

(2. コラム 行く・読む・感じる)

2-3. 信頼性のある研究のための事前登録の実践

水野 景子

近年、国内外の心理学分野においてオープンサイエンスの機運がますます高まっている（三浦, 2018; Nosek et al., 2022）。筆者の専門とする社会心理学でも、2022年に国内誌「社会心理学研究」にオープンサイエンスバッジが導入された。著者からの申請があった論文には、論文で報告されるデータが公に入手可能であること、実験・調査で用いたマテリアル（質問紙、実験刺激等の一次資料）が公に入手可能であること¹⁾、研究計画が事前登録されていることを示す3種類のバッジが論文タイトルの近くに表示される（日本社会心理学会, 2022）。

本コラムでは、オープンサイエンスの一つの要素である研究の事前登録とは何なのか、どのようなメリットやデメリットがあるのかについて、筆者の経験を踏まえて紹介する。筆者の行っている研究では、実験の中で小グループに分けられた参加者同士が相互作用をするタイプの実験²⁾を扱う。コラム内では、このようなタイプの実験ならではの事前登録のしかたにも触れる。また、近年議論されている二次データ分析研究の事前登録についても取り上げる。

事前登録の意義

事前登録は、実験や調査の実施前、二次分析であれば分析をする前に、研究計画（おもに仮説、方法、分析プラン）をオンライン上の事前登録用プラットフォームに登録することである。登録するとタイムスタンプが押され、以後編集ができなくなる。追加で資料をアップしたり、修正をしたりした場合は、その都度タイムスタンプが押される。すなわち、事前登録とは時系列に沿って自身の研究を記録しておき、それを読者に伝えるための枠組みといえるだろう。

事前登録の最も大きな意義は、信頼ある研究結果として読者に自身の研究を伝えられる点にある。研究が論文という形で発表された後では、研究者がどのタイミングで仮説を思いついたのか、もともと考えていた分析計画がどのようなものであったのかといった時系列を確認することができない。「効果のあった」データや分析結果を見た後に論文を書こうとすると、そのつもりはなくとも、あたかも初めからその効果があると考えていたかのようにストーリーや研究の仮説を組み立ててしまうことがある（Nosek, Ebersole, DeHaven, & Mellor, 2018）。また、データの分析方法についても、変数をどのように選択してどう変換するか、分析アプローチをどうするかについて多くの選択肢があり、方法によって全く異なる結論が得られることがある

1) オープンデータについては武藤（2022）に詳しい。

2) 研究例として Li et al. (2018) や Titlestad, Snijders, Durrheim, Quayle, & Postmes (2019) がある。

(Silberzahn et al., 2018)。すると、多くの分析を試したうえで「効果のあった」ものだけを報告することが起こり得る。多くの研究で使われる帰無仮説検定では偽陽性率を5%以下におさえる（すなわち有意水準を5%以下にする）ことが通例だが、上記のようなことを行えば偽陽性率は高くなる（Simmons, Nelson, & Simonsohn, 2016）。事前登録を行うことで、結果を見る前の事前予測と、結果を見た後の事後予測を研究者自身が明確に区別することができる（Nosek et al., 2018）だけでなく、事前予測と事後予測にずれがないことを読者に示すことができるのである。

また、上記の内容と関連するが、事前登録には、無自覚な QRPs（疑わしい研究実践）を防ぐ機能もある。「疑わしい」とあるためあたかも研究者が悪意を持って不正をしているかのようと思われるかもしれないが、分析結果を見た後に意図せず結果に沿った仮説を書くことや、効果のあった分析結果のみを報告することも QRPs に含まれる（John, Loewenstein, & Prelec, 2012）。

筆者の行った事前登録

事前登録のプラットフォームとして Open Science Framework (OSF) を使用した。OSF には事前登録用のテンプレートがいくつか用意されており、必須事項を埋めることで事前登録ができるようになっている³⁾。具体的な操作方法や、各項目に書くべきことについては長谷川ほか(2021)を参考にした。当プラットフォームが英語なのもあり「事前登録は英語でするもの」というイメージが強かったが、母国語の日本語で登録した。英語で書いておくに越したことはないが、投稿予定の論文誌や自身の書きやすさに応じて、どの言語で登録してもよいだろう。もしあとから英語版が必要になった場合には、事前登録をそのまま翻訳し、添付ファイルとしてアップする方法がある。

