

心理学におけるオープンサイエンス¹⁾

—「統計革命」のインフラストラクチャー—

三 浦 麻 子

関西学院大学／大阪大学

Open science in psychology: Infrastructure of a statistical revolution

Asako MIURA

Kwansei Gakuin University/Osaka University

Open science is an academic movement to transform scientific research into a more open endeavor. For the purposes of this paper, we define open science as the infrastructure of a “statistical revolution,” which indicates the transformation of fundamental ideas in psychometrics. The four basic components of open science are: (1) open access, (2) open data, (3) data-centric science, and (4) civil science. With regard to open data and open access, we surveyed the present situation of psychological research and considered how it is associated with the statistical revolution internationally and domestically in Japan. To date, the introduction of open science to academic psychology circles has arrived relatively late in Japan. The future prospects of psychology in open science are discussed, especially with regard to what to do and how to do open science as a researcher.

Key words: open science, open access, open data, pre-registration, statistical revolution

キーワード：オープンサイエンス、オープンアクセス、オープンデータ、事前登録、統計革命

1. はじめに

本稿では、「統計革命」のインフラストラクチャーとしてのオープンサイエンスについて、その取り組みの基本理念は何であるかを論じ、心理学界における現状を紹介するとともに、将来を展望する。

オープンサイエンスとは、科学研究をより開かれた活動へと変革していく運動、すなわち、より多くの人々がアクセスでき、より多くの人々が協力し、より多くの人々を巻き込み、より多くの人々から信頼されるものにするための取り組みである（北本, 2018）。近年、心理学界においてオープンサイエンスへの動きが急になってきた理由は主に2つある。まず学術界全般に共通する理由として、インターネットの発展によって、こうした

取り組みを容易に実現できるプラットフォームが整備されるようになり、たとえ大学や学会といった研究組織のサポートがなかったとしても、研究者自身が個人的にそれに着手できるような環境ができたことがある。もう1つは、それを活用する側の心理学界にある程度特有の問題である。ここ数年で、著名研究者によるデータの捏造を伴うような重大な研究不正が発覚したり（e.g. Callaway, 2011）、研究結果の再現性の低さが問題にされたり（Open Science Collaboration, 2015）するなど、社会問題にまでなるような数々の事件が発生し、科学としての信頼性を再構築するための積極的施策が急務とされるようになった。そのうち重要な1つが研究の透明性（transparency）と開放性（openness）の確保であり、具体的にはオープンサイエンス（特にオープンデータとオープンアクセス）の試みである。心理学界には、psychology's renaissance とまで言われるほどにドラスティックな変化（Nelson, Simmons, & Simonsohn, 2018）が

1) 本稿の一部は、科学研究費（15K13122）「社会心理学研究の再現可能性検証のための日本拠点構築」の助成を得た共同研究者たちとの議論と実践に基づくものである。

求められるような、良くも悪くも、オープンサイエンスの必要性が強く認識されるようになってしるべき現状があるということである。

ここでは、心理学界でどんな問題が起こったのか、またそこに根強くあると考えられる研究実践上の種々の問題 (questionable research practices; QRPs) について詳しく述べることはしない (それらについては、例えば本誌 59 巻 1 号「心理学の再現可能性」特集号所収の池田・平石 (2016)、あるいは樋口・藤島 (2018) などを参考にされたい)。本稿で詳しく述べるのは、心理学界がそれらを踏まえた対応策としてどのような試みに着手しているのかである。また、オープンサイエンスはこうした現状の問題点の解消のみを目的とするものではなく、科学にこれまでにない新たな展開をもたらすものである。こうした点についても論じる。既に始まっているサービスや諸学会の対応について、なるべく現時点 (2018 年 1 月) で最新のものまで紹介するが、国際的にはほとんど日進月歩と言ってもよい状況なので、すぐに情報としての鮮度は落ちてしまうだろう。このような潮流がある (あった)、という目でお読みいただきたい。

2. インターネットと科学の出会いと融合

前節の定義にあるとおり、オープンサイエンスという用語は、科学を「開かれた」ものとするための多面的な活動を複合的に表現したものである。ここでまずもって重要なのは、オープンサイエンスが、社会のインターネット化に伴う科学の新しい形であるということである。そこにはインターネットのもつオープン性が強く反映されている。インターネットとはネットワークを形作る情報通信技術 (例えば、電子メール、ウェブ、ファイル転送など) の総称であるが、その大きな特徴は、仕様がすべて公開されていることである。オープンな技術であることで、多種多様なネットワーク間の通信方式の差異が緩和され、相互接続が可能な状態が維持されている。こうした特徴をもつインターネットと科学が会うことにより、従来の科学研究の諸プロセスを包摂するだけではなく、さらに拡大させる 4 つのルート— 1) オープンアクセス、2) オープンデータ、3) データ中

