



心理学の再現性危機

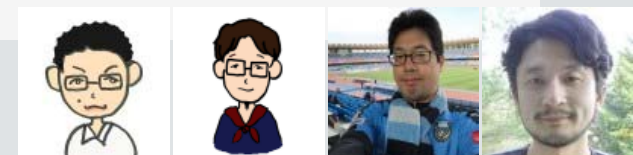
三浦麻子
(関西学院大学・大阪大学)

日本での取り組み

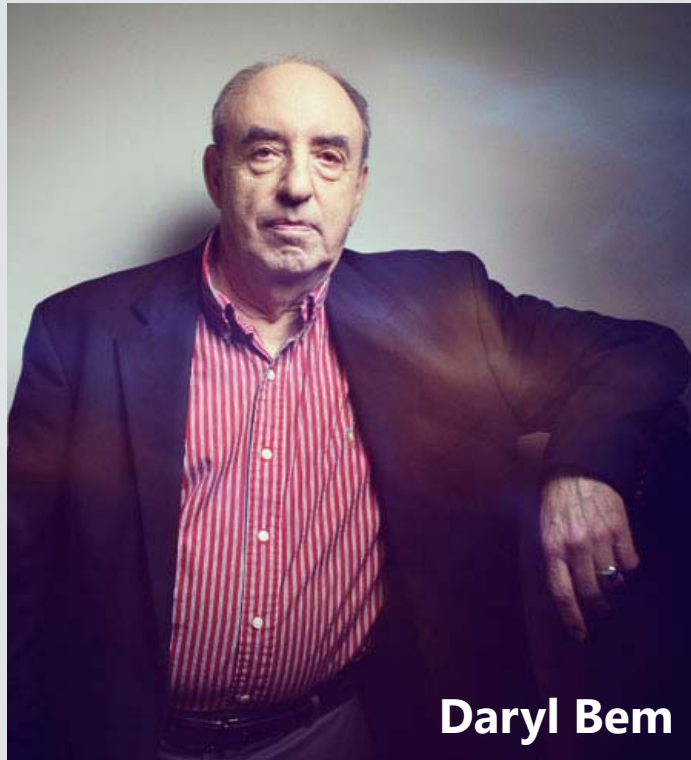
社会心理学研究の再現可能性検証のための日本拠点構築

研究課題/領域番号	15K13122
研究種目	挑戦的萌芽研究
配分区分	基金
研究分野	社会心理学
研究機関	関西学院大学
研究代表者	三浦 麻子 関西学院大学, 文学部, 教授 (30273569)
研究分担者	平石 界 慶應義塾大学, 文学部(三田), 准教授 (50343108) 樋口 匡貴 上智大学, 総合人間科学部, 准教授 (60352093) 藤島 喜嗣 昭和女子大学, 生活機構研究科, 准教授 (80349125)
研究期間 (年度)	2015-04-01 - 2018-03-31

<https://sites.google.com/view/asarin1003/replicability>



心理学における再現性問題 きっかけ その①



Daryl Bem

Journal of Personality and Social Psychology
2011, Vol. 100, No. 3, 407–425

© 2011 American Psychological Association
0022-3514/11/\$12.00 DOI: 10.1037/a0021524

Feeling the Future: Experimental Evidence for Anomalous Retroactive Influences on Cognition and Affect

