

**Open access as a mechanism of revolutionary science and the limitations of data**

Costa Vakalopoulos

hinemoa@bigpond.net.au

**Richmond Hill Medical Centre**

# **Abstract**

Data is the cornerstone of the modern academic industry, like a constant production line of consumable goods packaged with a veneer of statistical techniques. Biomedical and psychological sciences invert the traditional logic of the physical sciences where hypotheses were tested against data rather than data against hypotheses. This is likely to reflect the immaturity of the new paradigms that barely cope with the data output of a precocious enterprise. However, the current state leads to distortions of the practice of science and entrenchment of its dysfunctional politics reflected in the science itself. With the debate now on the reproduction of data and statistical sleights of hand accompanying many if not most studies, what is lost in the debate is the preeminent role of good theoretical discovery. This is partly as a result of the structure of funding, the nature of reporting in biomedical fields based on the symbiotic relationship of high ranking institutions and journals. Open access could fill a gap in traditional publishing literature which has entrenched a culture of highly restrictive practices at a time when revolutionary science is required. OA is not in danger of lowering the standards of science as its critics claim as the studies on non replication show that they do not discriminate against OA papers specifically. Because it is ‘open’ to new and radical ideas that would never see the light of day otherwise, it may well provide a rejuvenating energy. OA is a self-regulatory response to the consequences of the distorted and inhibitory contemporary practice of science as the popularity of OA journals testify to.

# **Introduction**

In her essay on the crisis of replicability Spellman (2015) notes a revolution of sorts: a political one, but not scientific. Her failure to recognize it as part of the current process for a

scientific revolution in the psychological sciences reflects a broader lack of concern for the nature of the data generating enterprise and only spells out the crisis in terms of quality assurance. There is however, something intrinsically wrong with the entire project of scientific research as currently practiced and its narrow pursuit of data collection. This data driven concern is paralleled by traditional journal policies where the ongoing impetus is to publish more and better data as a means to scientific enlightenment. This approach fails to take into account the historical means by which science has made progress in large leaps based on innovative theory, not the gradual accumulation of data. The recent developments in open access have offered a new fillip to this re-emerging method of scientific practice that has been subordinated to data generation in the life sciences.

Open access is more than a debate about models of access to the research literature and the burden of cost. It is driven by a broader agenda that challenges the control of ideas in science. This is demonstrated by the focus of marketing by OA journals on ‘fair’ and open peer review that encourages a more constructive interchange. Traditional anonymous peer review can vary from pampering an author to outright hostility dependent strictly on the provenance of a manuscript.

In the neurosciences it is as much about presenting new paradigms that wouldn't see the light of day if traditional means of publication stood unchallenged. Thus, it is a necessary part of a political shift in scientific control, a step for progress in science which has become rather stale, repetitive and dominated by a restricted set of agendas. These agendas have barely offered solutions, in spite of the prestige of provenance and venues of expression, to a number of outstanding questions in science. The most obvious is the mind-body problem and its significance for the single most outstanding issue in the psychological sciences. Thus,

refining clinical descriptions in DSM V remains observational, but is not anchored to the wealth of behavioural, pathophysiological and genetic data. Indeed, it represents a disjointed science with the unerring, but misguided belief that all we need is more data.

A brief survey of the practice of science is encapsulated by the patent wars where new technologies such as DNA editing are promising all too-frequently that they will ultimately revolutionize treatments for autism, schizophrenia and major depression. However, these conditions are poorly understood both from a clinical perspective and from a genetics point of view. So the technology has outstripped good theoretical models that might make the techniques useful. The dependence if not, addiction in the biomedical and psychological sciences to data-driven studies has marginalized theoretical models based on a broad scientific understanding as a way forward, even as these have been revered in the ‘first’ scientific disciplines like physics, which could not function without them.

### **The status of theory**

Many of the psychological and biological sciences have coopted the apparent methodologies of the physical sciences with variable degrees of success in spite of a profound immaturity in its disciplines. There has been an increasing trend for data in publications over the last 3 decades and this has been at the expense of theoretical discussion (Vale 2015). An oft-quoted example is the paper published in *Nature* by Watson and Crick which presented a speculative structural model of the DNA helix without data. The currently low status of theory is implied by *Nature* guidelines for authors under the section ‘Other material published in *Nature*’:

### *10. Hypothesis*

These articles are published rarely, only about once a year. They concern issues of immense and fundamental importance. They are peer-reviewed. Enquiries should be sent through our online submission system with ‘Hypothesis:’ inserted before the title. If the enquiry is not answered within 2 weeks, authors should assume that *Nature* is not interested...

The policy of delegating theory to a fringe activity except as a lesser adjunct to data may have a number of root causes. Revolutionary ideas are just too rare to worry about in the daily life of a journal or else, sociopolitical motives are at hand and these are not necessarily mutually exclusive. However, the emphasis on data can be restrictive because they assume a normative rather than revolutionary course for science. The state of play belies the true nature of the science in which data and statistical validation have hollow or at least very narrow theoretical and at best, incongruent frameworks on which to work.

Of course, the proposal for a double helix presented by Crick and Watson was not a dataless hypothesis. They had at their disposal crystallographic photographs of Rosalind Franklin. Pauling used Astbury’s blurry photographs and according to one point of view was a reason in his failure to determine the helical structure of DNA before Crick and Watson (<https://paulingblog.wordpress.com/2009/07/09/the-x-ray-crystallography-that-propelled-the-race-for-dna-astburys-pictures-vs-franklins-photo-51/>).

It is clear from this example that the strength and generalizability of a theory is dependent on the quality of empirical data it serves to explain and the tools available to measure these. The two scientists’ epiphany, though relates to an understanding of the significance of the famous photograph 51, not self-evident in its clarity. In other words, the language the image portrays already exists in the model used to interpret it rather than independent of it and is a measure of success of the theory itself. What this further suggests, however is that a theory does not

have to be completely formative or predictive of what it might find through observation, but ought to be successful at explaining and adapting to data that might prove informative... and there are empirical observations that just cannot be guessed at no matter how powerful the theory might be. This might appear like a concession, but the relationship between data and theory can be a messy.