図1に、筆者が行った事前登録の「データ収集方法」セクションの一部を示した。筆者が行っている実験では、集まった参加者を4名のグループに分け、参加者同士でゲームを行ってもらおう。4名の参加者がそろわなければ実験が成立しないが、当日の実験開始時間になるまで何名の参加者が集まるかわからない。そのため、全パターンについて記載することで対応した。

3) 9つある（2023年現在）テンプレートのうち長谷川ほか（2021）で紹介されていた OSF Preregistration テンプレートを使用した。他には、社会心理学に向けて「操作チェック」などの項目が用意されており、かつ各項目に何を記載すればよいか van't Veer & Giner-Sorolla, (2016) で詳しく解説されている Preregistration in Social Psychology (van't Veer & Giner-Sorolla, 2016) テンプレートや、必須項目が最小限で、方法まで書いた原稿をそのままアップして事前登録ができる Open-Ended Registration テンプレートなどがある。

Data collection procedures

実験日時
 2021年11月19日(金)～2021年11月26日(金)の8日間で、計30回の実験を行う。各実験回の最大人数は10名である。26日(金)の30回目の実験が終了した時点で60グループ分のデータが得られていなければ、11月27日(土)～11月30日(月)に追加の日程を用意し、同じく最大人数10名で実験を行う。

参加者
 授業に関するお知らせを通して、関西の私立大学で2021年度後期の心理学関連の授業を受講する学生に実験参加を依頼し、参加者を募集する。学生は、実験に参加することで授業の追加の授業成績と実験謝礼を得ることができる。実験謝礼は公共財ゲームの結果によって変動し、電子版Amazonギフトカードをメールで送信することにより支払われる(謝礼の計算方法については「手続き」セクションに記載している)。
 各実験回に集まった参加者の数によって、以下のように実験を実施する。

- ・3名以下の場合: 全員に個人で行う課題を割り当てる。
- ・4名の場合: 本実験課題を1グループ分行う。
- ・5名～7名の場合: 早く入室した4名の参加者を1グループ分本実験課題に割り当て、その他の参加者には個人で行う課題を割り当てる。
- ・8名の場合: 本実験課題を2グループ分行う。
- ・9名、10名の場合: 早く入室した8名の参加者を2グループ分本実験課題に割り当て、その他の参加者には個人で行う課題を割り当てる。

図1 著者が行った事前登録の一部

事前登録において重要な要素が、サンプルサイズを事前に決めて登録しておくことである。サンプルサイズは統計的検定の精度にかかわるためである。サンプルサイズの決め方には様々な方法がある⁴⁾が、検出力に基づくサンプルサイズ設計が多くの研究で行われる。筆者が事前登録を行った研究は追試研究であった。追試研究の場合には、検出力にもとづくサンプルサイズ設計を行うこと、そして可能であれば十分な検出力のサンプルサイズのデータを得ることが特に重要になる。なぜなら、仮に統計的に有意な効果が確認されなかったという結果が得られたときに、それが追試の失敗を意味するのか、検出力不足であったのかが判別できなくなるためである。

検出力に基づくサンプルサイズ設計では、効果量、有意水準、検出力⁵⁾から必要サンプルサイズを計算する。加えて、筆者が行っている、小グループに分けられた参加者同士が相互作用をするタイプの実験では、参加者同士の行動がグループ内で似通ってきて⁶⁾、グループ内の参加者の行動や尺度への回答の相関(級内相関)が高くなる可能性がある(Titlestad et al., 2019)。級内相関があることは、参加者数に対する情報量が人数分ないことを意味する。そのため、級内相関がないときよりも必要なサンプルサイズが大きくなる。よって、級内相関を考慮したサンプルサイズ設計が必要となる。そのために、切片にグループ間の変量効果を仮定した線形混合モデルのサンプルサイズ設計を行った。2群の平均値差についてのサンプルサイズ

4) 村井・橋本(2018)に詳しい。

5) 詳しくは大久保・岡田(2012)や村井・橋本(2018)を参照されたい。

6) 他のグループメンバーの行動を見て自分の行動を変えるため。

設計であれば、宇佐美 (2011) 内の式 (22) を用いて必要サンプルサイズを算出することができる。

図2は、筆者が行った事前登録のうち、サンプルサイズについて記載している部分の一部である。何名の参加者が実験に参加するかが不明であったため、検出力が0.8の場合、0.9の場合の2つのパターンについて必要サンプルサイズを計算し、それぞれ下限と上限にした。また、300名を超える参加者を確保することは難しい旨も率直に書いた。

