

Statistics and Methodology

#PSBigData: Big Theory

Author Rick Dale
Other Posts By Rick Dale
July 18, 2019 | No Comments

This commentary is in two parts. As a contributor to the special issue, I enjoyed reading these commentaries, and felt compelled to synthesize that enjoyment. That's the first part.

The ideas sparked by these commentaries lead to the second part. There I consider a potential next step in progress in our field, and link it to big data. That next step is posed as a question: *Is there a theory crisis in our field?* I review some recent discussion and papers suggesting the answer is "yes."

Part 1: Themes in the commentaries

There appears to be consensus in the commentaries: Large and naturally occurring data sets are great, for a host of reasons. This is refreshing. In recent years, I've spoken to many colleagues who report once facing skeptical questions about the noisiness or awkwardness of these data. It can take some convincing that such data are in some ways more "natural" than their laboratory counterparts.

For example, we analyzed [Yelp](#) review data in [our own article](#) in the [Lupyan and Goldstone special issue](#) (led by Dave Vinson and Mike Jones). Writing a *Yelp* review is quite "natural" and "ecological". Many people often sit at computers and type up heartfelt responses to their experiences. It is what I'm doing right now. There are [issues with big data](#), and I discuss some below.

But there are plenty of potential issues with lab data too.

So I share the enthusiasm for *big data* or *naturally occurring data sets* (I like the acronym BONDS, see [Alexandra Paxton's commentary](#), and will use it sometimes here). The wonderful discussion in these commentaries identifies several interesting themes in the special issue and beyond. I highlight three.

Improved power and breadth in data sets. [Gureckis and Griffiths](#) note that the special issue involves data from almost 2 million individuals, not even including text analyses of books and so on. That's awesome. But there's more. As [Mullett notes](#), it is not just an increase in sample size at the participant level, but also within-participant sampling. [Thorstad and Wolff](#) have data from single participants that span years. [Monaghan](#) summarizes how [Li and colleagues](#) bridge cultural and psychological levels of analysis. [WEIRD concerns](#) may surely persist, but these data go well beyond a narrow selection of college students getting course credit in the lab.

Testbeds for expanding new methods and statistics. As [Gray observes](#), BONDS permit us to revisit and reanalyze data to expand new and sophisticated methods, as illustrated in NBA data by [Vaci and colleagues](#) in the issue, and in [Sangster's work](#). In a similar vein, [Molly Lewis](#) describes how [Frey and folks](#) use a large social gaming dataset to explore cause-effect relations in communication. I was compelled by her points that using big data sets for causal analysis is an exciting future direction, considering the importance of causal reasoning in science and in minds themselves.

Open and ethical science practices. [Tim Mullett comments](#) on the assistance that big data can bring to such issues as replication and so on. I agree, and his remarks imply that it is not just by increasing power, but also by increasing complexity ("heterogeneity," as termed in [Lupyan and Goldstone's introduction](#)). Complexity in big data is not an irritating bug, but an irritating feature.

The array of potential proxy variables in BONDS can be leveraged to render insights difficult to obtain in the lab. So complex large-scale data sets may help us to understand what moderates small effects. The relationship between the lab and BONDS could be an exciting synergy.

Big data also raise new ethical concerns. [Alexandra Paxton](#) discusses the paper by [Dennis and colleagues](#) in the special issue. She comments on how to be maximally open in our science, while also preserving participant privacy. Her comments are compelling, and imply that academics could develop better models for data ethics, privacy and ownership than prominent [recent examples outside academia](#).

Interlude: Concerns with big data

Big data and open science (especially replication) are two big themes that have greatly influenced psychological science lately. There is much exciting promise on the horizon by combining lab work with natural data, and by improving our research practices *together*.

For example, BONDS have innumerable degrees of freedom for effect seeking, risking [HARKing](#) in particular. So adopting recommendations set by the important [open science movement](#) can help—such as preregistration of analyses on BONDS, and transparency in exploratory vs. confirmatory designs.

