

Data is not available upon request

Ian Hussey

Many journals now require data sharing and require articles to include a Data Availability Statement. However, several studies over the past two decades have shown that promissory notes about data sharing are rarely abided by and that data is generally not available upon request. This has negative consequences for many essential aspects of scientific knowledge production, including independent verification of results, efficient secondary use of data, and knowledge synthesis. I assessed the prevalence of data sharing upon request in articles employing the Implicit Relational Assessment Procedure published within the last 5 years. Of 52 articles, 42% contained a Data Availability Statement, most of which stated that data was available upon request. This rose from 0% in 2018 to 100% in 2022, indicating a change in journals' policies. However, only 27% of articles' authors actually shared data. Among articles stating that data was available upon request, only 17% shared data upon request. The presence of Data Availability Statements was not associated with higher rates of data sharing ($p = .55$), indicating a lack of adherence to journals' policies. Results replicate those found elsewhere: data is generally not available upon request, and promissory Data Availability Statements are typically not adhered to. Issues, causes, and implications are considered.

Verifiability is a cornerstone of the scientific method. Sharing research data is crucial for the advancement of scientific knowledge through verification. Data sharing promotes transparency, reproducibility, and credibility in scientific research, which are essential for maintaining the integrity of the scientific process (Munafò et al., 2017). One way in which journals have sought to promote data sharing is through the use of Data Availability Statements, which provide information on the availability and accessibility of research data. In this article, I examine the prevalence of both Data Availability Statements and actual data sharing upon request and the relationship between them.

It is increasingly common for journals to require Data Availability Statements to be reported in submissions. Typically, these journal policies require that a URL to the publicly available data is reported in the manuscript or, failing that, the authors state that the data is¹ available upon request. Naturally, policies also allow for situations where it is not possible to share the data for stated reasons. These policies therefore build on the same principle that many funding

organisations have built their data-sharing policies around, namely that data should be "as open as possible, and as closed as necessary" (European Commission, 2023).

Journal policies requiring and explicating data sharing are to be applauded (e.g., Evans, 2022), as data sharing is essential to independent verification of results, efficient secondary use of data, and knowledge synthesis. Journals are also joined by both professional societies (e.g., APA, APS, and ACBS) and the world's largest research funding agencies (e.g., NIH and EU) in encouraging or requiring data sharing (Nunes, 2021).

The motivation for this study, and the choice of the specific literature that I examined, came from my own experience of the difficulty of obtaining data upon request. Recently, I received peer reviews for a manuscript I wrote that meta-analysed the reliability of the Implicit Relational Assessment Procedure (IRAP; for reliability generalisation meta-analysis see Hussey & Drake, 2020). Reviewers raised the concern that the data, which came from two research groups, may not be representative of IRAP data collected in other labs. To try to address this point, I contacted

¹ I use data as singular throughout, following both its modal usage (Google Trends, 2023) and the recommendations of leading style guides for the last decade (Rogers, 2012).

other researchers who had published research using the IRAP to ask them to contribute their data to the meta-analysis. Often, authors could not be contacted, didn't reply, or declined, even when their published articles contained explicit statements that they would share data upon request.

Examination of the literature showed that this anecdotal experience was the norm rather than the exception. Over six decades ago, Wolin (1962) observed that only 24% of 37 articles published in APA journals, which had policies requiring data sharing upon request, actually shared data upon request. Nearly two decades ago, Wicherts et al. (2006) replicated this effect, observing that only 25.7% of 249 articles published in APA journals shared data upon request. More recently, Tedersoo et al. (2021) examined data sharing upon request across many fields. Inspection of their openly available dataset revealed that 40.4% of 57 psychology articles shared data upon request. A sufficient number of such studies examining the rate of data sharing upon request across many areas of science have been conducted that there is, as of very recently, a systematic review of them. Hamilton et al. (2023) concluded that across 13 studies the observed rate of data sharing upon request ranges from 0% to 37% (no meta-analysis was conducted due to methodological differences between studies). Finally, rather than treating the rate of data sharing upon request as a static property, Vines et al. (2014) demonstrated that it decays over time: for every one-year increase in article age, the odds of the data set still being reported as extant (either shared upon request or declined to shared but reported to exist) decreased by 17%.

Only one study to date has examined the prevalence and efficacy of Data Availability Statements within the behavioral research community, finding a very low prevalence (6.2%: Lear et al., 2023). However, this study was limited to the inspection of a single journal, the *Journal of Contextual Behavioral Science*. This study seeks to provide convergent evidence via different means, by studying data sharing in publications using a task generated by and mostly used by the behavioral research community, agnostic to which journal it is published in. This study goes beyond Lear et al. (2023) by studying whether Data Availability Statements translate into actual data sharing upon request. Additionally, whereas Lear et al. (2023) examined articles published in a single year, the current work examines a five-year period to examine trends across them.