サンプルサイズの算出にはR3.6.0 (R Core Team, 2019) および宇佐美 (2011) の式(22)を使用した。検出力が0.8の場合のサンプルサイズは184名(4名×46グループ)、0.9の場合のサンプルサイズは240名(4名×60グループ)であった。1グループにつき1名多く募集すると300名の参加者が必要だが、300名を超える参加者を確保することは難しいため、検出力(1-β)の上限は0.9とした。

よって本研究のサンプルサイズは、上記の効果量d、有意水準α、級内相関(ICC)のもとでの検出力(1-β)の下限が0.8、上限が0.9になるように、少なくとも184名(46グループ)、目標は240名(60グループ)に設定する。募集人数は230名~300名である。

Stopping rule

1回目の募集で240名(60グループ、募集人数は300名)が集まらなかった場合は再度参加者を募集し、2回目の募集でも集まらなかった場合はその時点で的人数で実験および分析を行う。

図2 筆者が行った事前登録のサンプルサイズセクションの一部。図中の **Stopping rule** (停止規則) はデータ収集をやめる基準のことで、事前登録には必ず書く必要がある。

もし、検出力に基づくサンプルサイズ設計をした結果、現実的に不可能な数が算出されてしまった場合はどのようにすればよいのだろうか。実際に、筆者がある研究のサンプルサイズ設計をした結果、 $N = 6390$ という値が算出されたことがある。実験リソースの観点から、800名の参加者に協力を依頼するのが限界であった。このときは、限られたリソースで実現できることを *As a result, the required sample size would be $N = 6390$. However, we can only collect 800 data sets due to funding limitations. With this sample size, we can only detect $d = 0.568$ with an 80% probability* と事前登録に記載した⁷⁾。

事前登録の副次的なメリット

データ収集から分析、研究結果の公表までが非常にスムーズに進んだ。たしかに、特にサンプルサイズ設計と分析プランの登録においてあらかじめ考えておかなければならないことが多かったが、何をどのように分析してどのような結果が出れば仮説が支持されたといえるのかがあらかじめ決めて分析用のコードを用意しておけば、その通りに分析をしてすぐに研究結果を報告できる。

7) 「母集団において $d = 0.568$ の効果量があった場合には 80% の検出力を確保して有意な結果を得たいが、それよりも小さい効果しかない場合は、(実質的な効果としては注目に値するほどでもない) 有意な結果が得られなくてもよいという立場」(村井・橋本, 2017) をとっている。

二次データ分析の事前登録

既存のデータを分析する研究における事前登録についても、議論が進んでいる。Van't Veer & Giner-Sorolla (2016) は、事前登録の基本である仮説設定の期間と、その後の調査や探索⁸⁾の期間を明確に区別することができる限り原則として事前登録を利用することが可能であると述べている。しかし、二次データ分析の事前登録には、自身が同じデータセットを使って分析をしたことがある場合や他の研究者がそのデータセットを使って分析した研究を見たことがある場合にバイアスが生じることや、多くの事前登録のフォーマットが二次データ分析研究に合わないこと、分析の柔軟性が損なわれて論文の発表や引用の獲得に支障をきたす懸念があることなど、特有の困難さがある (Baldwin, Pingault, Schoeler, Sallis, & Munafò, 2022)。Baldwin et al. (2022) は、それらの困難さを指摘したうえで、データセットへのアクセス記録の報告や二次データ分析用のテンプレートの使用 (van den Akker et al., 2021⁹⁾)、データセットを探索用のデータと検証用に分けるなどの解決策を紹介している。

おわりに

先に述べたように、事前登録の最も大きな意義は、信頼ある研究結果として読者に自身の研究を伝えられる点にある。よくある誤解として「事前登録をしたら登録外の分析ができなくなる」というものがあるが、探索的分析ができなくなるわけではない。探索的分析の結果を研究の結論にしない¹⁰⁾ことと、探索的分析であることを明記することを守れば報告できる。大切なのは、どこまでが事前に予測されていた結果なのか、どこからは事後に考えたことなのかを明示することである。