One might expect that for data to support good theory in the life sciences, it could well follow the example of the physical sciences where there are often time lags of decades. Good data doesn't automatically suggest breakthrough ideas, but ideas have a semi-autonomous course in their own right. That is there remains a 'magical' dimension to good empirical theory that transcends the data. In other words, there is a certain minimum number of empirical observations from which a model is necessarily drawn, but there comes a point where data becomes improbable for a foundation of science without a good model.

Common to the experience of *Nature* most journals do not offer a dedicated or severely restrict hypothesis manuscripts. The best avenues they offer are review type articles and many journals will not accept unsolicited papers. Peer review, as traditionally practiced and by the very nature of a science as an incremental process, generally discourages innovative ideas and the right to present theoretical articles generally depends on a track record for data-driven prior publications. That is expertise in the field is judged by experimental publishing record, even though history in the physical sciences dictates that theory and data can be generated by separate sources. Indeed, as is argued elsewhere those that produce data are not often the best placed to construct good theory. Although traditional subscription journals can publish theory, these are often 'allowed' to high profile individuals of whom historical precedents teach us are much less likely to produce revolutionary ideas. Many of these

individuals are well placed as members of journal editorial staff and can police the distribution of ideas that may be in conflict with their own. Expertise in this sense manifests a double-edged sword, those apparently best placed to decide on innovative ideas are those with the greater motivation to suppress challenging ideas. Or, to take a non-conspiratory angle, might even miss significance because they are unduly wedded to a standard, albeit dysfunctional model.

The pursuit of data and its generation appears to offer little relief from an ossified system and by its very nature follows the dictates of precedent studies at small incremental levels. Innovation in theory is far more threatening to the status quo and might explain why it has been marginalized, even though progress in science cannot function without it. Uncovering bad data found to be supported by statistical sleights of hand wont necessarily change this self-serving culture because there is less incentive that might be believed to reduce the number of even poorly designed studies. After all, if we remove a conservative 50% of studies from the traditional literature as potentially non replicable, that represents a not insignificant threat to the academic food chain. That is, even with the best intentions there remains an inherently insurmountable conflict of interest. Good theory alone can change the data generating enterprise into a more productive process, but not before it seriously challenges entrenched hierarchical interests and the way it is disclosed to the scientific community.

OA appears, at least currently, to show a greater tolerance for hypothesis papers from lesser status sources precisely because they are currently less organized into rigid hierarchical structures. Although, several OA venues exclude hypothesis based submissions, OA is not a monolithic enterprise where all advocates have the same objectives. The irregularities in

publishing criteria are important here for revolutionary science. Such attempts by several publishers to emulate subscription journal policies reflect the the greater status (read here: predicated on broader acceptance) of experimental papers, but represents a similarly restrictive practice in one sense. They are none the less far more inclusive and allow greater liberties to theoretical speculation in their discussion sections. The role of OA in this endeavor is a reaction to an innate conservatism that has entrenched itself in prestigious journals. Youth could well be a key element in this process harking back to a paraphrasing of Max Planck's famous statement of progress being made by one funeral at a time.

# **A practice in data without theory**

A good example of science's contemporary limitations and their pervasiveness in the modern 'alchemist's' claim of transforming data into scientific knowledge is the ubiquitous use of functional neuroimaging and the lack of real meaning of many of the findings, which resemble a cognitive phrenology, a common criticism (Faux 2002). However, there is an unwarranted belief that this will be addressed by "large-scale data mining approaches" and a network mapping-type of phrenology repackaged as a 'cognitive ontology' (Friston 2002; Poldrack 2010).

Several studies of fMRI analysis of resting state data have shown with a nominal family wise error rate of 5% the standard parametric statistical methods to be conservative for voxel-wise inference and invalid for cluster-wise inference for both single subject and group studies (Eklund et al. 2012; 2015). The authors suggest a non parametric solution to false positives. However, with some 28,000 published papers according to PubMed (fMRI in the title or abstract) and parametric tests of the null hypothesis being a common method of analysis, the



question arises not so much whether the findings for many studies are now implied to be obsolete, but whether if this is the case, it dramatically changes our understanding of cognition i.e. is it likely to seriously challenge the known theoretical foundations of psychology? I would argue probably not. This is not because fMRI studies add nothing of value to the investigation of cognition but because as discussed below, there is an insufficient theoretical edifice in the first place.

It is interesting to first briefly discuss what is actually assessed by these fMRI studies. The language is one of 'blobs' (clusters of fMRI generated voxels) and looking at significant levels of activity above baseline spatiotemporal parameters. It is not clear what a 'blob' as methodology could possibly represent when correlated with cognitive or behavioural parameters. A recent debate around a study looking at the role of the anterior cingulate cortex in pain perception nicely illustrates the degree of arbitrariness of results that can be attributed to many fMRI studies. In fact, activation of the dorsal anterior cingulate (dACC) is selective for pain according to the study published in *Proceedings of the National Academy of Sciences* (Lieberman & Eisenberger 2015). A blog post challenged this conclusion arguing for a far broader role and demonstrating a failure to take into account posterior probabilities (<http://www.talyarkoni.org/blog/2015/12/14/still-not-selective-comment-on-comment-on-comment-on-lieberman-eisenberger-2015/>). In fact, the latter analysis suggests an indiscriminate role for dACC in cognition.

How can analysis of fMRI lead to such polar views of the same data? The point is that the data is not self-interpretable and even if the broader functional thesis is ultimately seen as authoritative, the data is more spurious than the discussions admit to. The conflict between the two points of view is an endemic issue in the fMRI literature where concordance is often

difficult to attain largely because it is not bounded by generalizable theory. Indeed, the highly technical nature of statistical analysis of data gives a false impression of a sophisticated scientific endeavor.

No matter how it is dressed such methods remain correlational, not explanatory. Or even constructing the new ‘Tower of Babel’, the connectome as if demonstrating full connectivity patterns under various behavioural paradigms will provide a compelling logic for scientific discovery. It is simply a derivative of the inductive logic criticized by Popper. This does not mean that these technologies are not useful or will not become more so, but their full potential cannot be achieved without better understanding of what they are looking at. They appear to have become a practice in themselves not so much as a tool to progress in science, but one of career enhancement. Thus, the reservations about parametric maps in MRI studies are unlikely to even cause much consternation amongst most academic research labs that are grinding out uninformative data.