Method & Results

Data availability statement

All data and code to reproduce the results reported here are available at osf.io/8sp7e. All IRAP datasets that could be publicly shared are also available at that link. Some datasets that, due to the original author's requirements, cannot be publicly shared but can be shared privately are available upon request (yes, really!) from the author through my personal email

address (ian.hussey@icloud.com). This address should remain active over time even if I move between institutions. Other datasets were provided to me on the agreement that I would not share them with others, even privately. These datasets may be available from the original authors upon request. The existence and availability of each dataset, as well as contact details for the responsible persons, are documented in a separate OSF project (osf.io/nugzb).

Preregistration

This study was not preregistered and its results should therefore be considered to be exploratory.

Ethical approval

Whether or not this study required ethical approval was discussed extensively with peers and with consultation of ethical guidelines (i.e., those by the American APA: American Psychological Association, 2016; the German DPGs: Deutschen Gesellschaft für Psychologie, 2022; and the British BPS: Oates et al., 2021) prior to its conduction. The consensus among these guidelines was that this work did not constitute "human subjects research" and therefore did not require approval. Previous studies assessing data sharing upon request have adopted a similar position (e.g., Vines et al., 2014). Meta-science research such as this is very similar to the conduction of a meta-analysis. In most jurisdictions, including that in which this work was conducted, meta-analyses do not require ethical approval because they have no human subjects. To draw a closer comparison that provides an intuition pump: individual participant data meta-analyses (IPD-MA) involve contacting authors of original studies for access to the participant level data, but also do not require ethical approval. The human subjects in an IPD-MA are the participants in the original studies, not the researchers being asked to share that data. The act of contacting researchers to ask them to share data, and reporting the rate of data availability, involves no human subjects and therefore does not require ethical approval for human subjects research. Furthermore, requiring ethical approval to request data from authors, or quantify the rate of availability, is not in line with any common set of ethical guidelines (e.g., American Psychological Association, 2016; Deutschen Gesellschaft für Psychologie, 2022; Oates et al., 2021). This manuscript does lightly summarise some of the types of reasons that people gave for not being able to share, but this is brief, anonymous, and descriptive, and is in line with the norms of our field: we would not typically require ethical approval to quote or paraphrase from another researcher's correspondence with us (e.g., a tweet, email, or letter).

Separately, it is useful to recognise that all but one of the journals represented in the dataset have explicit data-sharing policies that the authors agreed to prior to publication. That is, by publishing in these journals, the authors agree to the policy that they will share data upon request. This encouragement or requirement to share data wherever possible is echoed by professional

bodies (e.g., APA, BPS, DGPs) and funding bodies (USA NIH, EU Horizon). Lastly, it is also reflected in the Research Data Management policies of many universities, including those with which many researchers contacted as part of this study were affiliated.

Article search

I aimed to contact the authors of every IRAP publication published in the previous 5 years with a data-sharing request. In order to choose the articles that I would attempt to obtain the data for, I reused an existing published systematic search of the published IRAP literature (2006 to 2022, in English, listed in the Web of Science or PsycINFO databases). Full details of that systematic search, including Boolean search strings, all materials necessary to reproduce, reuse, or update the search, all data, and R code to reproduce the analyses are available in that publication (Hussey, 2023). Given that data has a half-life, insofar as it becomes increasingly hard to obtain over time, I considered only articles published within the last 5 years (i.e., those with a publication date of 2018 to 2022). I excluded articles that either I or Chad Drake was a co-author of, as I already had the data for these studies (i.e., they are included in Hussey & Drake, 2020). I found 52 such articles. The references for these publications can be found in the Supplementary Materials (osf.io/ympv6). The number of IRAP articles per journal can be found in Table 1.

Table 1. Number of IRAP articles by journal.

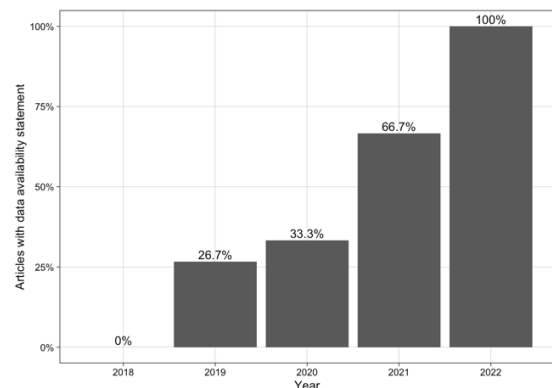
Journal	N articles
The Psychological Record	24
International Journal of Psychology & Psychological Therapy	8
Journal of Contextual Behavioral Science	5
Frontiers in Psychology	4
Behavior and Social Issues	2
Behavioural Processes	2
Dementia: The International Journal of Social Research and Practice	1
Emotional & Behavioural Difficulties	1
International Journal of Environmental Research and Public Health	1
Journal of Behavior Therapy and Experimental Psychiatry	1
Journal of Eating Disorders	1
Motivation and Emotion	1
Social Psychology of Education: An International Journal	1