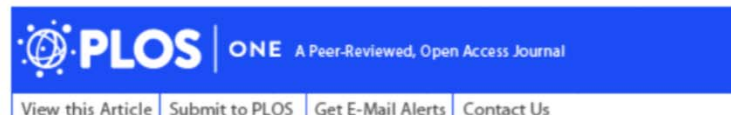
Daryl J. Bem
Cornell University

The term *psi* denotes anomalous processes of information or energy transfer that are currently unexplained in terms of known physical or biological mechanisms. Two variants of *psi* are *precognition* (conscious cognitive awareness) and *premonition* (affective apprehension) of a future event that could not otherwise be anticipated through any known inferential process. Precognition and premonition are themselves special cases of a more general phenomenon: the anomalous retroactive influence of some future event on an individual's current responses, whether those responses are conscious or nonconscious, cognitive or affective. This article reports 9 experiments, involving more than 1,000 participants, that test for retroactive influence by "time-reversing" well-established psychological effects so that the individual's responses are obtained before the putatively causal stimulus events occur. Data are presented for 4 time-reversed effects: precognitive approach to erotic stimuli and precognitive avoidance of negative stimuli; retroactive priming; retroactive habituation; and retroactive facilitation of recall. The mean effect size (*d*) in *psi* performance across all 9 experiments was 0.22, and all but one of the experiments yielded statistically significant results. The individual-difference variable of stimulus seeking, a component of extraversion, was significantly correlated with *psi* performance in 5 of the experiments, with participants who scored above the midpoint on a scale of stimulus seeking achieving a mean effect size of 0.43. Skepticism about *psi*, issues of replication, and theories of *psi* are also discussed.

Keywords: *psi*, parapsychology, ESP, precognition, retrocausation

心理学のトップジャーナル(JPSP)に、
かつて大理論を打ち立てた著名な社会心理学者による、
「人間には予知能力がある」という論文が掲載された。

きっかけ その①



PLoS One. 2012; 7(3): e33423.

Published online 2012 Mar 14. doi: [10.1371/journal.pone.0033423](https://doi.org/10.1371/journal.pone.0033423)

PMCID: PMC3303812

PMID: [22432019](https://pubmed.ncbi.nlm.nih.gov/22432019/)

Failing the Future: Three Unsuccessful Attempts to Replicate Bem's 'Retroactive Facilitation of Recall' Effect

Stuart J. Ritchie,^{1,*} Richard Wiseman,² and Christopher C. French³

Sam Gilbert, Editor

[Author information](#) ► [Article notes](#) ► [Copyright and License information](#) ► [Disclaimer](#)

This article has been [cited by](#) other articles in PMC.

Abstract

Nine recently reported parapsychological experiments appear to support describe three pre-registered independent attempts to exactly replicate or 'retroactive facilitation of recall', which examines whether performance influenced by a post-test exercise. All three replication attempts failed to (combined $n=150$; combined $p=.83$, one-tailed) and thus do not support

Journal of Personality and Social Psychology
2012, Vol. 103, No. 6, 933–948

© 2012 American Psychological Association
0022-3514/12/\$12.00 DOI: [10.1037/a0029709](https://doi.org/10.1037/a0029709)

Correcting the Past: Failures to Replicate Psi

Jeff Galak
Carnegie Mellon University

Robyn A. LeBoeuf
University of Florida

Leif D. Nelson
University of California, Berkeley

Joseph P. Simmons
University of Pennsylvania

Across 7 experiments ($N = 3,289$), we replicate the procedure of Experiments 8 and 9 from Bem (2011), which had originally demonstrated retroactive facilitation of recall. We failed to replicate that finding. We further conduct a meta-analysis of all replication attempts of these experiments and find that the average effect size ($d = 0.04$) is no different from 0. We discuss some reasons for differences between the results in this article and those presented in Bem (2011).

Keywords: psi, precognition, ESP, researcher degrees of freedom, meta-analysis

追試による再現性検証論文
（再現に失敗）をJPSPが「追
試は掲載しない」という理由
でリジェクトし大騒ぎに

↓
結果的に別の研究者らによる
再現性検証論文（再現に失
敗）がJPSPに掲載された

心理学における再現性問題 きっかけ その②



Diederik Stapel

SCIENCE NEWS NOVEMBER 2, 2011 / 8:41 PM / 7 YEARS AGO

Dutch psychologist admits he made up research data

Kate Kelland

3 MIN READ



LONDON (Reuters) - A Dutch psychologist has admitted making up data and faking research over many years in studies which were then published in peer-reviewed scientific journals.

Diederik Stapel, a psychologist working at Tilburg University in the Netherlands, said he had “failed as a scientist” and was ashamed of what he had done, but had been driven to falsifying research by constant pressure to perform.