A similar criticism can be made with regards to EEG correlates of cognitive function especially relating to the significance of frequency bands, but also specification of the role of event related potentials. For the most part studies using electroencephalography are correlational, but the underlying process is poorly defined. Even with robust epistemic data as in the the role of the hippocampus in anterograde declarative memory there remains an inadequate decades old theoretical model trying to explain the accumulation of half a century of data and observation. The models themselves are rather implausible and lack elegance, but no alternatives are proffered or at least those that are are barely discussed or promoted even though they may be better. Why?

The latest glam technology to assail neuroscience is optogenetics. Its touted revolutionary capacity has led to imminent claims of nobel prizeworthy discoveries. It has taken some time for doubts to emerge with regard to its specificity in determining neural function (confirming that no neuron is an island) and confidence now only exists for those findings that corroborate previously known functions from the crudest of tools, lesion studies (Südhof 2015). Like other techniques it appears that this methodology is prone to the same caveats of other means of study i.e. it is unlikely in itself to 'illuminate' unponderable questions. Compare this premature exuberance with the ground and space projects LIGO and LISA, designed to detect the existence of gravitational waves, a prediction made close to a century ago. There is a far more random application of technologies in neuroscience and clinically than the scientific community would perhaps admit. In fact, taking this latter example the technology makes sense only in the context of the theory.

This also underlies the problem faced in the pursuit of genetic signatures of complex mental disorders as alluded to at the start of this introduction. Genome-wide studies often demonstrate multiple and varied polymorphisms of minor significance where very large cohorts and statistical sophistication have largely been failures. The common excuse is that most of these disorders are polygenic and thus, no one mutation adequately correlates to risk. But even when mutations of a large effect size are found in a family cohort such as the DISC1 in schizophrenia and well-defined syndromic cases of autism, where the genetic abnormality has been well-defined and replicated in animal models, good theoretical models are conspicuously absent. The latter examples if anything, imply that accumulation of more data will not alone determine solutions. The relative failure of the genomics of mental disorders puts paid in turn to a commitment to big data mining as an alternative paradigm for scientific discovery. In part this results from the confluence of multiple genetic pathways.

252

253 The problem with the general type of proselytizing to the cause of more and ever expansive  
254 data sets is that there ultimately must be a confluence in mechanistic aetiology between  
255 genes, relevant pathophysiology and studied behavior. This level of explanation will not  
256 come from the data sets themselves though as the crystallographic example shows, but  
257 requires a level of creativity that even eminent philosophers of the history of science  
258 themselves struggled to explain.

259

260 The practice of data generation creates a rigid food chain with those at the top having little  
261 incentive to change it. The enterprise thus described has little need for a useful scientific  
262 outcome and can and does function in parallel to real progress. If one accepts this there is  
263 little surprise that such a massive proportion of studies fail replication and as long there is a  
264 perceived effort of correcting the record the system can continue to function as it is. In other  
265 words, motives that drive it are not scientific but careerist. The elimination of spurious results  
266 does not necessarily correlate with progress in science. The concept of an ecological food  
267 chain is critical to understanding why it is so difficult to move outside standard scientific  
268 practice and to perceive and adopt better models that can propel a line of study.

269

270 At every level of investigation and particularly in the mind sciences there is a failure of the  
271 modern scientific enterprise to provide solutions and the enterprise itself has degenerated into  
272 metric adornments for career enhancement that is dissociated from real progress in scientific  
273 discourse. The prestige of journals and its exclusionist policies are part of the causal chain in  
274 this distorted picture of modern science. Precisely because scientific production is indexed by  
275 factors that can be modulated and are only indirectly associated with judgment of true

quality, they have become detrimental to the enterprise. OA promises an innovation, albeit imperfect, in the practice of science.

## Fallacies

There is still a pervasive fallacy that journals of high-impact factor (HIF) only publish good quality articles and open access generally poorly screened ones (Osborne 2015). The latter author espouses what one could only refer to as the ‘benevolent dictator model’. A small number of experts in any field will determine the significance of a paper. Increasingly peer review does not actually perform peer review, but the editors make a quick judgment about a paper based on other factors that go beyond careful analysis of content merit. Most likely the perceived quality is based on personalities and institutions behind the work as a default measure of likely impact and desirability in the context of a culture of too few highly significant papers.

Osborne naively claims that what the ‘high quality’ peer review of traditional HIF journals with low acceptance rates does is relieve the scientific community of having to sift through unnecessarily a large amount of low quality papers and that any potentially good work that is missed will be in the minority. This misconception is embraced by those proponents of the current hierarchical structure of journal impact factors with traditional peer review. What it fails to recognize is that it also represents a tool for the control of the ideological landscape and it is highly unlikely that altruism guides the motives of peer reviewers with a stake in this power structure.

In fact, the claim here is that the main impetus for open access is not so much for free availability of state-funded studies, but the hiatus in ideas left by these controlling stakes. Science self-regulates and the traditional enterprise has not met the expectations of a burgeoning community both as outlet, but also for a failure to provide powerful principles that make science successful beyond careerist imperatives.

### **Distortions in the art of science**

What this means and is generally not acknowledged, is that it is not simply a matter of more data or better data or even replication, as important as these regulating activities appear to be (Nosek 2012). It is about the efficiency of the data, that is the targeting of new heuristic principles of research. The research enterprise now requires innovative paradigms based on better theories. This is the major anomaly in traditional, by subscription literature i.e. a disconnect between the manufacturing of data and its interpretation. Hypotheses are generally poorly developed and do not accommodate a wealth of findings, either intra- or inter-disciplinary. Most often data drives hypotheses rather than the reverse case, where good theory creatively spawns experimental design. There are pending questions to solutions, which data as an instrument just can't solve.

This argument parallels somewhat that made about distorted statistical intuition and the importance of informed judgment in comparing Bayesian statistics to the predominance in psychology of a mechanical null hypothesis significance testing (Gigerenzer 1993). It offers an alternative interpretation of the example of *p*-hacking and/or publication bias proposed for non-confirming replication studies of the effects of romantic priming on risk taking behavior (Shanks et al., 2015).

325

326 “... *Good data need good hypotheses and theories to survive.* We need rich theoretical frameworks  
327 that allow for specific predictions in the form of precise research hypotheses. The null hypothesis of  
328 zero difference (or zero correlation) is only one version of such a hypothesis -- perhaps only rarely  
329 appropriate. In particular, it has become a bad habit not to specify the predictions of a research  
330 hypothesis, but to specify a different hypothesis (the null) and to try to reject it and claim credit for  
331 the unspecified research hypothesis. Teach students to derive competing hypotheses from **competing**  
332 **theoretical frameworks**, and to test their ordinal or quantitative predictions *directly*, without using  
333 the null as a straw man” (Gigerenzer 1993).