Journal policies

Interestingly, three of the top four journals have data-sharing policies that require data sharing. The Psychological Record requires data sharing except in circumstances that must be justified at the time of submission: “A submission to the journal implies that materials described in the manuscript, including all relevant raw data, will be freely available to any researcher wishing to use them for non-commercial

purposes, without breaching participant confidentiality... All original research must include a Data Availability Statement.” (The Psychological Record, 2023). The International Journal of Psychology & Psychological Therapy does not have a data-sharing policy (International Journal of Psychology and Psychological Therapy, 2023). The Journal of Contextual Behavioral Science requires data sharing except in circumstances that must be justified at the time of submission: “It is expected that all authors who publish in the Journal of Contextual Behavioral Science will share data upon reasonable request. Therefore, we ask authors who do not already have their data openly available to the public to include an author note indicating ‘Data is available upon reasonable request’.” Authors can request to leave this note out if they can provide an adequately strong justification for not doing so in the cover letter.” (Journal of Contextual Behavioral Science, 2023). Finally, Frontiers requires data sharing except in circumstances that must be justified at the time of submission: “Frontiers requires that authors make the ‘minimal data set’ underlying the findings described and used to reach the conclusions of the manuscript, available to any qualified researchers.” (Frontiers, 2023).

Figure 1. Percent of articles reporting a Data Availability Statement by year.



Prevalence of Data Availability Statements

Of the 52 articles, 21 (42.3%) contained a data-sharing statement. Encouragingly, the proportion of articles increased from 0% in 2018 to 100% in 2022 (see Figure 1). It is worth noting that it is difficult to define a precise date when these policies came into effect. For example, from speaking to the editor of JCBS, these policies were progressively rolled out through the different levels of the journal’s article-handling processes over time. Regardless, results demonstrate that data-sharing statements have moved from absent to ubiquitous over these years.

Prevalence of data sharing upon request

I sent a data-sharing request to the authors of every article via email. A copy of the email can be found in the supplementary materials (osf.io/aetpq). In

summary, it stated that I wished to obtain the data from publications using the IRAP published in the last 5 years; that data would be screened for any personally identifying information and then posted to a project on the Open Science Framework; and that I hoped that authors could reply within two weeks to indicate whether they are able and willing to share the data.

In some cases, authors replied that they could not allow data to be made public, in which cases I replied that I was also willing to obtain the data and not make it public. I also noted that I was willing to sign any data-sharing agreements that authors felt were necessary. The strategy was therefore to request data to make it openly available in the first instance, and to request it be shared with me but not made public as a fallback option.

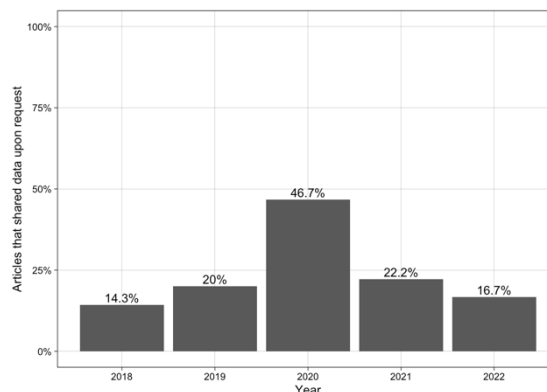
In the first instance, I attempted to contact the corresponding author using the email listed in the published article. If I did not receive a response offering to share the data, I then contacted the first author, if they were also listed as the corresponding author. If I again did not receive a response, I contacted other authors in order of apparent seniority (e.g., contacting professors before PhD students). Note that the IRAP literature does not, at the time of writing, conform to the norm present in some fields that the final author is considered to be the “senior” author, so seniority did not easily map onto authorship order. In many cases, a given author instructed me to speak to other co-authors to obtain the data. This was therefore a highly iterative process of sending email requests. At least two co-authors of every article were contacted. In total, 56 researchers were contacted.

In order to define a study endpoint, results were finalised 60 days after sending the first email to each author. This number was based on previous work by Tedersoo and colleagues (2021) who demonstrated that authors who shared data upon request tend to do so relatively quickly, and the probability of data being shared per day of waiting falls rapidly over time (50% within 15 days, 75% within 28 days, 90% within 34 days, 95% within 48 days, 97% within 59 days, and 99% within 90 days).