心科学、4) 市民科学 — でオープンサイエンスが進行し始めた。ここではそれらをより具体的に理解するために、武田 (2018) の論考に沿って概略を述べる。日本あるいは世界各国で国家的課題として言及される際 (e.g. 内閣府, 2016) の「オープンサイエンス」は前 2 者を指すことが多いが、武田 (2018) はより俯瞰的に整理している。なお、制度面、技術面におけるオープンサイエンスの基盤については、本特集号の大向 (2018) をご覧いただきたい。

2.1 オープンアクセス

オープンアクセスは、インターネット上で論文などの学術情報を誰でも無償で自由に利用できるようにすることである。学術雑誌の出版が印刷媒体のみに限られていた頃、論文はそれが掲載されている雑誌を購入している図書館に出かけるか、複写を依頼して送付してもらう以外に閲覧の手段がなかった。また、serials crisis (雑誌の危機) とも呼ばれるほどの雑誌購読料の高騰により、その手段も制約されるようになってきた。学術情報を電子媒体によってインターネット上の資源とすることは、こうした問題を解決し、情報アクセスの平等化が推進されるばかりではなく、研究成果の共有と再利用が進み、学際的な研究やイノベーションの創出が促進され、その成果を社会に還元する波及効果が期待される。それを「誰でも無償」で実現することが「オープン」たるゆえんである。著者が論文投稿料を支払う等の方法でオープンアクセスを実現するジャーナル (OA 誌) が急増し、また著者自身が所属機関のリポジトリやプレプリントサーバを利用することでオープンアクセスを実現できる環境も整備されてきた。

2.2 オープンデータ

オープンデータは、デジタルデータの活用促進である。これは学術領域のみならず、国や自治体レベルで蓄積されている公共データの市民による利活用という文脈 (例えば Government 2.0 など) で語られることも多い。Open Knowledge International (2018) の Open Data Handbook は、オープンデータとは表 1 に掲げた 3 点を満たすものであると定義し、これらを前提としてデータの相互運用性を重視している。相互運用性とは、多様なシ

表1 オープンデータの条件（Open Knowledge International（2018）より引用）

条件	説明
利用できる、そしてアクセスできる	データ全体を丸ごと使えないといけないし、再作成に必要な以上のコストがかかってはいけない。望ましいのは、インターネット経由でダウンロードできるようにすることだ。また、データは使いやすく変更可能な形式で存在しなければならない。
再利用と再配布ができる	データを提供するにあたって、再利用や再配布を許可しなければならない。また、他のデータセットと組み合わせて使うことも許可しなければならない。
誰でも使える	誰もが利用、再利用、再配布をしなければならない。データの使い道、人種、所属団体などによる差別をしてはいけない。たとえば「非営利目的での利用に限る」などという制限をすると商用での利用を制限してしまうし「教育目的での利用に限る」などの制限も許されない。

ステムや組織が共同で作業を進められることを意味し、オープンデータの場合は、多様なデータセットを組み合わせることで混ざって使えるということを表す。Open Knowledge Internationalは、これを「バベルの塔」を回避する」という表現で比喩的に説明している。「バベルの塔」とはすなわち、大量のデータセットが存在しているのにそれらを組み合わせることで真の価値を持つ大規模なシステムを作れない状態に陥ることを指す。いくらデータセットが大量に公開されても、それら同士のコミュニケーションができなければ知の塔を構築する作業はいつまで経っても終わらない。そうなることを防ぎつつ、より優れた知見を生み出すためには、オープンデータこそが必要だという主張である。

2.3 データ中心科学

武田（2018）は、オープンサイエンスの普及範囲は、こうした従来の科学的手続きやその所産のオープン化に留まらないと指摘する。まず、研究方法そのもののオープン化である。科学自身がデジタル化・インターネット化することにより、既存の研究方法が洗練されると共に、新しい研究方法が生まれることになった。これが科学の第4パラダイムとしてのe-Scienceである（Hey, Tansley, & Tolle, 2009）。第1の経験科学（自然現象を解明する科学）、第2の理論科学（理論構築的な科学）、第3の計算科学（複雑な現象をシミュレーションにより予測・再現する科学）に続く第4のパラダイムは、data centric science（データ中心科学）とも呼ばれる、いわゆる「ビッグデータ」の科学である。インターネット技術の普及は、探求の対象

とする現象に関してデジタル化された膨大なデータの取得を可能にさせた。そうであるがゆえに、その全体像を経験的に記述することが困難となり、また、理論に基づいた数理モデリングによる分析だけではその解釈が困難となった。そこで、前述3つのパラダイムを統合し、膨大な1次、2次データの存在と、遍在する計算能力の存在を前提とする研究方法としてデータ中心科学が生まれたのである。ここでは、研究者たちは、大量のデータが集約されたインターネット上のデータベースにアクセスすることで研究を進め、それらの統合的な成果が科学的発見に結びつく。