現役バリバリで活躍していた著名な社会心理学者が、
10年にわたりデータのねつ造・改ざんをしていたことがわかり、
論文は取り下げとなり、大学は解雇された。

きっかけ その②

RETRACTION

Retraction

D. A. Stapel, S. Lindenberg¹

+ See all authors and affiliations

Science 02 Dec 2011:
Vol. 334, Issue 6060, pp. 1202
DOI: 10.1126/science.334.6060.1202-a

Article

Info & Metrics

eLetters



Our Report “Coping with chaos: How disordered contexts promote stereotyping and discrimination” (1) reported the effects of the physical environment on human stereotyping and discriminatory behavior.

On 31 October 2011, Tilburg University held a press conference to announce findings of its investigation into possible data fraud on the part of author Stapel. These findings of the university's interim report (2) included fabrication of data in this *Science* paper. Therefore, we are retracting the paper, with apologies from author Stapel. Coauthor Lindenberg was in no way involved in the generation of the data and agrees to the retraction of the paper.



Stapelは後に自身が繰り返した不正を詳細に記述し、その時の心境を明かした書籍を出版している

ダメ押し

RESEARCH ARTICLE

Estimating the reproducibility of psychological science

Open Science Collaboration^{*,†}

+ See all authors and affiliations

Science 28 Aug 2015:
Vol. 349, Issue 6251, aac4716
DOI: 10.1126/science.aac4716

[Article](#)[Figures & Data](#)[Info & Metrics](#)[eLetters](#)[PDF](#)

You are currently viewing the abstract.

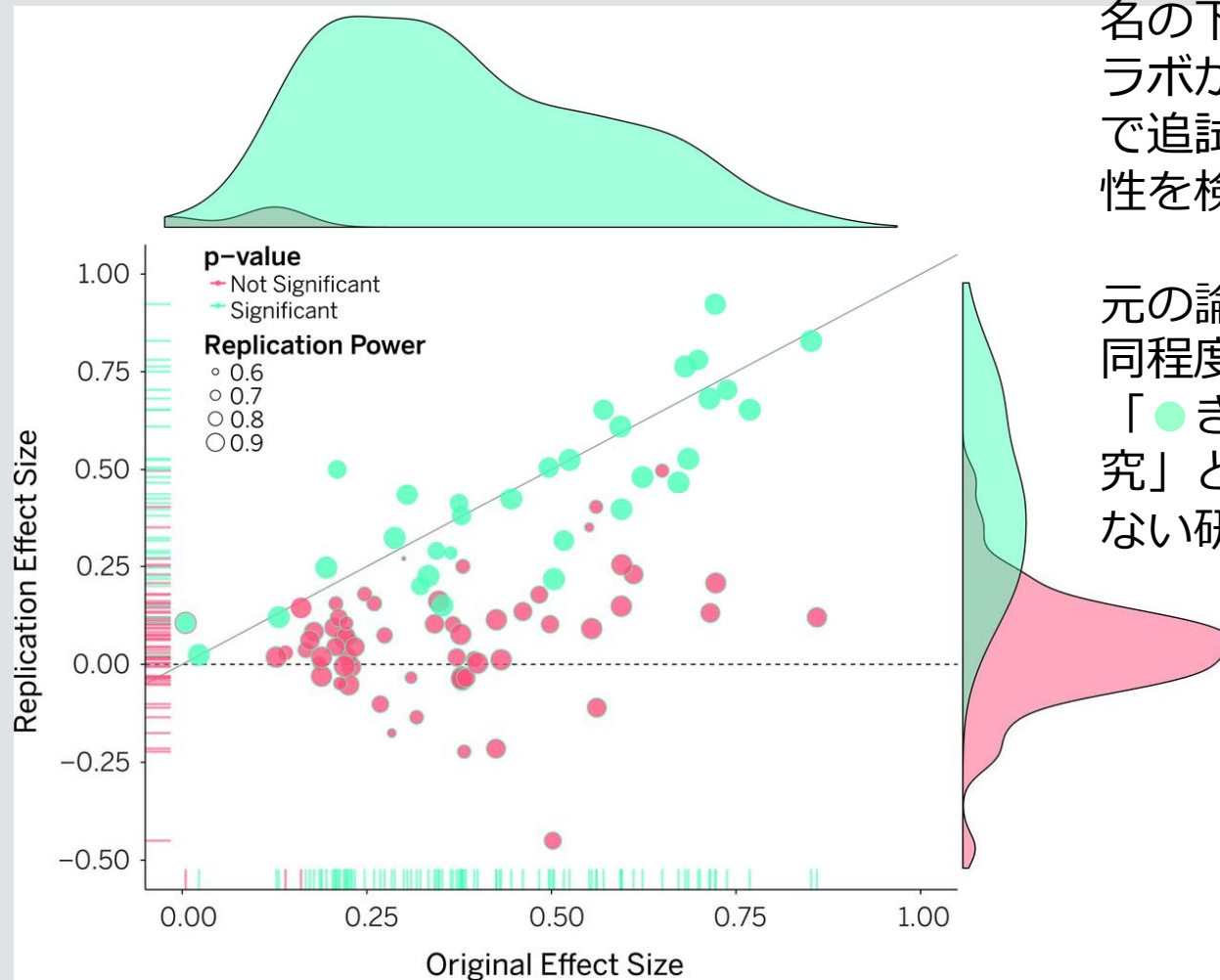
[View Full Text](#)