334

335 Although a very interesting exercise in uncovering disingenuous statistical aberrations, the  
336 focus of investigation in the romantic-priming papers is unduly narrow. Risk-taking is a  
337 complex behavioural phenomenon and to simplify the parameters, often necessary to reduce  
338 covariates to find statistical associations, is also likely to lead to much wasted effort.

339

340 Reexamination of the body of work by the authors was based on a perceived lack of the  
341 plausibility of generalizing prime induction to such a broad range of behaviours and there  
342 were already good reasons to doubt the significance of undertaking such an enterprise  
343 (Newell & Shanks, 2014). Risk behaviour in the studies referred broadly to those associated  
344 with driving, spending and consumption, gambling and even health related factors. The  
345 impetus for the non-replicating studies was largely motivated on theoretical grounds, in spite  
346 of the accumulative clout of 42 of the 43 targeted studies.

347

348 Perhaps, better theoretical development and open discussion might have avoided this  
349 although, such analysis ignores sociopolitical reasons for publication above those of ensuring

valid data. However, even as critical appraisal is available in some journals it tends to be a secondary concern, a shortcoming in itself.

In fact, it is far more difficult to create good theory that can generate good data and most hypotheses form a post script at the end of papers. New theory is practically censured in traditional journals. Generally, their purported reach is rather tame lest it trouble peer reviewers and compromise chances of acceptance. Worse it desperately attempts to show support or otherwise for often outmoded and simplistic models of reputed significance. Good theory in the mind and even biomedical sciences has yet to achieve the status of venerable art as in physics, but the scientific community still seems to ignore more fundamental causes of a disconcerting malaise, believing that better statistics not better theories will remedy all.

Quite apart from the lack of corroboration of quality based on IF as measured by replication and the matter of increased retraction rates in these journals, which are used as counter-arguments to the cultivated impression of superiority, the malaise of science is more fundamental. The underlying presumption that looks to maintain the status quo is that journals have taken the initiative to improve its peer reviewing guidelines based on statistical rigor and the most criticized because of their standing have acted most vociferously.

Important as this remains, it will not change the political landscape, aims to entrench it and even might have a perverse effect of increasing the number of non-replicable studies with HIF or conversely, reduce the likelihood of novel ideas or findings being considered. For example, pre-registered studies are suggested as one of the tools for combatting *p-hacking*, but appraisals of its success showed that a large proportion of pre-registered studies are either not published or change primary outcome measures (Chan et al., 2004; Mathieu et al., 2009).



It will become a bit like the Tour de France where ever more effective means of masking doping are developed and complicit regulators discovered. In any case, as claimed the approach itself wont save the life sciences from themselves.

This is because of the inherent incentive structure of academic publishing and the siren call of HIF journals and their models of practice discouraging innovative ideas that challenge established norms. There is a recent trend for HIF journals not to retract, but to modify or obfuscate if criticisms are raised (Brookes 2014) unless brought into the public domain. The recent coverage of an Imperial College London study published in the journal Science illustrates this perversion of incentive (Okoye et al., 2015). The critique is external to the journal in PubPeer and the response has been relatively muted, one of repeated figure modifications of obvious duplicates and the avoidance of retraction.

In fact, the ramifications are so serious for the current sociopolitical structure of academia that crisis denialists have emerged reinterpreting the statistics in an attempt to show low false discovery rates in the biosciences or even normal prediction intervals of effect sizes in unreplicated psychology papers (Jager & Leek, 2014; Leek et al., 2015). The debate is being played out in specialized statistical journals where the science itself is simply a backdrop: "We have a problem with the identification of *scientific* hypotheses as *statistical* 'hypotheses'" (Gelman & O'Rourke, 2014; Ioannidis 2014). In social psychology others deny that non-replication is even a valid scientific practice ([http://web.archive.org/web/20150421165018/http://wjh.harvard.edu/~jmitchel/writing/failed\\_science.htm#\\_edn2](http://web.archive.org/web/20150421165018/http://wjh.harvard.edu/~jmitchel/writing/failed_science.htm#_edn2)).

The response generally, then is to hide or ignore serious concerns about a study to protect the reputation of the journal rather than to illuminate methodological concerns. Although the industry of self-promotion beyond what most findings are actually capable of delivering is alive and well, the irony is that most excitement is being generated by fraudulent retractions and have become a distraction to the real cause of science's present malaise. Problems of replication, fraud or poor statistical practice are really just symptoms of the malaise not the cause.

Whether the findings in the abovementioned *Science* article ultimately are discredited and regardless of whether the errors are inadvertent is immaterial, the real issues of the paper are the plausibility of reducing a complex biological process to a simple mutation. The 'lymphocyte expansion molecule' (LEM) gene and expression of a protein has many more effects than just CD8<sup>+</sup> T cell expansion in reducing viral or cancer load. LEM has other effects not highlighted in the experiments including premature death of cell lines and is also found in organisms not possessing an immune system. The abrogation of complex interactions for the sake of a simple digestible story line makes great reading and sensational headlines, but also almost certainly underlies the problems detected by in Dijkstra's post in PubPeer (<https://pubpeer.com/publications/B74CF2D21C4A180A5685A30DC06D29>).

The 'darling' model of drug development represented by the *Bcr/Abl* tyrosine kinase and the treatment of chronic myeloid leukemia using a single targeting of the enzyme by imatinib is an exception (Druker et al., 1996). It is no wonder that attempts at replicating its extraordinary success have proved frustrating.

As a counterpoint in the physical sciences, when the large hadron collider (LHC) was constructed to test out well developed theories, the Higg's boson for example, was already part of the scientific vernacular and the experiments were largely confirmatory, not a discovery. Even with the serendipitous discovery of cosmic microwave background radiation (CMB) a theoretical model was already in place to interpret the findings. Serendipity will always be part of the scientific process of discovery, but the landscape has changed and data generation appears to be reaching a nadir with reduced productivity. Making penicillin-type findings because of unexpected moulds growing in forgotten mediums are much less likely. If results from observation vary from those predicted then physical scientists will generally recognize this based on a strong theoretical vantage point and even may be able to address the anomalies.