事前登録の副次的なメリットとして、データ収集から分析、研究結果の公表までが非常にス

8) アーカイブ研究やメタ分析の文脈で書かれていたため文献を探す期間のことを意味していると思われるが、二次データの分析におけるデータセットを探す期間や分析をする期間と読み替えることも可能だろう。

9) van den Akker et al. (2021) は二次データ分析のためのチュートリアルであり、二次データ分析の事前登録をするうえでおいに参考になるだろう。

10) 前の研究の探索的分析の結果から考えた仮説を自身の次の研究で (事前登録をして) 行うこともできる。

ムーズに進んだ。たしかに、特にサンプルサイズ設計と分析プランの登録においてあらかじめ考えておかなければならないことが多かったが、何をどのように分析してどのような結果が出れば仮説が支持されたといえるのかがあらかじめ決めて分析用のコードを用意しておけば、その通りに分析をしてすぐに研究結果を報告できる。

【引用文献】

- Baldwin, J. R., Pingault, J. B., Schoeler, T., Sallis, H. M., & Munafò, M. R. (2022). Protecting against researcher bias in secondary data analysis: challenges and potential solutions. *European journal of epidemiology*, 37, 1-10.
- 長谷川龍樹・多田奏恵・米満文哉・池田鮎美・山田祐樹・高橋康介・近藤洋史 (2021). 実証的研究の事前登録の現状と実践——OSF 事前登録チュートリアル——. *心理学研究*, 92, 188-196.
- John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological science*, 23, 524-532.
- Li, X., Jusup, M., Wang, Z., Li, H., Shi, L., Podobnik, B., ... & Boccaletti, S. (2018). Punishment diminishes the benefits of network reciprocity in social dilemma experiments. *Proceedings of the National Academy of Sciences*, 115, 30-35.
- 三浦麻子 (2018). 心理学におけるオープンサイエンス-「統計革命」のインフラストラクチャー-. *心理学評論*, 61, 3-12.
- 村井潤一郎・橋本貴充 (2017). 心理学のためのサンプルサイズ設計入門 講談社.
- 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.
- 武藤拓之 (2022). データ・マテリアル・分析スクリプトのオープン化が拓く心理学の未来. *科学*, 92, 800-805.
- Nockur, L., Pfattheicher, S., & Keller, J. (2021). Different punishment systems in a public goods game with asymmetric endowments. *Journal of Experimental Social Psychology*, 93, 104096.
- Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, 115, 2600-2606.
- Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., ... & Vazire, S. (2022). Replicability, robustness, and reproducibility in psychological science. *Annual review of psychology*, 73, 719-748.
- 日本社会心理学会 (2022). オープンサイエンスバッジ申請 https://www.socialpsychology.jp/kitei/doc/open_science_badges_sinsei.rtf (2023年1月30日)
- 大久保衛重・岡田謙介 (2012). 伝えるための心理統計：効果量・信頼区間・検定力 勁草書房.
- Silberzahn, R., Uhlmann, E. L., Martin, D. P., Anselmi, P., Aust, F., Awtrey, E., ... & Nosek, B. A. (2018). Many analysts, one data set: Making transparent how variations in analytic choices affect results. *Advances in Methods and Practices in Psychological Science*, 1, 337-356.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2016). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. In A. E. Kazdin (Ed.), *Methodological issues and strategies in clinical research* (pp.547-555). American Psychological Association.
- Titlestad, K., Snijders, T. A. B., Durrheim, K., Quayle, M., & Postmes, T. (2019). The dynamic emergence of cooperative norms in a social dilemma. *Journal of Experimental Social Psychology*, 84, 103799.
- 宇佐美慧 (2011). 階層的なデータ収集デザインにおける2群の平均値差の検定・推定のためのサンプルサイズ決定法と数表の作成-検定力および効果量の信頼区間の観点から-. *教育心理学研究*, 59,

385-401.

van den Akker, O. R., Weston, S., Campbell, L., Chopik, B., Damian, R., Davis-Kean, P., ... & Bakker, M. (2021). Preregistration of secondary data analysis: A template and tutorial. *Meta-psychology*, 5.

Van't Veer, A. E., & Giner-Sorolla, R. (2016). Pre-registration in social psychology – A discussion and suggested template. *Journal of experimental social psychology*, 67, 2-12.