

心理学研究の「常識」が変わる？

—— 心理学界における再現可能性問題への取り組み

関西学院大学文学部総合心理科学科 教授

三浦麻子 (みうら あさこ)

Profile—三浦麻子

1992年、大阪大学人間科学部卒業。1995年、同大学院博士後期課程中退。同大助手、神戸学院大学准教授などを経て、現職。博士（人間科学）。専門は社会心理学。特に、コミュニケーションやインタラクションが新しい「何か」を生み出すメカニズムを解明することに興味を持ち、大規模ログデータの解析、質問紙調査、実験室実験など、様々なアプローチによる実証的研究を行なっている。



「再現可能性」問題と心理学

著名な物理学者であり随筆家としても知られた中谷宇吉郎は、その著書『科学の方法』（1958）で「科学は再現の可能な問題に適用範囲に限られる」と述べた。ありとあらゆる科学（を標榜する学問）にとって、ある知見が信頼に足るものかどうかを確認するためには、一度得られた実験結果について必ずその再現可能性を検証する必要があることは自明である。このことは、2014年初頭以来世間を騒がせている「STAP細胞問題」によって、決して望ましい形だったとは言えないが、一般にも広く知られることとなった。

心理学は、経験科学的なアプローチによって心理、行動のモデル化を試みる学問である。特に、方法論的な意味での行動主義（行動を内的な心理状態のあらわれとみなす考え方）と統計学的検定を主軸として、多くの知見を積み上げてきた。これらのアプローチ、特に統計学的検定は、常に偽陽性（第1種の過誤；たとえば、誤って「差がある」と結論づけてしまうこと）と偽陰性（第2種の過誤；たとえば、誤って「差がない」と結論づけてしまうこと）の可能性をはらんでいるため、その結果の信頼性を評価する際は直接的な追試（同じ刺激や手続きを用いて別のサンプルを対象として実施する実験）が必須である。しかし、それは心理学がまだ「若い」学問であるからかもしれないが、常

にオリジナリティを、つまり新しくユニークな発見を目指そうとする強い推進力に比べると、学問領域としての地歩を固めるための知見の再現可能性検証の試みは（少なくとも相対的には）軽視されてきた。

ひたすらにオリジナリティを追い求めることは、諸刃の剣である。たとえば、人間が必ずしも常に合理的な思考ができるわけではないことを示す「認知的バイアス」に関する心理学の知見は、経済学における古典的な人間観の見直しにつながるような影響力をもち、新たな研究の潮流を作り出すきっかけとなった。しかしその一方で、過去を振り返らず進みつづけてきたことが心理学研究の信頼性の著しい低下につながる出来事が相次ぎ、その科学性は今や危機的状態にあるとあって過言ではない。近年になって「心理学の科学性」を揺るがすような大きな問題が続けざまに明らかとなった。その典型例が、社会的認知研究を「リード」してきたオランダの社会心理学者ディエデリック・スタベルによるデータ捏造の発覚である。ある研究領域のリーディングリサーチャーと目されてきたような研究者たちによる科学的に不適切な行為は、行為それ自体の責任は研究者本人に帰せられるべきものであるとはいえ、再現可能性の検証が（特に、当事者以外の研究者によって）適切になされてこなかったことが、それを長年にわたって看過する原因となったことは否めな

い。また、自己知覚理論で高名な社会心理学者ダリル・ベムによる「超能力」の実証実験論文がトップジャーナル『*Journal of Personality and Social Psychology*』に掲載された一件（本誌61号「超能力ってどうなのでしょう？」参照）や、社会的認知研究の世界的権威であるジョン・バージとノーベル経済学賞受賞者のダニエル・カーネマンの間で起こった「社会的プライミング論争」（カーネマンが社会的プライミング効果の实在に疑念を呈し、公開質問状によってバージに実験の手続きを公開するよう求めたが、バージがこれに応じなかった）も、一つの研究パラダイムを複数の研究者（グループ）によって実施し、忠実な追試の蓄積によって知見の信頼性を高めることの必要性を、われわれに改めて認識させた。

しかし、事態改善の必要性の認識がすぐさまそれを目指す行動の実践につながるとは限らないことを、われわれは経験としても、そして皮肉にも「心理学的知見」としてもよく知っている。そこにはおそらく長年にわたり、「すべき」と考えていても「できない」と判断せざるを得ない社会的ジレンマ状況があった。誰しも、再現可能性を検証することが「科学が科学たり得る」ための必須条件であることはよく知っている。しかし、原典の「オリジナリティ」を尊重しなければならない追試はその精度が高いほど新規性と独自性を失うために、追試をした研究者自身の業績として認められにくい。つまり、研究者が大好きな、自らの「オリジナリティ」成分が絶対的に不足している。また、出版バイアス（「主流」とされる理論やモデルに対して確証的な知見が出版されやすい一方で、それに反証を挙げる内容の研究は出版されにくい傾向）も、再現可能性検証の公平性を妨げてきた。人間誰しも、好きこのんで損な勝負はしたくないだろう。もちろん、勇者が誰一人いなかったわけではない。しかし、勇者による数少ない追試の例も、それぞれがそれぞれなりに追試対象を選び、それぞれなりの方法で実験をしてデータをとり、「結果が再現された（されなかった）」という知見をそれぞれが提出するも