Indeed, many purely hypothetical articles are deposited in the hugely successful preprint repository arXiv, but not in bioRxiv. PLOS eschews hypothesis manuscripts and PeerJ only accepts them as preprints, all pointing to the misconception of the infallibility of data and the delusion of quality control. So several innovators in the field misunderstand an important impetus for their existence and simply aspiring to HIF look-a-likes by restraining qualitative aspects of content remains inadequate. For scientists to "stop seeing what isn't there", paraphrasing Adam Gopnik (2015), they need a better theoretical framework.

Frontiers does publish hypothesis-based pieces and appears to be a covert reason for recent affronts by establishment tools countering what perhaps is perceived as theoretical vandalism of cherished ideas ingrained by the status of a collaborative artifact. This level of control in the literature is far more surreptitious and often difficult to quantitate as it is often practiced indirectly. But it can be easily surmised by anyone who has submitted an article to a

traditional or high IFJ only to to be directed to well established theories that one is required to pay homage to to even have a chance of their paper being seriously considered.

The reason why similar initiatives in the life sciences as in the physical sciences have so far been met with a muted response are complex and surely include a relative immaturity in some fields, but also the importance of value addition to the proposed publication of data that can trade on the prestige of a group or lab than the heuristic value of the study itself. Some researchers are in a better position to achieve the endorsement of 2-3 reviewers than open their data to the science community in general, which by way of publication has achieved its main goal even prior to open assessment. By ensconcing the data in a few well-traversed threads of theoretical explanation at start and end of a paper, the custodians of the field in question are labelled unchallengeable by HIF approbation and envy.

### **Data's existential crisis**

The simplistic models for data presentation are not necessarily a result of deceptive intent, but most often and more problematic, as there is no simple remedy, they are borne of a profound poverty in theoretical frameworks. The best examples of this existential crisis come from the mind sciences. Pervasive within published studies of data generation are superficially plausible models. A particular example from a study of binocular rivalry in autism shows that unlike normal subjects, those with ASD show a delayed transitional phase of observing mixed percepts (Freyberg, Robertson & Cohen, 2015). The interesting finding was used as evidence for excitatory-inhibitory imbalance model in autism tied to under-expression of GABA receptor subtypes. Quite apart from the fact that one could make the same argument in several other psychiatric conditions including schizophrenia and even

major depression, there is a massive leap in faith between a basic neural property and complex phenomenological findings.

The modern principle of good experimental design is that hypotheses are already formulated prior data collection and this might be used as an argument against the current thesis. This supposedly minimizes post hoc data manipulation. However, hypotheses are formulated with the prospect of data manipulation so there is in a sense a retroactive effect of expected data handling on initial conception of an hypothesis. So hypotheses are constructed not as a means to enlightened theoretical discovery but primarily to serve the purpose of generating simple variables for testing. proposed hypotheses are already compromised by a certain instrumental value upon which the entire publication roller coaster has evolved. Outcomes rather than explanations are prioritized with the inevitable effect of window dressing a prolific, but premature output. It leads to repetitive non-productive studies because it fulfils the criteria necessarily defined by academia including statistical standards, but not science.

It is worth analyzing the study of binocular rivalry in autism further as it it is an eminent illustration of the existential limitations of data driven science. First there is the behavioural association between ASD and binocular rivalry suggesting at least for a subgroup, reduced perceptual suppression. Then, there is an ill-defined association of a failure of association between GABA signalling and perceptual switching. Finally, known properties of GABA as inhibitory neurons is extrapolated as an explanandum of failure of suppression.

This leads to all sorts of methodologically related problems of interpretation. Not only were GABA levels not lower than controls, but 2 of the controls had lower suppression scores than the ASD cohort and there was large variation in performance (<http://drbrocktagon.com/2015/12/21/gaba-autism-and-the-correlation-that-wasnt-there/>). In fact, it is the absence of correlation which is in deemed as evidence for GABA dysfunction

(Robertson et al, 2015). There is strong evidence for the role of neocortical GABA in increasing signal to noise of sensory input exhibited as desynchronization. This basic function would be incompatible with its putative role in phenomenological suppression. This demonstrates how difficult it remains to build higher order theory from basic properties and how arbitrary theory has become to the exercise of data generation when second order models are ignored. These second order models have become default behavioural explanations.

The GABA theory of ASD is an ideal example of theory serving data rather than a converse situation and is based on a theoretical premise initially encapsulated as an increased ratio in ASD between excitation/inhibition (Rubenstein & Merzenich, 2003). It is extended to similar interpretative rationales of other data generating studies so that GABAergic dysfunction of inhibition is a simile for not only perceptual, but also spatial suppression deficits (Foss-Feig et al, 2013). This pervasive 'metaphorical' form of mind-body theory is pernicious and almost always very limited in its application to particular data sets. Distinguishing metaphorical explanations is the first step in delineating what constitutes bad theory. The structure of the poor argument is linguistic: a simile becomes a metaphor becomes an explanation.

Paradoxically, limited but diverse data sets can lead sometimes to an apparent proliferation of ideas, but these are generally restricted to, or surround idiosyncrasies of the methodology used in particular experiments. Reading the literature can give a false impression of competing views in science, but these actually rarely intersect and seem to dominate in parallel, self-interested readership i.e. research subgroups. Examples are theory of mind (ToM) and EEG or fMRI based regional cortical coherence studies. Surely, such incoherence is a sign of pre-paradigm. A caveat is the extension of the concept to open science, which

includes post publication reviews on sites like PubPeer. These debates, however are generally limited to detection of fraud not an analysis of the merits of particular theories.

‘Mock’ theories designed to shape research agendas rather than true discovery are endemic in science. After all, dominance in a field requires at least a semblance of commanding the theoretical space. Of course the ideological landscape is more uneven than this analysis suggests and there are many examples of models that attempt to understand the data being produced. Generally, they are redescrptions of observations, albeit often useful. Examples of these are theory of mind (ToM) deficits in autism or the structure of the visual hierarchy in terms of ventral and dorsal streams. They are part of the piecemeal process in data accumulation, but tend to lead to inductive type inferences. Many of the problems stem from the ghetto-like structure of the research effort where few scientists have a command of what would require extensive training in various disciplines. Most scientists are consumed by a narrow focus of interest and the studies they perform barely allow the time to consider broader theoretical issues.