Aggregating results across all co-authors of each article, I received a reply to my email(s) in 82.7%, 95% CI [69.7, 91.8] of cases (43 articles). Authors reported being able and willing to share their data in 42.3%, 95% CI [28.7, 56.8] of cases (22 articles). Authors actually shared their data in 26.9%, 95% CI [15.6, 41.0] of cases (14 articles). When authors initially reported that they would share the data but did not actually do so, the nature of these interactions and the reasons given took a variety of different forms: some authors simply stopped replying to emails, some raised ethical or legal issues, and others reported that the data could only be shared via a university platform that did not yet exist. Impediments to data sharing are discussed in more detail later.

Unlike the prevalence of data availability statements, no linear trend in data sharing upon request was observed between articles published in 2018 to 2022 (see Figure 2). This was somewhat surprising, given that (a) more recent years had greater coverage of Data Availability Statements, and (b) previous work has found that data availability decreases over time, with a half-life of between 6 and 9 years (Minocher et al., 2021).

Figure 2. Percent of articles sharing data upon request by year of publication.



Relationship between promissory Data Availability Statements and actual data sharing

It is useful to consider data sharing in the subsets of different types of Data Availability Statements. I make a distinction between statements of actual sharing at the time of publication (e.g., a URL included in the article that links to a data repository containing the data for the study, or reference to supplementary materials published alongside the article) and promissory data sharing (e.g., a statement that data is available upon request, or upon ‘reasonable’ request).

Three articles’ data-sharing statements represented claims of actual data sharing at the time of publication. Of those, only two (66.7%) actually provided the data as described, in both cases via links to the Open Science Framework and ResearchGate (note however that I would recommend researchers to use services offered by not-for-profit entities with longer-term data preservation structures in place, such as the OSF, over for-profit entities such as ResearchGate). One article stated that “All data generated or analysed during this study are included in this article and its supplementary information files”. However, no such supplementary materials were available on the journal’s website. In this case, an author was able to provide the data upon request, however the inaccuracy of the data sharing statement is still notable.

Of the 31 articles without a Data Availability Statement, 8 shared data upon request (25.8%, 95% CI [11.9, 44.6]). Of the 18 articles with a promissory Data Availability Statement, 3 shared data upon request (16.7%, 95% CI [3.6, 41.4]). That is, data sharing was descriptively lower in articles with a promissory data

sharing statement than no statement at all. A Chi-squared test suggested that the presence of a Data Availability Statement was not associated with a higher rate of actual data sharing upon request, $\chi(1) = 0.35$, $p = .55$.

Impediments to data sharing

This section contains some slightly more qualitative reflections on the replies that I received and the insights they provide into impediments to data sharing. In the below, I sometimes refer to the number of authors or cases of a given situation. It is important to note that these do not perfectly map onto the number of articles, because (a) some researchers were co-authors of many articles and conversely (b) sometimes more than one author replied to my emails regarding a given publication, and sometimes authors even contradicted one another regarding the existence of data or its shareability. As such, these cases should be interpreted as examples of situations that can arise and what can be learned from those situations, rather than any attempt to estimate the prevalence of such situations.

It is often not possible to correspond with corresponding authors. Some corresponding authors were simply impossible to find working contact details for: the email addresses listed in the article did not work, and no up-to-date details could be found online, from contacting their collaborators or scouring social media. Worryingly, this included authors of articles published within the last calendar year (2022). This finding is in line with previous work that has demonstrated that the rate at which working contact details can be found decays over time: Vines et al. (2014) found that the odds of being able to find working contact details to request data from authors decrease by around 4% per year.

In their replies to the data-sharing request, multiple authors stated that they were on maternity leave (2 authors), were retired (1 author), or had left academia (3 authors). Three other first or last authors were not contacted in the first place because I knew they had left academia. All are reasonable and common circumstances, however, all situations highlight ways in which promissory data sharing is ineffective due to extremely foreseeable circumstances. Similarly foreseeable circumstances that can disrupt data sharing include researchers moving between institutions. Whereas academics typically put plans in place to cover their research and teaching duties while on leave, less attention may be paid to handing over responsibility for data availability.

Sixteen authors were initially responsive to my email and stated that I should instead contact a different author for access to the data. Almost universally, these authors could not provide working contact details for the authors they suggested contacting (i.e. when I could not find working contact details myself). Data ‘available upon request’ policies therefore leave us not only at the whim of authors’

willingness to share data but also their willingness and ability to keep and share even more basic details such as who has the data and how they can be contacted.

Reasonable steps should be made to be able to ensure that we can in fact correspond with a corresponding author. For example, the use of email addresses that are not tied to employment at a specific institution, and a deeper understanding of and commitment to the lasting responsibilities that come with being the corresponding author.