のがほとんどであった。統一され標準化された刺激や手続きによらない実験には、それによる不可避の測定誤差が生じる。そうした誤差を含んだデータによる再現性の検証は、決して精度の高いものとはならない。また、散在する知見をメタ分析などの統計的手法を用いて統合する試みはあっても、統一され標準化された刺激や手続きによる追試が組織的に行われるようなことはなかった。

心理学界の取り組み

しかし近年噴出した数々の問題事例は、遂にこうしたジレンマ状況で「追試は重要だ。しかし決してやりたくはない」と座して動かぬ選択肢をとりつづけるわけにはいかない、という方向に研究者たちを突き動かした。Association for Psychological Science (APS) は、機関誌『*Perspectives on Psychological Science*』の2012年11月号で実験結果の再現可能性に関する特集を組み、この問題の重大性を広く世界的に周知した。そしてAPSは実際に行動を起こし、同誌に新たな論文カテゴリ「Registered Replication Reports」を新設して、所定の手続きに則って行われた追試を積極的に掲載する方針を明らかにした（図1）。学会主導で追試による再現可能性検証を奨励し、その取り組みを成果として認める姿勢に「転向」したのである。この動きには、『*Journal of Experimental Social Psychology*』誌が同種



図1 APSによる「Registered Replication Reports」

の特集号を企画するなど他誌も追随し、「論文になりにくい」最たるものであった追試が一躍脚光を浴びることになった。

前述したとおり、精度の高い追試において必要なのは、共通フォーマットに基づく大規模で組織的な実施による質が均一かつ高いデータの収集であり、それを保証するための仕組みが事前登録 (Pre-registration) システムである。今般の追試への取り組みにおいて特に強調されるのは、追試対象の選定から、その実施方法の詳細までを事前に登録することの必要性である。その背景には、心理学の科学性を危機に陥らせているもう一方の問題、すなわち疑問符のつく研究実践 (Questionable Research Practice; QRP) の問題がある。詳しくは本特集「心理学な心理学研究」に譲るが、心理学研究においてはこれまで、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数々のQRPが行われ、それが知見の頑健さを著しく毀損していることが指摘されてきた。追試においても同じ轍を踏んではならない。事前登録システムによってこうしたQRPを阻止し、再現された (もしくは、されなかった) ケースばかりが報告される歪みを回避する必要がある。

事前登録システムを導入した追試を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のが「Reproducibility Project: Psychology」や「Many Labs」といったプロジェクトである。われわれの研究環境がデジタル/オンライン/ネットワーク化されたことが、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を可能にさせている。たとえば、現在進行中のプロジェクトの一つ「Many Labs 2」のWebサイト (<https://osf.io/8cd4r/>) にアクセスすると、そこでどのような形で活動が展開されているかが一目で分かる。図2はその一部である。定められた手続きに則って事前登録された研究がずらりと並び、それに関する刺激や教示文など、あらゆる手続きに関する資料が提供されている。

追試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、それらはすべてプロジェクトで共有される。

既に成果を論文として公表したプロジェクトとして「Many Labs」を紹介しよう。参加メンバーはインターネット上で公募され、36 (うちアメリカ25のほかオランダ、ブラジル、マレーシアなど) の研究機関が参加した (Klein et al., 2014)。古典的な研究 (最も古いのは1936年刊行の「情報の認知や評価に情報源が大きな影響を与える」ことを示した研究) からごく最近の研究 (最も新しいのは2013年刊行の「金銭への単純接触が現在の社会システムへの是認を高める」効果を示した研究) まで、事前登録された13の研究で得られた効果について、36の研究機関で独立したサンプルによる追試が行われ、参加者総数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である。結果は図3のようにまとめられている。グラフ中の×印が、各研究の原典で得られた効果量 d (当該データにおける実験的操作の効果の大きさ) であり、○や●印がそれぞれの追試で得られた効果量である。○や●印に比し