2.4 市民科学

最後に、市民科学（シチズンサイエンス）の普及、つまり科学研究の場そのもののオープン化がある。市民科学とは科学への参加を専門の科学者以外に広げ、これまでになかった新たな科学を作り出すことやその成果を指す（保坂、2018）。例えば、多くの一般市民が研究者や研究機関と連携してデータの収集や分析に参加するような動きが拡大しつつある。市民科学には、研究活動への貢献以外にも、市民の科学への関心を高めたり、科学コミュニケーションを促進したりといった、様々な効果が期待されている。職業科学者という立場が確立した19世紀以来、その研究成果が学術的に認められるプロセスは、専門家たち以外を排除する形で進められてきた。研究成果は論文にまとめて雑誌に投稿され、査読を経て掲載されたものだけが業績としてカウントされる。査読による採否は社会的な有用性という観点ではなく学術的な価値で判断されるため、専門の科学者だけに

よって行われる科学はとかく社会から遊離しがちである。市民科学はこうした風潮を食い止める役割を期待されている。市民科学的な活動は例えば天文学のような領域においては以前から存在・活用されていたが、その場がインターネット化・デジタル化されることは、その発想そのもののオープン性を強化する仕組みとして機能している。「ユーザ参加型研究」を謳うニコニコ学会β (<http://niconicogakkai.tumblr.com/>) などがその典型例である(林, 2015)。

2.5 「研究＝データ」サプライチェーンの成立

以上、オープンサイエンスを構成する4つの要素について述べてきた。「オープン」というキーワードの下に、各要素が必然的に連携して発展しつつあるのがまさに現代の科学の姿である。従来の科学研究において、データ生成(収集)はある個別の研究の中で完結する、ある意味でブラックボックスであった。収集されたデータが「社会」(各研究の外)に出されることはまずなく、研究と社会との接点は、先行研究の文献調査と、ブラックボックスから取り出した内容の成果発表(論文)のみであった。しかしこれからは、論文とデータが一体化し、研究成果はすなわちデータ生成である、という研究＝データのサプライチェーンが成り立ち、その過程にあらゆる人々が関わるようになる、というのがオープンサイエンスの発想である。

3. 心理学におけるオープンサイエンス：国内外の動向

次に、オープンサイエンスの4つの要素のうち、心理学界でも既に大きな進捗がみられるオープンアクセスとオープンデータについて、現在までの国内外の動向をレビューする。

3.1 心理学におけるオープンアクセスの現状

心理学系の学術誌は、国際的に見ると、アメリカ心理学会(American Psychological Association)によるものを除くと、出版社が刊行しているものがほとんどである。それらの多くは研究機関あるいは個人で購読契約をしなければ無料で入手することはできない。ただし、論文の掲載決定後すぐに

最終版原稿(final draft)を研究者自身のウェブサイトや関係する機関リポジトリで公開することを認めるケースも増加しつつある。これにより、一定の条件下でオープンアクセス化が実現される。また、OA誌も多数あり、そのうち多くは著者が掲載料を支払うことによって維持されている。また、購読契約が必要な従来型の学術誌においても、一定期間(embargo)経過後に無料公開されたり、著者が掲載料を負担することでオープンアクセス化できたりする場合もある。

また、ResearchGate (<https://www.researchgate.net/>) や Academia.edu (<https://www.academia.edu/>) など、学術系ソーシャルネットワーキングサービス(SNS)と呼ばれるサービスに、登録者が自らの論文やワーキングペーパー等を掲載して、オープンアクセスを実現させているケースも数多く見られる。例えば池田・平石(2016a)の「追加的ノート」はResearchGateで公開されており、デジタルオブジェクト識別子(doi)も付されている(池田・平石, 2016b)。

しかし、当然ながら営利目的の学術出版社とこうしたオープンかつフリーのサービスのせめぎ合いの様相もあり、大手学術出版社や学会が連合して前述のResearchGateを訴える事案も発生している(Chawla, 2017)。これに応じたResearchGateが2017年11月にオンラインに掲載していた論文のうち、出版社の権利を侵害していた約170万本の論文へのアクセスを制限するという処置を講じたことも報じられた(Matthews, 2017)。一方、OA誌では著者が高額の論文出版加工料(Article Processing Charge; APC)を負担しなければならない場合もあり、これの高騰化が研究者間の格差を招くことが危惧されてもいる。ビジネスとして出版を行う以上、どこからその費用を捻出する必要がある、これまではそれが読者からだったのが著者からになったただけだ、というわけである。オープンアクセスを前提とした成果公開のあり方について、出版社と著者の間で新しいルールの策定と共有が進みつつある段階であると言えるだろう。

