Date: 2023-03-10

Subject: Request for data from your recent IRAP publications

Dear Prof [name],

Apologies for the slightly impersonal template email below, I have quite a few of these emails to send!

I'm writing to you to ask you to share the data from your recent IRAP publications (from 2018 to present).

Results of my systematic review of IRAP publications show that you are the [corresponding/senior] author of the following IRAP manuscripts published since 2018. If there are any missing, please let me know.

* [publications]

**Background**

Over the past few years, I have collated the data from my and Chad Drake's IRAP studies into a common and reusable FAIR format (<https://www.apa.org/pubs/journals/data-sharing-infographic.pdf>). Data from thousands of participants in dozens of domains is now openly available (<https://osf.io/v3twe/>) in multiple different formats, including trial level data, scored data (D, PI), and meta analytic estimates (e.g., Cronbach's alpha).  
  
This allows for anyone to conduct new analyses of this data, either for domain-level purposes (e.g., additional research questions not addressed within the original publication), meta-analytic purposes (e.g., my meta-analysis of reliability, currently under review at JCBS: <https://psyarxiv.com/ge3k7>), or for meta-scientific purposes (e.g., assessing the computational reproducibility of published analyses, as has been done elsewhere: <https://journals.sagepub.com/doi/full/10.1177/09567976221140828>).  
  
Since I began this project with my own data around 2017, there has been an encouraging shift by professional organisations and funding bodies towards encouraging or requiring data sharing of this kind (see further below). In summary, please note that two prolific IRAP researchers, Dermot Barnes-Holmes and Louise McHugh, recently co-authored and unanimously approved a report for ACBS which advocates for data sharing.  
  
**What data I'm asking for**

I'd therefore like to do the same data collation and sharing for all IRAP papers published within the last 5 years (i.e., from 2018). I'm writing to the senior authors and PIs of all IRAP publications from 2018 (as found in this systematic review: <https://osf.io/tna8s/>). I selected 2018 as it roughly coincides with the development and distribution of the GO-IRAP, which produces data files that that contain all three of (1) the procedural setup for that IRAP (stimuli, task parameters, etc), (2) basic demographic information, and (3) both trial level and scored data. This makes my life curating the collated data much easier.  
  
I deeply appreciate that we are currently not strongly incentivised to spend much time thinking about data management, curation, sharing etc. As much as possible, I have therefore tried to tailor this request to minimise any administrative burden on you. It should only take a few minutes, given that you can simply send me the files/folders that you have and I will do the vast majority of the work here.  
  
I'm asking for both of the following for each publication:  
  
1. The dataset used for the analyses reported in the published article.

    - This can be in any format you like, SPSS, excel, csv, etc. If you have it hosted online already, e.g., on OSF or GitHub, you can simply send me a link.

    - If you have scripts or syntax that implement the analyses in the manuscript, they would be very welcome too (e.g., R code, SPSS syntax, etc), however this is a secondary concern.

2. The raw data, trial level data files.

    - For example, the .csv file(s) produced by the GO-IRAP.

    - If you have item level data available for other tasks (e.g., self reports), you're welcome to share it too, however this is a secondary concern.  
  
**What will be done with the data**

It is quite unlikely that your datasets contain what GDPR refers to as "personal data", such as full names and addresses. It is unproblematic to share (between research groups, such as with me; or publicly) the irrevocably anonymised and non-personal data contained in the studies I have listed above (e.g., age, gender, and responses on the IRAP). This mitigates any concerns about sharing for reasons of privacy or GDPR compliance.  
  
That being said, my first step will be to screen the original data files that you send me for any identifying or personal data and permanently delete it. The resulting set of files I will refer to as the screened data (although note that these materials may also contain your scripts and code, not just data).  
  
The screened data will then be organised and posted publicly on the Open Science Framework and GitHub under a Creative Commons CC-0 licence (<https://creativecommons.org/share-your-work/public-domain/cc0/>). This data will then be openly available for others to use.  
  
I will then curate the screened data into the processed data formats already established for my own data (see <https://osf.io/v3twe/>). Screening the data should be quite quick. Curating it could take some time.  
  
**Timeline**

I'm asking that you respond within two weeks (by March 26th) to indicate your willingness to share the data. If you're not able or willing to share it, it is still useful for me to know that this is the case: please drop me a quick reply stating so. It would also be helpful if you could give me your reason(s) for not sharing.    
  
We can then discuss a reasonable timeline for you to share the data with me. Again, I'd like to reiterate that I will do the majority of the work screening, curating and publicly sharing the data. All that is required of you is to locate the files and send them to me. I'm therefore hopeful that four weeks from now (April 9th) will be enough time to send the data.

If you don't have access to a given dataset but, for example, another author of that publication is likely to have a copy, I would really appreciate your help liaising with your co-authors to obtain it and forward it to me.  
  
**Data sharing recommendations and requirements of various professional organisations and funding bodies**

* In 2021, ACBS's Task Force on the Strategy and Tactics of Contextual Behavioral Science Research produced its Open Science recommendations, which state that "We recommend open data and transparency whenever possible" (<https://contextualscience.org/news/adoption_of_open_science_recommendations>). The Open Science committee included Louise McHugh and Dermot Barnes-Holmes, and it unanimously approved these recommendations. ACBS's journal, JCBS, now states requires all authors to agree to share data upon reasonable request: "It is expected that all authors who publish in the Journal of Contextual Behavioral Science will share data upon reasonable request" (<https://www.elsevier.com/journals/journal-of-contextual-behavioral-science/2212-1447/guide-for-authors>).
* APA's data sharing principles (<https://www.apa.org/science/leadership/bsa/data-sharing-report>) state that "Sharing data promotes scientific progress" and "Sharing data within the larger scientific community encourages a culture of openness and accountability in scientific research".
* The EU's Horizon Europe funding program now "requires all authors to share the underlying data which relates to their article" <https://open-research-europe.ec.europa.eu/for-authors/data-guidelines>

If you have any questions, please let me know.

best wishes

Ian