). The Availability of Research Data Declines Rapidly with Article Age. *Current Biology*, *24*(1), 94–97. https://doi.org/10.1016/j.cub.2013.11.014

Wang, B., Ding, X., & Wang, F.-Y. (2017). Determination of Polynomial Degree in the Regression of Drug Combinations. *Ieee-Caa Journal of Automatica Sinica*, *4*(1), 41–47. https://doi.org/10.1109/JAS.2017.7510319

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

Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R, Second Edition*. CRC Press.

Zoltowski, D. M., & Pillow, J. W. (2018). Scaling the Poisson GLM to massive neural datasets through polynomial approximations. *Advances* *in neural information processing systems*, *31*, 3517–3527.

Zotero. (2020). https://Zotero.org