日本国内においては、多くの心理学系学術誌は学会が刊行元となっており、電子媒体はたいていJ-STAGE (<https://www.jstage.jst.go.jp/>) を利用して公刊されている。それらのほとんどは冊子体の刊

行から時期をおかず、あるいは1～2年程度の embargo を経てオープンアクセスとなっている。ページ数の超過やカラー印刷など、特別な場合を除いて著者が論文掲載料を支払うケースはほとんどなく、刊行に際する費用の多くは会費によって賄われているので、その対価として会員にのみ投稿する権利を与えている場合が多い。今後はオープンアクセスの電子媒体のみの刊行になるものも増えるだろう(2018年4月現在で電子媒体を主としている(冊子体はオンデマンド印刷の)学術誌には『感情心理学研究』がある)。本誌『心理学評論』は、2018年3月にJ-STAGEで電子媒体による刊行(1年の embargo あり)を始めた。また、著者自身の責任の及ぶ範囲あるいは所属機関のリポジトリにおいて個別の論文の電子媒体を公開することは認められているので、本特集号は特設サイト(<http://team1mile.com/sjpr61-1/>)からそれらへのリンクを張り、掲載論文を一覧することを可能にしている。ただし、心理学評論刊行会は学会ではなく、雑誌刊行にかかる経費は基本的に大学図書館や個人からの購読料によって賄われている。その代わり投稿に会員資格は必要とされない。となれば、オープンアクセス化は単純に購読者の減少につながるかもしれない。

こうした事情を見ると、時代の趨勢がオープンアクセスに向かっていることは間違いなく、そうしなければ著者にとって不利に働くこともありえる。しかし、これまで長らく運用されてきた学術雑誌の出版形態との軋轢が解消されるにはまだしばらくの時間がかかりそうである。

3.2 心理学におけるオープンデータの現状

心理学者にとっての「データ」というと、「ある特定の研究目的のもと収集された、研究者が分析対象とする情報(事実や数値の集合)」が想定されるかもしれないが、ここで言うデータは、より広義に、研究に関するあらゆる資料(狭義のデータ以外にも、研究材料(実験に使った刺激画像、調査票に含めた項目など)や詳細な手続き(実験の教示文や観察した行動を分類する際のカテゴリーとその具体例など)、分析方法(変数の作成方法や分析プログラムのソースコードなど)を指すものとする。

心理学研究にとって喫緊に必要な変革(Nelson

et al., 2018)として強く叫ばれていることの1つがこうした情報の開示であり、これによって、第三者が研究に関する(論文に掲載されている以外の)情報を詳しく知ることができ、研究の信頼性や妥当性を検証できる環境を整えることができるようになる。オープンデータが実現することにより、ある研究成果について、いつでも誰でもその結果の再現可能性を検証したり、2次分析に付したり、あるいはメタ分析に必要な情報を入手することができるようになる。

心理学系の学術誌におけるオープンデータの実現は、電子媒体で公刊された論文の付録(supplementary material)によるものが主流であった。近年では、研究プロジェクトのWebサイトで公開されたり、Center of Open Scienceが運営するオープンサイエンスのためのプラットフォームOpen Science Framework(OSF; <https://osf.io/>)など外部サービスを利用して共有される場合もある。ただしオープンデータを義務化しているケースはほとんどなく、まずそれらの実践を呼びかけつつ普及を目指している段階である。たとえばAssociation for Psychological Science(APS)の刊行している学術誌(例えばPsychological Science誌)では、オープンデータのための試みをOpen Practicesと総称して強く勧奨し、オープンデータ(APSは、狭義のデータを公開することをOpen Data、それ以外の情報を公開することをOpen Materialsと称している)、後述する事前登録のいずれかを満たす論文には図1のようなバッジをつけてそれを明記している。

日本では、例えばJ-STAGEに「電子付録」というオプション機能がある。ただし現時点では、日本国内の心理学系の学術誌で積極的にオープンデータに取り組んでいるところはほとんどない。多くの学術誌がJ-STAGE等を活用して論文のオープンアクセスは実現させているが、「電子付録」を活用しているものはまだ見当たらない。本稿の著者が編集委員長を務めている日本グループ・ダイナミクス学会『実験社会心理学研究』では、2017年10月に投稿規程を改訂して「研究に用いた調査票、動画、音声、高解像度の写真、ローデータなど、本文と図表に含めるのは困難な資料や、審査の際に有用な資料を、付録として添付することができる。」という条項を追加し、審



図1 APSが刊行する学術誌でOpen Practicesを実践している論文に付されるバッジ。左から順にデータ公開、研究素材公開、事前登録が行われていることを意味する（APSのウェブサイト（<http://www.psychologicalscience.org/publications/badges>）より引用）。