These themes, even the sometimes spicy discussion online, have been very helpful. For example, my own lab is now aspiring to the bold edict that data should be "[born open](#)." But even with open data and reproducible results, another concern is that BONDS can yield innumerable *subtle statistically significant effects*. But which effects matter? Which effects should be sought in preregistration? Which explorations should be conducted, and which potential findings are beckoning for confirmation in our large data sets?

Answers to these questions may be clear in specific domains and situations, and the special issue contains wonderful examples. But in more general terms, some recent discussion suggests that systematic answers to these questions—how we come to answer them in a more general sense—is an important next step. Some researchers have recently argued that improving research resources and practices alone may not yield the scientific progress we seek, and not even when coupled with big data. Methodological improvement, according to these researchers, should also accompany a very strong dose of *improvement in theory*.

Part 2: Theory crisis?

The concern may be best expressed by the intriguing title of this 2013 manuscript by [Hasselman and colleagues](#): "So you confirmed, replicated and emptied your file drawer—now what?" They argue that even if we achieved the various methodological aims discussed here, expanding resources and improving methods, we'd still be stuck. They observe that in a [special issue in Perspectives in Psychological Science](#) on the replicability crisis, a relatively small percentage of papers mentions theory at all. Their paper offers an elaborate historical and philosophical discussion of the role of theory, ultimately endorsing a kind of "structural realism," a big-picture account of what theory is meant to accomplish, such as orienting replication efforts, why specific replications are important, and so on.

[Iris van Rooij](#) also offered an account of the importance of theory in this very forum, earlier this year. She kicks off her own discussion with a quote from [Cummins](#):

"(...) a substantial proportion of research effort in experimental psychology isn't expended directly in the explanation business; it is expended in the business of discovering and confirming effects."

Her commentary converges on a lament for the lack of tools for theory development in our field. A stark observation she shares is that standard psychology curricula require students to learn rigorous experimental and statistical methods, but rarely require training on the design of theories, especially formal ones.

Two recent papers further highlight this concern. In a perspective paper in *Nature Human Behaviour*, [Muthukrishna and Henrich](#) argue that broader theoretical frameworks, especially formal ones, are needed to render clearer prediction, and better understand what makes a result surprising or not. Again a theme here is that replication issues can be facilitated by a greater focus on *cumulative progress* guided by theory.

A second illustration is in [Oberauer and Lewandowsky](#), their paper just accepted in *Psychonomic Bulletin & Review* (in press) entitled "Addressing the Theory Crisis in Psychology." Their discussion overlaps in some important ways with the discussion mentioned above. They elaborate on the theory crisis in a tour de force discussion that is compelling and practices what it preaches—quantifying their own concerns about theory in a Bayesian framework. Their discussion is guided by contrasting two overarching research styles: discovery-oriented research, and theory-testing research.

The distinction can be illustrated, as they do, with embodied priming research. The very notion of embodied priming might manifest itself in a wide variety of forms—the elements of daily experience that can be simulated in the lab is rife with *possible*, but *unrealized* embodied and conceptual associations (sense modalities, spatial orientation, etc.). Discovery-oriented research searches for instances of these priming effects, without necessarily knowing *why* particular contexts might show this. If an effect is not found, it does not challenge the overarching theory too sorely—we just keep searching for *where* the priming holds.

Theory-testing research develops stronger links between the elements in a theory and the behaviors we measure—it motivates experimental or other tests in *very specific circumstances*. It is a stronger and more direct test of the theory; if the result does not hold, it more directly challenges the theory that predicted it.

Oberauer and Lewandowsky develop an elaborate kind of Bayesian "meta model" for the evidential structure of these two research strategies. The paper shows how strengthening a theory-testing approach may be the best way to overcome the issues in replication and so on, because robust theory testing acts as an important constraint about what we expect from our studies, how surprising are their outcomes, and so on. Oberauer and Lewandowsky agree on the importance of recommendations made in the open science approach. But they also argue that the impacts of these recommendations will be limited without strong theories to guide research designs and interpreting their outcomes.