The ethics of data sharing should be considered holistically. Data sharing has ethical implications, but not sharing also has ethical and research integrity implications. Four authors stated that the data could not be shared on ethical grounds despite having included data sharing upon request statements in their publications. In each case, I followed up with a request that it be shared with me privately without subsequently being made public and that I was happy to sign any necessary data-sharing agreement. However, in three of these four cases, these initial ethical concerns were then made redundant when either replied the author replied that data was in fact lost, or the author stopped replying. This may represent a selective deployment of caution: much caution placed on the ethical requirement not to share data in certain ways, and not enough on research integrity, such as ensuring that results can be independently verified and uphold prior commitments to data sharing (e.g., to journals, funders, and professional bodies).

In four other cases, I was able to appeal to a Research Data Manager at the authors’ institutions and ask them to apply the university’s data sharing policy. In one of these cases, the author had been non-responsive to my emails. In a second, the author replied to my emails but was ambiguous about sharing the data, and then stopped replying. In the two others, the authors replied to my emails but directed me to the Research Data Manager to work out the legalities and ethics of how data sharing could be accomplished. It is worth noting that, in all four of these cases, the authors had stated that data was available upon request, and both the journal and university policy required sharing, but apparently no actual mechanisms for data sharing had been put in place to accomplish this until my request. At the time of writing, eight months after the first emails were sent, negotiations with the Research Data Manager regarding the specifics of the data-sharing agreements are still ongoing for three of these cases. The terms of these data-sharing agreements being debated set a remarkably high bar, such as the requirement the data-sharing agreement would need to be signed by the president of both of our respective universities (arguably not a scalable solution if data-sharing was to be as commonplace as funders and journals wish it to be), and the use of the data would be limited to reproducing exactly the same analyses reported in the original publications and no others (therefore no robustness tests could be examined, nor

could the data be reused for other purposes). The fourth case is apparently unresolvable as, despite university policy to (a) retain data and (b) share it upon request, the university has little power to enforce its own policies once researchers are no longer employed by that institution.

Unfortunate and sometimes untimely data losses occur. One researcher noted that they did have the data for multiple studies until very recently but they mistakenly wiped the hard drive of the laptop on which they were stored. Given the high concentration of authorships of IRAP papers by a small number of authors (Hussey, 2022) – a concentration of authorship that is likely to also be found in many small subfields – even a single data loss can involve the loss of data associated with large proportions of the literature, as was the case here.

Non-adherence to journal data-sharing policies, institutional data-sharing policies, and public statements about data sharing. In addition to journals' data-sharing policies, institutions, and funding bodies increasingly also assert their own Research Data Management policies regarding the retention, storage, and access to data by those seeking to verify results. In addition to non-adherence to journal Data Availability Statements, at least seven researchers were observed violating their institutions' research data management policies. This is a lower bound of the occurrence of this: no exhaustive search of institutions' data retention and sharing policies was conducted, and only the policies of institutes with a larger number of authors or IRAP publications were inspected. One author was also observed to contradict their own recent public position about the importance of data sharing: a co-author of the Association for Contextual Behavioral Science's recent Open Science recommendations report, which states "we recommend open data and transparency whenever possible." (Task Force on the Strategies and Tactics of Contextual Behavioral Science Research, 2021), declined to share any data.

Finally, there were instances of apparent inclusion of tokenistic or misleading Data Availability Statements. One author of several publications – who was listed as the contact person in the Data Availability Statement in one of them – when contacted stated that not only did they not currently have access to the data but that they had never been in possession of it. Data that we have never possessed cannot be shared, and should not be promised to readers upon request.

Discussion

Results demonstrated that the prevalence of Data Availability Statements in IRAP articles has risen from 0% in 2018 to 100% in 2022. This is encouraging and the journals should be applauded for embracing these policies and investing in the administrative infrastructure to implement them.

However, results also demonstrated that very few authors of recent IRAP publications share data on request (25.0% of 52 articles). Results were therefore very similar to those observed over 60 years ago by Wolins (1962: 24% of 37 articles). Hamilton et al.'s (2023) systematic review of studies of the rate of data sharing upon request observed rates between 0 and 37% (no meta-analysis was possible due to methodological heterogeneity). Worryingly, data sharing was no higher in articles that stated that data was available upon request (16.7%, 3 of 18 articles) than those that included no Data Availability Statement at all (25.8%, 8 of 31 articles). Disappointingly, results did not suggest that Data Availability Statements are associated with higher actual data sharing.

Limitations

Generalisability. The top-line conclusions of all previous studies covered in Hamilton et al.'s (2023) review, and indeed the results of this study, are in close agreement: data sharing upon request occurs in the minority of cases. However, precise estimates of the rate of data sharing are as yet unknown. The generalisability of the current estimate of data sharing upon request to other areas of psychology, or indeed other areas of science, is unknown at this time. The recent systematic review of the rate of data sharing upon request by Hamilton et al. (2023) noted that no meta-analysis of the rate of sharing across studies was possible due to methodological differences between the component studies, such as differences in journal policies between fields and differences in how old the publications were. The estimate provided by this study is perhaps more likely to correspond with (a) fields of similarly small size, (b) fields with comparable journal policies (e.g., requiring Data Availability Statements but not mandatory sharing at the time of publication), and (c) studies considering a similar timeframe (e.g., publications within the last 5 years).