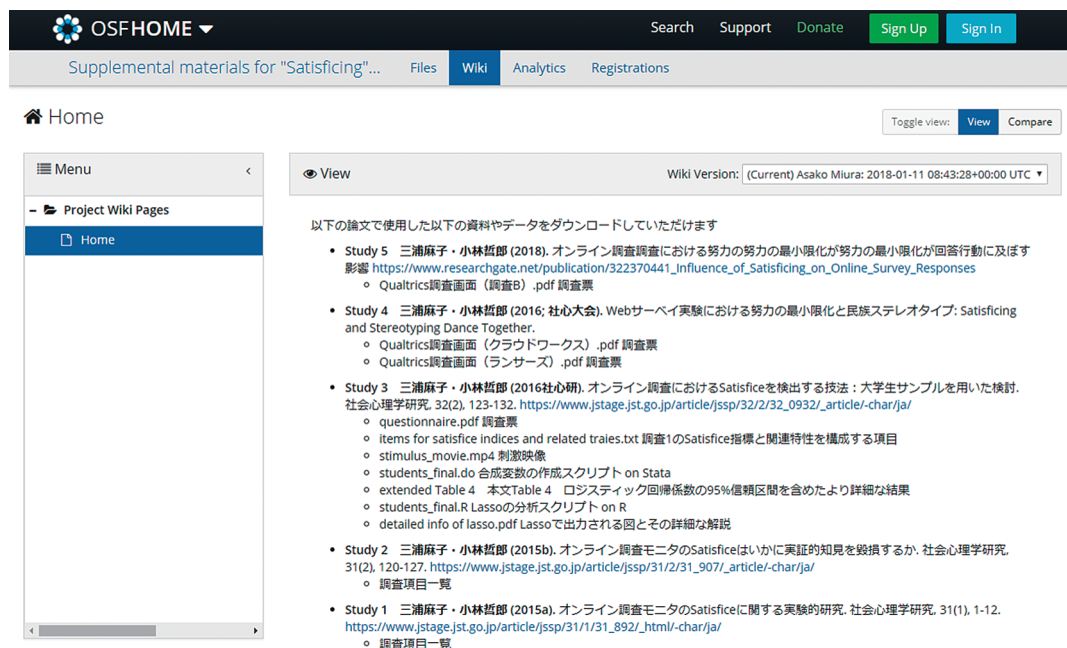


図2 Open Science Framework を利用した研究者個人によるオープンデータの実践

査を経ることを条件としてオープンデータを勧奨しているが、2018年4月現在でまだ利用例はない。学術誌の体制がそうであることが示すように、日本の多くの心理学者にはまだ（国際誌で明示的に求められる場合は別として）オープンデータに積極的に取り組もうという意識は浸透していないというのが実状であろう。

このように日本国内の心理学系の学術誌のオープンサイエンスに向けた試みはまだ貧弱と言わざるを得ず、実践は研究者の自由意思に任されている状態である。制度は整備されるべきであるが、学術誌側に制度がないからできないというわけではなく、オープンデータ実践の場は個人にも開かれている。前述のOSFは個人利用ができるので、例えば本稿の著者は、OSFで研究材料や手続き

を個人的に公開して（図2）、公開された論文の脚注としてそのURLを含めることをよくする。また、社会調査のローデータと調査票を東京大学社会科学研究所附属社会調査・データアーカイブ研究センターによるSSJDA（Social Science Japan Data Archive）に寄託したりもしている。

オープンデータと密接な関連を持つ取り組みが事前登録（pre-registration）である。「オープン」という視座から表現すれば、研究目的や仮説・手続き（取得予定のサンプルサイズ、実験デザイン、統計分析手法など）の着手前段階でのオープン化、ということになる。これらを第三者に向けて「事前に」公開することで、「正」の研究手続きをあらかじめ定めておくのである。OSFなどで研究計画の内容をタイムスタンプと共に事前

OSFHOME

My Projects Search Support Donate Asako Miura

Pre-analysis plan for the "Media Priming Replication Project"

Contributors: Tetsuro Kobayashi

Date created: 2015-05-17 03:46 PM | Last Updated: 2015-05-18 11:19 AM

Category: Uncategorized

Wiki

This pre-analysis plan describes a replication of Experiment 1 in Iyengar, Kinder, Peters, and Krosnick (1984). The current experiment was designed to ask participants to watch edited news videos, in which the degree of coverage of a specific issue (the U.S. base relocation in Okinawa, Japan) was to be manipulated. We plan to collect pre- and post-treatment measures from the participants, including...