# **Psi effects, statistics and bodiless data** (preliminary thoughts)

A surreal situation has grown from a debate surrounding a body of literature on the *psi* effect, which refers to extra-sensory perception and psychokinesis. Parapsychology remains unexplained by known physical (except for potentially scalable quantum effects) or biological mechanisms, but has hinged on the statistical met-analyses of precognition studies to gain legitimacy (Bem 2011; Bem et al. 2015). Precognition is "the anticipation of future events that could not otherwise be anticipated through any known inferential process".

Most of the criticisms revolve around methodological issues, poor power and publication bias, although lack of a theoretical framework for positive findings is commented upon making the initial hypotheses implausible (Schwarzkopf 2014; <https://onedrive.live.com/view.aspx?resid=DF3F7227F3844BE2%21100063&id=documents&authkey=!ALOGGrXhpTQYuiIk&wdparaid=%20null>). To the non-statistician the debate may appear opaque, esoteric even because the discussion revolves around the relative merits of novel statistical techniques. It is easy to miss however, that what is actually at stake here is the decisiveness of statistics as scientific method precisely because it is vying for proof of concept in uncovering pseudoscience. In principle this seems like a reasonable assumption. It remains unlikely though that statistics alone will result in resolution, if no other reason that it is very difficult to control multiple confounds and a broader appeal to plausibility or lack thereof will be required.

The prominence of the debate in conservative academic circles serves to illustrate an extreme, but perhaps inevitable conclusion to a data-driven bodiless research culture (mind without body). Plausibility itself will depend on emergence from a pre-paradigmatic stage and resolution of the mind-body problem rather than belief systems as to what constitutes a valid science, not perhaps as an absolute proof against *psi*, but as a better heuristic model.

### **Inefficient science**

A case in point illustrates the distortions in scientific agendas discussed in his book 'Failures' by Firestein (2015) and the work so far conducted and published on A $\beta$  protein as a potential cause of Alzheimer's disease since 2000. From a few papers this has grown incrementally to



5000 a year looking at amyloid- $\beta$  lowering treatments. Teplow found 2000 references to the term in the last 12 months (personal communication). A $\beta$  protein and its role in dementia, is now considered rather simplistic (Kirkitadze, Bitan & Teplow, 2002). This persists in the face of a poor correlation between presence of plaques with level of disease and repeated failures do not appear to have dimmed the ‘perseverative’ pursuit of ‘proteinopathies’ (Castellani et al., 2009). It is even suggested that amyloid- $\beta$  assembled fibrils could be protective.

It is possible if not likely that impact factor journals and the opinions of a small community with a committed interest in the theory dissuade the publication of dissent and thus, mold the agendas of new dependent players. Even if alternative better argued theories are available they are often marginalized because they lack the imprint of validation of HIF journals and influential individuals.

The collaborative elite university-HIF journal industry demonstrates a penchant for simple causal associations without facilitating greater theoretical discussions that might lead to alternative ideas and experimental paradigms. Well-developed competing theories might have fared more productively and a lack of sponsored critical evaluation is probably part of the reason for lack of efficiency that plagues dysfunctional science. Complex ideas require ingenuity and inventive strategies for testing and are less appealing than simple ideas to test and publish outcomes. But the latter also function like a tabloid media of easily digestible facts ignoring analysis of inconvenient data. None the less sponsoring of alternative views would ideally occur by those institutions and journals with the most influence, but require a top-down regulatory ethic that moves beyond individual reputations and authoritarian practices.

600

601 The premise for greater openness is that even with the authoritative power of a finding in a  
 602 HIF journal it rarely translates into a compelling theory irrespective of the perceived standing  
 603 of its authors. Original proponents who come to command the intellectual space associated  
 604 with the dissemination of ideas are resilient irrespective of productivity in the progress of  
 605 research agendas. There is certainly a strong case to be made that the technical innovators in  
 606 a field shouldn't have as much power over the scientific mind because rarely do they provide  
 607 an ongoing heuristic that proves as productive for the field as their achieved status suggests.  
 608 In fact, it appears to hinder the emergence of better ideas coming from elsewhere explaining  
 609 their own data. Ownership of data published in HIF journals has the unfortunate effect of  
 610 giving the right to set research agendas, but mostly based on modest theory.

611

# **612 The cause of data is distorting**

613

614 The distortions of science are compounded by the problem that the gold standard of  
 615 replication is more troublesome than it is generally believed even when motivated by the  
 616 integrity of the exercise (Pulverer, 2015). So the same sources of error may be levelled at  
 617 negative findings and confidence in the absolute authority of non-replicating studies is not  
 618 straightforward (Morrison et al., 2014). Pulverer states of replication:

619

620 “The aim has to be to formally publish only those scientific findings for which we have compelling  
 621 support”.

622

623 But this is unlikely to suffice for the reasons stated here and in the same article and there is a  
 624 danger of not publishing what will be eventually discovered to be valid and important  
 625 findings. So part of the solution might be: to publish where there is compelling theoretical

support i.e. data comments on good theory, but where this is lacking to encourage critical discussions of the merit of individual ideas. The danger remains that apparent agreement on the guiding principles of dominant theoretical models rests more on the limits imposed by the authoritative structure, which takes on an aura of being factual and is difficult to disengage from both politically and intellectually. This in spite of the usual criticisms of inelegance or an accumulation of anomalies from the *outset*.

Yet the data driven culture of HIF publication behaves contrary to this apparent discrepancy allocating discretionary and privileged access to the means of communicating to a broader scientific community by a difficult right of passage. The latter is often dictated by influential mentorship and is for all practical purposes a *sine qua non* of a successful career.

The problem is not unique to HIF journals, but systemic to the all traditional anonymous peer-reviewed journal publications, where the ideological landscape is heavily policed, a form of covert censorship. Peer review as currently practiced outside OA is often seen with some justification as a self-referential opportunity to dictate the course of scientific exposition. The advantage of OA peer review to scientific dissemination of ideas and the bane of OA critics is its regulatory imposition of a non-discriminatory process, at least in theory. Reviewers can suggest improvements, but not arbitrarily block the publication of papers, in fact there is pressure on them to act constructively and fairly. The top down imposition on peer review practice is mandatory if a more open science is the expected result.