The impact of personal acquaintance on sharing. The fact that I was personally acquainted with many of the researchers contacted may have influenced the rate of data sharing in some way. I have written critiques of IRAP research in the past, and it is possible this could have made some researchers more reticent to share. On the other hand, the fact that I was acquainted with many of the authors prior to contacting them for their data may have lent a degree of ecological validity to the request: within many small-to-medium-sized fields data requests are likely to come from another researcher who is already known to you. To the best of my knowledge, only one previous study has examined this question: Tedersoo et al. (2021) examined data sharing upon request and quantified whether they were acquainted with the authors prior to the data requests they sent (i.e., 'no contact', 'have met', 'have collaborated with'). An analysis of their open data suggests that data sharing upon request in their sample was slightly higher among researchers

already known to the authors (50% of 12 articles) than those who were not (40% of 272 studies).

The impact of the phrasing of a data sharing request on sharing. It is of course possible that the rate of data sharing upon request may be moderated by the nature of that request, e.g., the phrasing used in the email. Authors may perceive data sharing as a matter of interpersonal trust, and if they perceive the request to be overreaching, impudent, or otherwise distasteful they may be more likely to refuse, ignore or disengage from the request. To the best of my knowledge, no work to date has experimentally manipulated the phrasing of data sharing requests in order to assess its impact on the efficacy of such requests. Future research may wish to do this. However, at the same time, we should be wary of making compliance with data sharing requests the responsibility of the requester rather than the authors, where those requests are a matter of compliance with professional standards, journal policy, institutional policy, funding body policy, and/or the authors own written prior commitment to do so. Given that some authors are simply disinclined to share their data despite the requirements to do so, it may be more effective for the field to attempt to enforce these existing requirements rather than attempt to solve the issue by finding the optimal way of phrasing the request.

Data usability and computational reproducibility. Of course, not all shared data is useful, and data sharing is not an end in itself. Data sharing merely (a) enables verifiability of published analyses, and (b) enables reuse for novel purposes. Even when data is technically complete and shared, poor documentation (e.g., the absence of high-quality codebooks) can limit the degree to which it can be used for verification or novel purposes (Horstmann et al., 2020). Separately, recent research has demonstrated that even when data and code are openly shared at the time of publication, the results of relatively few publications can be precisely reproduced (e.g., only 1 of 14 articles: Crüwell et al., 2023). Additional research would be required to estimate the computational reproducibility of results reported in IRAP publications based on shared data.

Implications

While the overall rate of data sharing is disappointing, the non-adherence to journals' data sharing policies – which authors explicitly agree upon submission – is obviously unacceptable. If authors are shown to routinely disregard this specific journal policy (and in some cases also their institutional Research Data Management Policies and their own public positions on data sharing), this raises the question: has the research integrity of other as-yet unexamined elements of the research process also been undermined?

Unfortunately, perhaps these results are less surprising when viewed through the lens of the incentive structures in science. The contingencies that govern scientific research generally stop at the

publication of a given article. Publications typically function as reinforcers. Curating data and code to make it openly available, or even genuinely sharable upon request, has few reinforcers: it is more work for little reward. Nonetheless, there are now many resources which provide practical guidance to researchers on how to share data more easily (Gilmore et al., 2018; Meyer, 2018). Research elsewhere has considered other specific elements of the research process that make data sharing easier, such as the content of consent forms in light of the EU's GDPR data legislation (Hallinan et al., 2023), how-to guides on using data sharing platforms such as the Open Science Framework (Soderberg, 2018), and tools to easily create data codebooks that allow others to interpret and use shared data (Horstmann et al., 2020).

Equally, there are currently few punishers for failing to adhere to Data Availability Statements. As employers, institutions have more scope to enforce Research Data Management policies among their employees as a matter of research integrity. Having spoken to them about the unfulfilled data requests described here, many institutions' Research Data Management and Research Integrity offices seem to have a growing interest in defining and enforcing such policies. However, as yet, journals have asserted relatively fewer demands on authors. Some journals go further than requiring Data Availability Statements and actually require data and code to be shared as a condition of publication. A small number of journals even check the computational reproducibility of results prior to publication (e.g., Meta-Psychology). No journal as yet has established any punishment mechanisms for breaches of data sharing agreements, such as a policy of rejecting future submissions to the journal if they receive and verify a report of a refusal to abide by the data sharing agreement in a previous publication. These policies and others would all likely be extremely effective in increasing data sharing, however, they also require yet more investment from already under-resourced journal staff, most of whom are volunteers. No solution to these problems is trivial, but in my opinion, the current state of affairs is untenable. As stated in previous similar articles, I conclude that "statements of data availability upon (reasonable) request are inefficient and should not be allowed by journals" (Tedersoo et al., 2021). The presence of Data Availability Statements that are not adhered to or enforced in any way risks giving rise to what is referred to as 'Open Washing': the appearance of transparency without adequate follow-through (Villum, 2014).