Read More

Citation

osf.io/szak6

Tags

pre-analysis plan

Files

Name	Modified
Pre-analysis plan for the "Media Priming Replication Project"	
OSF Storage	
pre-analysis_plan20150517.docx	2015-05-17 03:47 PM

Recent Activity

- Tetsuro Kobayashi registered Pre-analysis plan for the "Media Priming Replication Project" 2015-05-18 11:19 AM
- Tetsuro Kobayashi made Pre-analysis plan for the "Media Priming Replication Project" public 2015-05-18 11:03 AM
- Tetsuro Kobayashi registered Pre-analysis plan for the "Media Priming Replication Project"

図3 Open Science Framework を利用した研究の事前登録例

登録することによって、その時点以降の改変が不可能になるよう「凍結」する。

事前登録によって、研究遂行中のQRPs, 例えばデータ分析後にその結果に合致するような仮説を立てる、統計的検定が有意になるまでデータを収集する（有意になった時点で収集を打ち切る）、相関が有意になった変数だけを報告するといった行為の抑止が期待される。例えば、Iyengar et al. (1984) のメディアのプライミング効果に関する実験を追試した Kobayashi, Miura, and Inamasu (2017) は、OSF で先行研究に合わせた実験デザイン、および具体的な変数の作成方法や回帰分析のモデルを事前登録してから実施している（図3；<https://osf.io/bgk29/>）。

再現可能性の検証を目的とした集合的な追試研究に事前登録システムを導入し、インターネット上のオープンなプラットフォームで実現している例に Many Labs と名付けられた一連のプロジェクトがある（例えば <https://osf.io/wx7ck/>）。プロジェクトが追試対象として選定した研究に関する刺激

や教示文など、手続きに関するあらゆる資料が提供され、参加者はこれに忠実に従ってデータを収集することが求められ、それらもすべてプロジェクトで共有される。参加者ごとに独立したサンプルによる追試が行われ、総数 6,344 名にのぼるデータによって再現可能性のバリエーションを検証したものが論文としてまとめられた (Klein et al., 2014)。事前登録とオープンデータの組み合わせによって、こうした取り組みを世界的規模で迅速に、また誰にでも開かれた形で行うことができる。Many Labs の成果が大きな反響を呼んだ後、*Psychological Science* 誌は Registered Replication Reports という論文カテゴリーを新設した。ここでは追試プロジェクトを事前審査した上で公開し、原典の著者と密にやりとりをしながら進めることが可能となっている。こうした実践は、前述した「研究＝データのサプライチェーン」に通じる。

研究機関による倫理審査の次のステップとして、研究としての信頼性と妥当性を当該領域の専

門家によって事前に認められることを目的に事前登録を行うためのシステムが、事前審査付き事前登録制度である。この制度では、これまでの論文は研究が終了してから学術誌に投稿されるものだったのを、データ取得以前に投稿することを求める。つまり、研究者は事前登録の内容を投稿し、それが審査される。審査を通過したらその内容が学術誌によって公開される。その後実際にデータが収集され、事前の計画に沿った分析結果がどのようなものであろうと、その成果がその学術誌に掲載される。つまり、従来生じがちだった「仮説を支持する結果でなければ掲載されにくいので支持されなければお蔵入りさせる」という公開バイアスをなくす方向に働く。この制度を導入している学術誌は Center for Open Science のウェブサイト (<https://cos.io/rr/#RR>) にリストされており、心理学系のものもいくつかある。

事前登録は研究実施前にその計画を詳細に公開することにほかならない。前述した追試プロジェクトのような目的の下であればともかく、新しい研究について事前登録をすれば、それを目にした同業者が研究を「横取り」して先に成果を挙げてしまうリスクはゼロではない。例えば OSF では、事前審査中は非公開とするオプションが設けられているが、それでも審査者という第三者に実施前の計画を提示することにはなる。事前審査付き事前登録制度は、当該研究者コミュニティ内で「私の仲間たちは横取りなどしない（できない）」という信頼（ないしは、安心）が共有されていることで成り立つシステムではある。

4. 「統計革命」とオープンサイエンス

本稿で「インフラストラクチャー」と副題をつけたとおり、こうしたオープンサイエンスの潮流は、本特集号で主に扱う様々な「統計革命」と密接な関係にある。例えば、心理学のデータ解析における新しい潮流を正当に評価するために、オープンデータは必要不可欠である。

清水（2018）を皮切りに、特集号所収の多くの論文で採り上げられている統計モデリングでは、データが何らかの分布から発生しているというアイデアのもと、確率モデルをデータに当てはめて現象の理解と予測を試みる。伝統的な統計学に