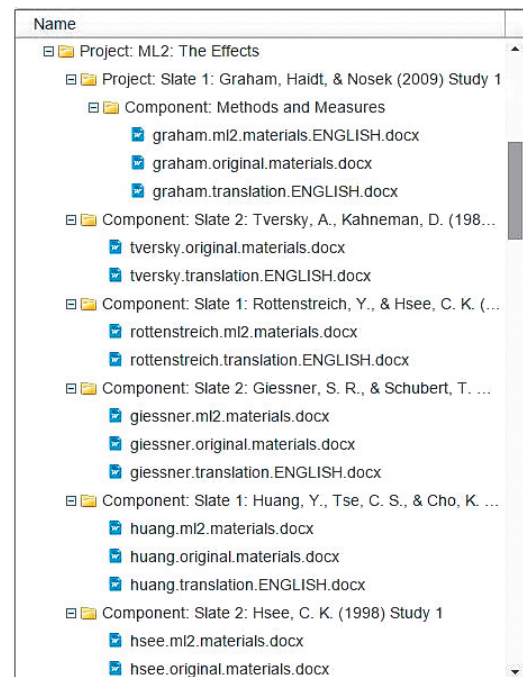


図2 Many Labs 2に事前登録された研究とその資料

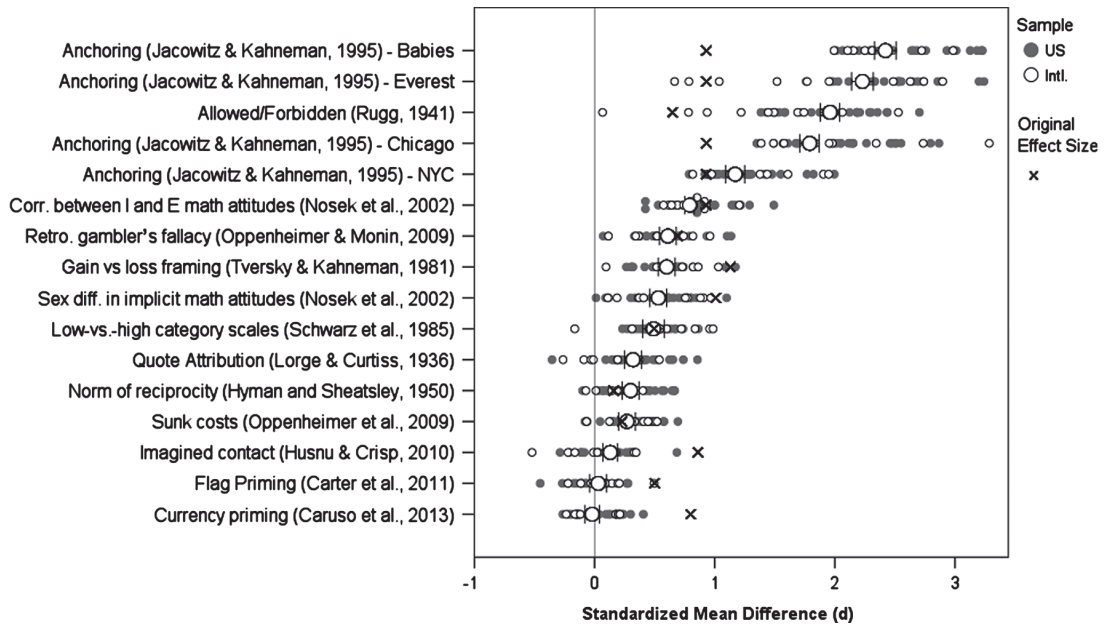


図3 Many Labs による再現可能性検証結果 (Klein et al., 2014)

て×印が右側（大きな値）に偏っているほど、追試では実験的操作の効果が原典ほどは見られなかったことを、すなわち再現性に疑問符がつくことを示している。今現在も、こうした取り組みが着々と、目に見える形で進行している。

心理学研究の「パラダイムシフト」

本稿では、心理学において実験結果の再現可能性検証が喫緊の問題としてクローズアップされてきた過程と、それに対応する学界の取り組みについて述べてきた。前述したとおり、精度の高い追試の実施にまずもって必要なのは、追試者は原典（オリジナル）に忠実であること、すなわち自らのオリジナリティを追求しないことである。われわれ心理学者は、その成長過程において「オリジナリティの高い研究を目指せ」という精神をたたき込まれ、それをほとんど唯一の行動規範としてきた。そのことと再現可能性の検証への取り組みとの間に大きな葛藤が生じることは想像に難くない。その意味で、心理学者には今、ある種のパラダイムシフトが求められている。誰にとっても、これまでとは違った世界の探索に向けて一歩足を踏み出すことには勇気がある。しかし今が昔と違うのは、

既にそれが科学としての心理学にとって有為であることを保証する地固めがなされており、今すぐにでもそれに参加することができる環境が準備されていることである。

心理学者に必須のスキルの一つとして、追試による再現可能性の検証が着実に進めることが求められるようになるのも、そう遠い将来ではないかもしれない。

謝辞

本稿は、平石界（安田女子大学）、池田功毅（中京大学）、藤島喜嗣（昭和女子大学）、樋口匡貴（上智大学）、平井啓（大阪大学）の各氏との議論に基づくものである。記して謝意を表す。

文献

中谷宇吉郎（1958）『科学の方法』岩波書店
 Klein, R.A. et al. (2014) Investigating variation in replicability: A “Many Labs” replication project. *Social Psychology*, 45, 142-152.