HIF journals to some degree have brought the problem on themselves by cultivating a coveted reputation of data reproduction above any potential or post hoc criticism, after all they are the arbiters of significance. The standard tools of ensuring quality in the face of the

difficult task of proper and competent review is a tacit understanding with high ranking universities, precisely what appears to be failing them currently.

There are implications for a broad area of research that touches on the delivery of innovative medical practice and accountability:

“Publication in such a journal can transform a career or influence millions of dollars or more in sales of a product. That concentration of power exerts substantial influence over perspectives and information that are disseminated broadly in the press, and that guide the public and policy makers. In the future, the scientific community may prefer that such influence is more broadly and openly distributed, rather than placed in the hands of the few.” (Krumholz, 2015).

### **Changes are afoot and a program for the future**

The argument in this paper is not simply to rehash a well known Kuhnian perspective, although it remains astounding that an emerging science doesn't appear to reflect on historical precedent, but to make the case for OA as a revolutionary tool for theoretical development of a data driven culture. It is a fascinating case study in the preparadigmatic science in which we are immersed and the reason many players behave as if oblivious to this fact. The case still needs to be remade in the current context and with all the details of mechanisms that define resistance and the the sociopolitical sources thereof. It is also a time for prospective employment of these insights. Perhaps current methodology is even more insidious than in past centuries because the modern psychological sciences function with all the high tech trappings of a real science as if good paradigmatic infrastructure is superfluous. The naïve underlying presumption is that more and better data will unlock the secrets of a field and this myth is propagated by an exclusive hierarchy.

677

678 A lack of adherence to initial intent means that the language of paradigms has partly lost its  
679 significance and fosters a reductionist science of narrow interests and ideologies. Although,  
680 one may argue that Kuhn's paradigms were not defined as an absolute and allowed the  
681 emergence of a narrower perspective there are key broader questions in science which must  
682 be answered before progress in that science is deemed proper or in the case of the  
683 psychological enterprises, that a science even exists.

684

685 There are obviously many teething problems and the resistance to OA is more than an issue  
686 of quality assurance, but one of fear of having one's dominance usurped by new and non-  
687 endorsed groups. Critics of OA are not interested in reform, but silencing a political foe in the  
688 best traditions of Kuhnian thought. With the almost altruistic aims of the OA movement its  
689 limitations should be highlighted but better tolerated by the research community for what it  
690 can offer it beyond traditional means of doing science. This is not merely naïve thinking and  
691 it is readily acknowledged that like a good Marxist theory OA can be usurped for less than  
692 honourable reasons. In any case unwelcome outcomes like predatory publishing, is not  
693 simply a fault of the OA enterprise itself, but emerges from a failure of traditional publishing  
694 and its hierarchical structure, where acceptance rates of 10% or less reflect as much  
695 aspirations of entrenching prestige as quality control.

696

697 Thus, greater accountability in OA could come from the traditional marketplace itself and  
698 counterintuitively opening up subscription-type HIF journals to greater diversity and quantity  
699 of publications without a loss in 'quality control' as currently defined and reduce its  
700 intolerance for radical theory. This does not equate to greater hybrid OA, which fails on the  
701 same grounds, but refers to better access for potential authors without fear of reprisal for new

ideas, not just readers. Such a push would surely reduce the attraction to predatory publishing. As an enduring or interim measure it can function in parallel with other objectives of the open access movement.

## Conclusion

It's easy to get lost in the simple economics of OA but there is a window of opportunity here perhaps largely unappreciated. Looking back in years to come it might be seen as an essential part of the new scientific enterprise. OA is a core part of a likely revolutionary science and that window may well close as individual journals garner more prestige and increasing discretion in what gets published.

This is not meant to reify OA as a tool of disruptive science, which could in principle emerge from other sources. OA may in time not differentiate itself from traditional journal practices. But academia is governed by its own set of rules that have little to do with the science itself and more with careerist objectives and uses covert tools of intimidation and fear to guarantee conformity. Fear of being an outsider, fear of being rejected by an anonymous source of peer review, fear of losing one's livelihood or prematurely ending a chance at tenure. In this sense OA at its best mitigates these dark arts of a shadowy academic practice, but it is naïve to believe it a guarantee of such as other factors come into play. It is quite easy to ostracize novel published ideas by an insidious failure to cite or by demoting the publishers themselves by anti-OA advocates (<http://scholarlyoa.com/2013/11/05/i-get-complaints-about-frontiers/>).

Good science itself is not the only rationale for open science and its democratization, as it no doubt creates its own problems requiring separate solutions. But, the mechanism of good

science as it is currently understood is failing to deliver outcomes as it becomes a stated  
 proxy for vested interests. The economics of subscription publishing and the invective against  
 inappropriate profiteering although may be valid, is not the main issue or even impetus for  
 open access. It is part of the fabric of a new revolution in science and the ideas it generates.  
 This discussion is not meant to deny the importance of data for good theory, but that data  
 driven research will fail to attain its goals without a rebalancing toward better theoretical  
 models that should gain a preeminent status in research publications in the same way they  
 represent the pinnacle of research in the physical sciences. Exceptional theory making is as  
 the name suggests the exception as data production became the default position of practicing  
 science, but an open scientific culture could promote better ideas to be tested. Empirical  
 findings are the foundation of great theories but data should not be treated as a substitute. The  
 replicability crisis is not so obviously a crisis in ideas and will be solved by a revolution in  
 better theoretical models and how they are judged.

## References

- Begley CG, Ellis LM. Raise standards for preclinical cancer research. *Nature* 2012; 483:531-533.
- Bem DJ. Feeling the future: experimental evidence for anomalous retroactive influences on cognition and affect. *J Pers Soc Psychol* 2011;100:407-425.
- Bem D, Tressoldi P, Rabeyron T, Duggan M. Feeling the future: A met-analysis of 90 experiments on the anomalous anticipation of future events. *F1000Research* 2015;4:1188. doi: 10.12688/f1000research.7177.1.

752

753 Brookes PS. Internet publicity of data problems in the bioscience literature correlates with  
754 enhanced corrective action. PeerJ 2014;e313. doi 10.7717/peerj.313.

755

756 Castellani RJ, Lee H.-J, Siedlak SL, Nunomura A, Hayashi T, Nakamura M, Zhu X, Perry G,  
757 Smith MA. Reexamining Alzheimer's Disease: Evidence for a protective role for amyloid- $\beta$   
758 protein precursor and amyloid- $\beta$ . Journal of Alzheimer's Dis. 2009;18:447–452.