よる分析が、データを何か特定の確率モデルにもとづく分析手法に当てはめることによって遂行されるのとは異なり、どのようなモデルで表現するかを分析者自らが決めるので、格段に自由度が高い。言い方を変えれば、伝統的な統計学による分析とは異なり、統計パッケージソフトウェアが提供するコマンドに当てはめることができなければ分析そのものが不可能、という事態にはならないかわりに、どのような分析をしたのかを詳細に報告する必要性が高い。となれば、確率的プログラミングによって記述されたモデル（分析コード）の公開は必須だろう。また、コンピュータの計算速度の向上が計算量の膨大さを乗り越えさせ、プログラミング言語の進歩がモデリングの試行錯誤を容易にさせた。確率的プログラミングで用いられる言語には、文法さえ理解すれば数学やプログラミングについての高度な知識を必要とせずその意味を理解することができるものが多い。例えばベイズ統計モデリングであれば、Stan という言語で確率モデルを記述しさえすればマルコフ連鎖モンテカルロ法（MCMC）によって自動的にパラメータが推定できる。つまり、分析コードとデータの両方がオープンデータ化されていれば、査読者や読者が分析結果をトレースしたり、さらにアレンジを加えた分析をすることが容易になる。

また、サンプルサイズ設計（村井・橋本，2018）が、事前登録による研究計画の着手前段階でのオープン化と密接な関わりがあることは言うまでもない。先行研究の結果等を手がかりとしてデータ収集開始前にサンプルサイズ設計をしておくことは、特に統計的仮説検定を核とする伝統的な統計学においては、推定の精度を保証し、あるいはデータ収集後の QRP を抑止するための不可欠な要素である。事前登録をするなら、収集すべき適切なサンプルサイズとその算出根拠は必ず含めるべき情報となる。

さらに、前述したとおり、統計モデリングは分析の自由度が高いが、その代わり計算量が膨大で多くの試行錯誤を必要とする場合が多いので、オープンデータを多くの研究者が「よってたかって」解析することがよりよい研究成果のために効果的である。こうしたオープンデータ解析は、発想としてはデータ中心科学に近く、R Markdown, Github などのツールを活用して実施され始めてい

る。こうした解析の現場への非専門家の参入は従来よりも容易で、それはまさに市民科学にもつながる。分析結果を活用した論文執筆、あるいはウェブスクレイピング（ウェブサイトから情報を抽出するソフトウェア技術）によるものも含めたデータ収集そのものもオープンな場で行われるようになり、一連の研究プロセスのほとんどすべてがオープンになるのもそう遠くない将来かもしれない。

こうして「統計革命」とオープンサイエンスの関わりを将来像も含めて考察すると分かるのは、とかく伝統的な統計学（あるいはここしばらくの心理学の統計解析におけるローカルルール）に対するアンチテーゼとしてのみ受け取られがちな「統計革命」は、オープンサイエンスと同期的に進行することによって、心理学の前途を大きく拓く可能性を持っているということである。そうやってようやく統計的データ解析は、ごく卑近な表現をすれば「統計パッケージの使いこなすスキル」のような、限られた専門家にしかできない秘技のようなものではなく、誰しもが活用できる汎用のツールとして、心理学（あるいはそれを越えた科学全体）の新たな発展に貢献しうる存在となるだろう。

5. オープンサイエンスとしての心理学

本稿では、オープンサイエンスの基本理念とそれを実現する際の4つのルートについて概説した上で、心理学における国内外の現況を紹介し、本特集の主眼である「統計革命」との関連について論じた。心理学者にとってオープンサイエンスは、「統計革命」がそうであると同様に、自分たちが創出する科学的知見の信用を担保することに資するとはいえ、現時点ではやや挑戦的にも見える試みかもしれない。しかし、既にそれは絵に描いた餅ではなく、着実に実行すべき時が来ている。最後に、どういうスタンスで研究に臨めば心理学者はオープンサイエンス時代を生きる科学者として幸せになれるのか、私見をまとめる。

まず、結果的に何が検証されたかではなく、何を検証すべきだと考えていたかというアイディアに科学的価値を置くために、自分の研究は何のために何をやるものなのかを、着手前にできる限り

明確かつ具体化することが強く求められる。その手続きとして事前登録は有用で、その段階で審査を受けることができればより精緻化される。これは探索的研究を制約したり排除したりするものではなく、どこまでの範囲の分析が検証的で、どこからは探索的か（つまり今後の検証が必要とされるか）の線引きを明確にすることが重要だという主張である。研究に区切りがついてから遡及的にその過程をオープンにするのは困難だという以前に単純にQRPとなる可能性が高いことは、人間には認知バイアス（例えば後知恵バイアス）があることをよく知っている心理学者なら当然理解できるはずである。

そして、研究成果が論文として公刊される際には、データや関連資料も同時に公開して、自分の研究は何のために何をしたものなのかをできる限り明確かつ具体化する。オープンデータにより、その研究の展開可能性は拡大する（とはいえおそらくロングテール現象が生じるのでほとんどの研究は展開しない可能性の方が高い）。ここで重要なのは、その際の実施者を選ばず、むしろ見知らぬ誰かにいつでも始めてもらえるようにしておくということである。展開があればその研究もオープンデータとなることで、研究＝データのサプライチェーンがより密に繋がっていく。