Hardwicke et al. (2018) studied the impact of a change in policy at the journal *Cognition* from encouraging data sharing to mandating open data sharing at the time of publication on the actual rate of data sharing. In one sense, results were encouraging: sharing of open data alongside the publication rose from 22% (of 104 articles) to 62% (of 136 articles). In another sense, it is somewhat dismaying that even a journal

requirement to make data available prior to publication – a policy that is in principle enforceable by the journal, which could withhold publication until the requirement is met – did not produce universal data sharing.

Of course, data sharing is not a panacea or an end itself, but merely one step towards increasing the reproducibility of findings and unlocking data reuse potential. Recent research has demonstrated that even when articles share their data, the results reported in the articles can only be precisely reproduced in a small minority of cases. If future research is to become increasingly reproducible, it will have to become not only increasingly verifiable through transparency but also increasingly verified through actual checks, at least in a proportion of cases. It may also be the case that even a small but non-zero expectation that others may actually ask for our data or check our reported results for their accuracy may increase their reproducibility. If so, although studies such as the current one do not show support for the efficacy of low-intensity data-sharing interventions such as Data Availability Statements, it is possible that the act of conducting research such as this also acts as an intervention. Historically, the probability of being asked to share one's data is quite low. Studies such as the current one, or indeed the potential for future journal-wide audits of data sharing held on a regular basis (e.g., Lear et al., 2023), may serve not only to test but also to establish normative expectations of data sharing.

Author notes

Correspondence: Ian Hussey, University of Bern, Switzerland (ian.hussey@icloud.com). ORCID [0000-0001-8906-7559](https://orcid.org/0000-0001-8906-7559).

Conflict of Interest and Funding: This research was supported by the META-REP Priority Program of the German Research Foundation (#464488178).

Author Contributions: IH was responsible for all aspects of this project.

References

- American Psychological Association. (2016). *Ethical principles of psychologists and code of conduct*. APA. <https://www.apa.org/ethics/code>
- Crüwell, S., Apthorp, D., Baker, B. J., Colling, L., Elson, M., Geiger, S. J., Lobentanzer, S., Monéger, J., Patterson, A., Schwarzkopf, D. S., Zaneva, M., & Brown, N. J. L. (2023). What's in a Badge? A Computational Reproducibility Investigation of the Open Data Badge Policy in One Issue of Psychological Science. *Psychological Science*, 34(4), 512–522. <https://doi.org/10.1177/09567976221140828>
- Deutschen Gesellschaft für Psychologie. (2022). *Berufsethische Richtlinien*. <https://www.dgps.de/die-dgps/aufgaben-und-ziele/berufsethische-richtlinien/>
- European Commission. (2023). *The EU's open science policy*. https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science_en
- Evans, T. R. (2022). *Developments in Open Data Norms* (1). 10(1), Article 1. <https://doi.org/10.5334/jopd.60>
- Frontiers. (2023). *Policies and publication ethics*. <https://www.frontiersin.org/guidelines/policies-and-publication-ethics>
- Gilmore, R. O., Kennedy, J. L., & Adolph, K. E. (2018). Practical Solutions for Sharing Data and Materials From Psychological Research. *Advances in Methods and Practices in Psychological Science*, 1(1), 121–130. <https://doi.org/10.1177/2515245917746500>
- Google Trends. (2023). *Comparison of the relative frequency of usage of “data is” vs. “data are.”* <https://trends.google.com/trends/explore?date=today%205-y&q=data%20is,data%20are&hl=en-GB>
- Hallinan, D., Boehm, F., Külpmann, A., & Elson, M. (2023). Information Provision for Informed Consent Procedures in Psychological Research Under the General Data Protection Regulation: A Practical Guide. *Advances in Methods and Practices in Psychological Science*, 6(1), 25152459231151944. <https://doi.org/10.1177/25152459231151944>
- Hamilton, D. G., Hong, K., Fraser, H., Rowhani-Farid, A., Fidler, F., & Page, M. J. (2023). Prevalence and predictors of data and code sharing in the medical and health sciences: Systematic review with meta-analysis of individual participant data. *BMJ*, 382, e075767. <https://doi.org/10.1136/bmj-2023-075767>
- Hardwicke, T. E., Mathur, M. B., MacDonald, K., Nilsson, G., Banks, G. C., Kidwell, M. C., Hofelich Mohr, A., Clayton, E., Yoon, E. J., Henry Tessler, M., 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>
- Horstmann, K. T., Arslan, R. C., & Greiff, S. (2020). Generating Codebooks to Ensure the Independent Use of Research Data. *European Journal of Psychological Assessment*, 36(5), 721–729. <https://doi.org/10/ghmt9r>
- Hussey, I. (2022). Reply to Barnes-Holmes & Harte (2022) “The IRAP as a Measure of Implicit Cognition: A Case of Frankenstein’s Monster.” *PsyArXiv*. <https://doi.org/10.31234/osf.io/qmg6s>
- Hussey, I. (2023). A systematic review of null hypothesis significance testing, sample sizes, and statistical power in research using the Implicit Relational Assessment Procedure. *Journal of Contextual Behavioral Science*, 29, 86–97. <https://doi.org/10.1016/j.jcbs.2023.06.008>
- Hussey, I., & Drake, C. E. (2020). The Implicit Relational Assessment Procedure demonstrates