Consider preregistration. While noting general merits of preregistration, Oberauer and Lewandowsky also remark on significant evidential limitations of preregistration when not accompanied by strong theory:

"These priors should depend on how strongly each hypothesis follows from a theory, and not on how many hypotheses a researcher plans to test in the same data set. The role of preregistration in the Bayesian approach is to make researchers think about their priors without being biased by the data, but the act of preregistering a hypothesis does not increase its prior, and therefore has no impact on its posterior."

Ultimately Oberauer and Lewandowsky argue for stronger formalization, especially through computational modeling. Formal or computational models can accomplish the specificity and explicitness in the theory-testing approach. But this introduces a host of other issues relevant to modeling, such as numerous free parameters in our models, or seeking computational precision in some fields or phenomena that may pose special challenges. These issues are addressed in an extended conclusion, in which they offer some ideas about how one could formalize embodied priming research, and make it more aligned with explicit theory-testing.

Conclusion: Big data and theory

What does this potential theory crisis mean for the special issue and big data? Similar lessons can be drawn. As noted above, BONDS afford many degrees of freedom, especially in the number of possible comparisons or correlations or coefficients that can be computed. Improving theory can help narrow which of these is most interesting. BONDS can produce large numbers of significant such tests—an ocean of significant regression coefficients or pairwise correlations and so on. But theory could help make it clearer when weak significant effects may in fact be much more interesting and theoretically impactful than gigantic significant effects, which may be more intuitive or just "theory-agnostic."

None of this is to deny that each paper in the special issue brings its critical granules of theory. The papers are often rather directly motivated by big-picture ideas about human culture, human sociality and decision making, mental health, and more. The lesson that I myself have drawn though, as someone fascinated by BONDS and their promise, is to try to focus as much on the development of theory alongside the many other exciting methodological improvements we're seeking in our field.

I'm convinced by the work reviewed in the prior section that we could chew a bit more on the theoretical/philosophical aspects of our work. This might help render that theoretical explicitness, even if it eventually will come from formal or computational treatment. It might even help us bridge the many divides among our many, many models and competing frameworks, a fractionation that has not gone unnoticed by many.

For example, to draw from [Cummins](#) again, who tells us not to worry too much, as long as we all dive into the modeling rabbit hole:

"The ordinary practice of good science will take care of disunity eventually. There is a far greater danger in forcing more unity than the data warrants. Good experimentation, like good decision making generally, can tell us which of two models is better, but it cannot tell us how good any particular model is. The best strategy, then, is to have a lot of models on offer on the grounds that, other things equal, the best of a large set is likely better than the best of a small one."

Author

[Rick Dale](#)



The Psychonomic Society (Society) is providing information in the Featured Content section of its website as a benefit and service in furtherance of the Society's nonprofit and tax-exempt status. The Society does not exert editorial control over such materials, and any opinions expressed in the Featured Content articles are solely those of the individual authors and do not necessarily reflect the opinions or policies of the Society. The Society does not guarantee the accuracy of the content contained in the Featured Content portion of the website and specifically disclaims any and all liability for any claims or damages that result from reliance on such content by third parties.

Previous Post	Next Post
-------------------------------	---------------------------

You may also like

- July 12, 2019

[The cognitive psychologists are coming! #psynom22](#)
- November 8, 2022

[#BayesianPsych: Preventing miscarriages of justice and statistical inference](#)
- February 18, 2018

[#PSBigData: The Power of Accidental Data: Replicating lab studies without experiments](#)



Follow Us



Subscribe to Blog via Email

Enter your email address to subscribe to this blog and receive notifications of new posts by email.

EMAIL ADDRESS

Categories

Select Category

Archives

- ▼ 2022
 - November
 - October
 - September
 - August
 - July
 - June
 - May
 - April
 - March
 - February
 - January
- 2021
- 2020
- 2019
- 2018
- 2017
- 2016
- 2015
- 2014

Latest Twitter Feed

Psychonomic Soci...
@Psychonom... · 39m

Replying to @Psychonomic_Soc

Current state of the project: testing and expanding the experimental toolkit/documentation--if you're interested in providing feedback, reach out.

Scan the QR below to get updates!