最後に、科学研究を専門家だけの特権にしないことも意識したい。オープンサイエンスは知の生産者と知の消費者の関わりを密にし、将来的には両者の境界をなくす試みである。そして心理学は本来、科学知と日常知が近くあるはずの学問領域である。常にその生態学的妥当性の確保が懸案とされながらも実現が困難であり続けてきたように、課題解決を科学者だけでするのは難しいこともある。市民科学の導入は、それを助けるだけでなく、より実り多い知見を導くかもしれない。

文 献

- Callaway, E. (2011). Report finds massive fraud at Dutch universities. *Nature*, 479, 15. <http://doi.org/10.1038/479015a>
- Chawla, D. S. (2017). Publishers take academic networking site to court. *Science Magazine*. <http://doi.org/10.1126/science.aag1560>
- 林 和弘 (2015) オープンな情報流通が促進するシブズンサイエンス（市民科学）の可能性 科学技術動

- 向, 150, 21–25. <http://data.nistep.go.jp/dspace/bitstream/11035/3097/1/NISTEP-STT150J-21.pdf> (2018 年 1 月 26 日)
- Hey, T., Tansley, S., & Tolle, K. M. (2009). *The fourth paradigm: Data-intensive scientific discovery (Vol. 1)*. Redmond, WA: Microsoft Research. <https://www.immagine.com/eLibrary/ARCHIVES/EBOOKS/M091000H.pdf>
- 樋口匡貴・藤島喜嗣 (2018) アスタリスク～真実の石を求め (すぎ) て ヒューマンインタフェース学会誌, 20, 12–16. <https://osf.io/zua7d> (2 月公刊)
- 保坂直紀 (2018) 科学者と市民の協働を目指すシンチズンサイエンス. http://scienceportal.jst.go.jp/news/newsflash_review/review/2018/01/20180117_01.html (2018 年 1 月 26 日)
- 池田功毅・平石 界 (2016a) 心理学における再現可能性危機：問題の構造と解決策 心理学評論, 59, 3–14. https://doi.org/10.24602/sjpr.59.1_3
- 池田功毅・平石 界 (2016b) 池田・平石 (2016) 「心理学における再現可能性危機：問題の構造と解決策」に関する追加的ノート. <http://doi.org/10.13140/RG.2.1.3393.6247/1> (2018 年 1 月 26 日)
- Iyengar, S., Kinder, D. R., Peters, M. D., & Krosnick, J. A. (1984). The evening news and presidential evaluations. *Journal of Personality and Social Psychology* 46, 778–787. <http://doi.org/10.1037/0022-3514.46.4.778>
- 北本朝展 (2018) オープンサイエンス. (Open Science) <http://agora.ex.nii.ac.jp/~kitamoto/research/open-science/> (2018 年 1 月 26 日)
- Klein et al. (2014). Investigating variation in replicability: A “many labs” replication project. *Social Psychology*, 45, 142–152. <http://doi.org/10.1027/1864-9335/a000178>
- Kobayashi, T., Miura, A., & Inamasu, K. (2017). Media priming effect: A preregistered replication experiment. *Journal of Experimental Political Science*, 4, 81–94. <http://doi.org/10.1017/XPS.2017.8>
- Matthews, D. (2017). Publishers push ResearchGate harder in copyright battle. *Time Higher Education*. 2017 年 11 月 9 日付記事. <https://www.timeshighereducation.com/news/publishers-push-researchgate-harder-copyright-battle> (2018 年 1 月 26 日)
- 村井潤一郎・橋本貴充 (2018) 統計的仮説検定を用いる心理学研究におけるサンプルサイズ設計 心理学評論, 61, 116–136.
- 内閣府 (2016) 第 5 期科学技術基本計画 <http://www8.cao.go.jp/cstp/kihonkeikaku/5honbun.pdf> (2018 年 1 月 26 日)
- Nelson, L. D., Simmons, J., & Simonsohn, U. (2018). Psychology’s renaissance. *Annual Review of Psychology*, 69, 511–534. <http://doi.org/10.1146/annurev-psych-122216-011836>
- 大向一輝 (2018) オープンサイエンスと研究データ共有 心理学評論, 61, 13–21.
- Open Knowledge International (2018). *Open Data Handbook*. <http://opendatahandbook.org/guide/ja/what-is-open-data/> (2018 年 1 月 26 日)
- Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science*, 349, aac4716. <http://doi.org/10.1126/science.aac4716>
- 清水裕士 (2018) 心理学におけるベイズ統計モデリングの可能性 心理学評論, 61, 22–41.
- 武田英明 (2018) ORCID とオープンサイエンス 平成 30 年第 17 回分類学会連合公開シンポジウム (2018 年 1 月 6 日・国立科学博物館) 講演資料. <https://www.slideshare.net/takeda/orcid-85786283> (2018 年 1 月 26 日)

— 2018. 2. 3 受稿, 2018. 3. 7 受理 —