759

760 Chan A, Hróbjartsson A, Haahr M, Gøtzsche P, Altman D. Empirical evidence for selective  
761 reporting of outcomes in randomized trials. JAMA 2004;291: (20) doi:  
762 10.1001/jama.291.20.2457.

763

764 Cook D. *et al.* Lessons from the fate of AstraZeneca's drug pipeline: a five-dimensional  
765 framework. Nat Rev Drug Disc. 2012;13: 419-431.

766

767 Druker BJ, *et al.* Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of  
768 *Bcr-Abl* positive cells. Nat. Med. 1996;2:561-566.

769

770 Eklund A, Andersson M, Josephson C, Johannesson M, Knutsson H. Does parametric fMRI  
771 analysis with SPM yield valid results? NeuroImage 2012;61:565-578.

772

773 Eklund A, Nichols T, Knutsson H. Can parametric statistical methods be trusted for fMRI  
774 based group studies? arXiv:1511.01863v1 2015.

775



- 776 Faux S. “Cognitive neuroscience from a behavioral perspective: A critique of chasing ghosts  
777 with Geiger counters.” *The Behavior Analyst* 2002;25:161-173.
- 778 Foss-Feig JH, Tadin D, Schauder KB, Cascio CJ. A substantial and unexpected enhancement  
779 of motion perception in autism. *J Neurosci.* 2013;33:8243-8249.
- 780 Freyberg J, Robertson CE, and Baron-Cohen S. Reduced perceptual exclusivity during object  
781 and grating rivalry in autism. *J Vision.* 2015;15:1-12.
- 782 Friston K. Beyond phrenology: what can neuroimaging tell us about distributed circuitry?  
783 *Ann. Rev. Neurosci.* 2002;25:221-250.
- 784 Gelman A, O’Rourke K. Discussion: Difficulties in making inferences about scientific truth  
785 from distributions of published *p*-values. *Biostatistics* 2014;15:18-23.  
786 doi:10.1093/biostatistics/kxt034.
- 787 Gigerenzer G. The superego, the ego, and the id in statistical reasoning. *A handbook for data*  
788 *analysis in the behavioral sciences.* Hillsdale, NJ: Erlbaum, 1993. pp. 311-339.  
789
- 790 Gopnik A. *SPOOKED*. What do we learn about science from a controversy in physics? *The*  
791 *New Yorker* 2015:Nov 30.
- 792 Ioannidis JPA. Discussion: Why “An estimate of the science-wise false discovery rate and  
793 application to the top medical literature” is false. *Biostatistics* 2014;15:28-36.
- 794 Jager LR, Leek JT. An estimate of the science-wise false discovery rate and application to top  
795 medical literature. *Biostatistics* 2014;15:1-12.
- 796 Kirkitadze MD, Bitan G, and Teplow DB. Review. Paradigm shifts in Alzheimer’s Disease

797 and other neurodegenerative disorders: The emerging role of oligomeric assemblies. J  
 798 Neurosci Res. 2002;69:567–577.  
 799  
 800 Krumholz HM. The end of journals. Circ. Cardiovasc. Qual. Outcomes 2015;8:00-00.  
 801 DOI: 10.1161/CIRCOUTCOMES.115.002415.  
 802  
 803 Leek JT, Patil P, Peng RD. A glass half full interpretation of the replicability of  
 804 psychological science. arXiv:1509.08968v1 2015.  
 805  
 806 Lieberman MD, Eisenberger NI. The dorsal anterior cingulate cortex is selective for pain:  
 807 results from large scale reverse inference. Proc Natl Acad Sci USA 2015;112:15250-15255.  
 808  
 809 Morrison S. Time to do something about reproducibility. eLife 2014;3:e03981.  
 810  
 811 Nosek BA. An open, large scale, collaborative effort to estimate the reproducibility of  
 812 psychological science. Persp Psychol Sci. 2012;7:657-660.  
 813  
 814 Osborne R. "Open Access Publishing, academic research and scholarly communication".  
 815 Online Information Review 2015 39:637 - 648  
 816  
 817 Okoye I, Wang L, Pallmer K, Richter K, Ichimura T, Haas R, Crouse J, Choi O, Heathcote D,  
 818 Lovo, E, Mauro C, Abdi R, Oxenius A, Rutschmann S, Ashton-Rickardt PG. The protein  
 819 LEM promotes CD8<sup>+</sup> T cell immunity through effects on mitochondrial respiration. Science  
 820 2015;348:995-1001.  
 821

- 822 Poldrack RA. Mapping mental function to brain structure: How can cognitive neuroimaging
- 823 succeed? *Persp Psychol Sci.* 2010;5:753-761.
- 824
- 825 Pulverer B. Reproducibility blues. *EmboJ.* 2015;34:2721-2724.
- 826
- 827 Robertson CE, Ratai E-A, Kanwisher N. Reduced GABAergic action in the autistic brain.
- 828 *Curr Biol.* 2016, <http://dx.doi.org/10.1016/j.cub.2015.11.019>.
- 829
- 830 Rubenstein JL, Merzenich MM. Model of autism: increased ratio of excitation/inhibition in
- 831 key neural systems. *Genes Brain Behav.* 2003;2:255-267.
- 832
- 833 Shanks DR, Vadillo MA, Riedel B, Clymo A, Govind S, Hickin N, Tamman AJF,M.
- 834 Puhlmann LMC. Romance, Risk, and Replication: can consumer choices and risk-taking be
- 835 Primed by Mating Motives? *Journal of Experimental Psychology: General (In print).* 2015.
- 836
- 837 Spellman BA. A short (personal) future history of revolution 2.0. *Psychol. Sci.* 2015;10:886-
- 838 899.
- 839
- 840 Südhof TC. Experiments mismatch in neural circuits. *Nature* 2015. doi:10.1038/nature16323.
- 841
- 842 Schwarzkopf DS. We should have seen this coming. *Front Hum Neurosci.* 2014.
- 843 <http://dx.doi.org/10.3389/fn-hum.2014.00332>.
- 844
- 845 Vale RD. Accelerating scientific publication in biology. *Proc Natl Acad Sci.*
- 846 2015;112:13439-13446.