- poor internal consistency and test-retest reliability: A meta-analysis. *PsyArXiv*.
<https://doi.org/10.31234/osf.io/ge3k7>
- International Journal of Psychology and Psychological Therapy. (2023). *Authors Guidelines*.
<https://www.ijpsy.com/normas.html>
- Journal of Contextual Behavioral Science. (2023). *Guide for Authors*.
<https://www.ijpsy.com/normas.html>
- Lear, M. K., Spata, A., Tittler, M., Fishbein, J. N., Arch, J. J., & Luoma, J. B. (2023). Transparency and reproducibility in the journal of contextual behavioral science: An audit study. *Journal of Contextual Behavioral Science*, 28, 207–214.
<https://doi.org/10.1016/j.jcbs.2023.03.017>
- Meyer, M. N. (2018). Practical Tips for Ethical Data Sharing. *Advances in Methods and Practices in Psychological Science*, 1(1), 131–144.
<https://doi.org/10.1177/2515245917747656>
- Minocher, R., Atmaca, S., Bavero, C., McElreath, R., & Beheim, B. (2021). Estimating the reproducibility of social learning research published between 1955 and 2018. *Royal Society Open Science*, 8(9), 210450.
<https://doi.org/10.1098/rsos.210450>
- Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., 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>
- Nunes, L. (2021). Data Sharing for Greater Scientific Transparency. *APS Observer*, 34.
<https://www.psychologicalscience.org/observer/data-sharing-methods>
- Oates, J., Carpenter, D., Fisher, M., Goodson, S., Hannah, B., Kwiatkowski, R., Prutton, K., Reeves, D., & Wainwright, T. (2021). *BPS Code of Human Research Ethics* (p. bpsrep.2021.inf180). British Psychological Society.
<https://doi.org/10.53841/bpsrep.2021.inf180>
- Rogers, S. (2012, July 8). Data are or data is? *The Guardian*.
<https://www.theguardian.com/news/datablog/2010/jul/16/data-plural-singular>
- Soderberg, C. K. (2018). Using OSF to share data: A step-by-step guide. *Advances in Methods and Practices in Psychological Science*, 1, 115–120.
<https://doi.org/10.1177/2515245918757689>
- Task Force on the Strategies and Tactics of Contextual Behavioral Science Research. (2021). *Adoption of Open Science Recommendations | Association for Contextual Behavioral Science*.
https://contextualscience.org/news/adoption_of_open_science_recommendations
- Tedersoo, L., Küngas, R., Oras, E., Köster, K., Eenmaa, H., Leijen, Ä., Pedaste, M., Raju, M., Astapova, A., Lukner, H., Kogermann, K., & Sepp, T. (2021). Data sharing practices and data availability upon request differ across scientific disciplines. *Scientific Data*, 8(1), Article 1.
<https://doi.org/10.1038/s41597-021-00981-0>
- The Psychological Record. (2023). *Instructions for Authors*. Springer.
<https://www.springer.com/journal/40732/submission-guidelines>
- Villum, C. (2014). “Open-washing” – The difference between opening your data and simply making them available – Open Knowledge Foundation blog. <https://blog.okfn.org/2014/03/10/open-washing-the-difference-between-opening-your-data-and-simply-making-them-available/>
- 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>
- Wicherts, J. M., Borsboom, D., Kats, J., & Molenaar, D. (2006). The poor availability of psychological research data for reanalysis. *American Psychologist*, 61(7), 726–728.
<https://doi.org/10.1037/0003-066X.61.7.726>
- Wolins, L. (1962). Responsibility for Raw Data. *American Psychologist*, 17(9), 657–658.
<https://doi.org/10.1037/h0038819>