Empirically analyzing empirical evidence

One of the central goals in any scientific endeavor is to understand causality. Experiments that seek to demonstrate a cause/effect relation most often manipulate the postulated causal factor. Aarts et al. describe the replication of 100 experiments reported in papers published in 2008 in three high-ranking psychology journals. Assessing whether the replication and the original experiment yielded the same result according to several criteria, they find that about one-third to one-half of the original findings were also observed in the replication study.

Science, this issue [10.1126/science.aac4716](#)

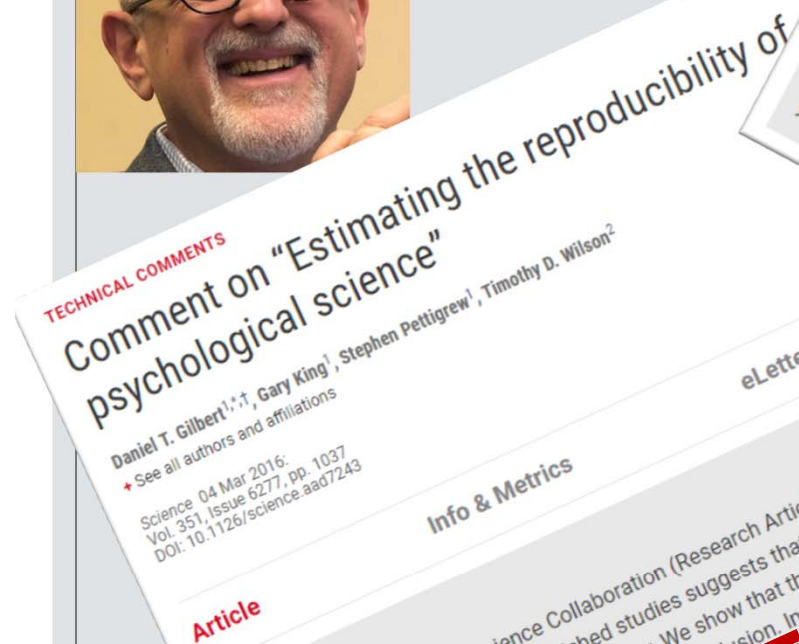
ダメ押し



Open Science Collaborationの
名の下に、世界各国の多くの
ラボが協力して、同じ手続き
で追試データを収集し、再現
性を検証

元の論文の効果量に関わらず、
同程度の効果量を得られる
「●きちんと再現できる研
究」と「●まったく再現でき
ない研究」に二分される印象

そして、百家争鳴



Abstract

A paper from the Open Science Collaboration (Research Articles, 28 August 2015, aac4716) attempting to replicate 100 published studies suggests that the reproducibility of psychological science is surprisingly low. We show that this article contains three statistical errors and provides no support for such a conclusion. Indeed, the data are consistent with the opposite conclusion, namely, that the reproducibility of psychological science is quite high.

Info & Metrics

eLetters

Article

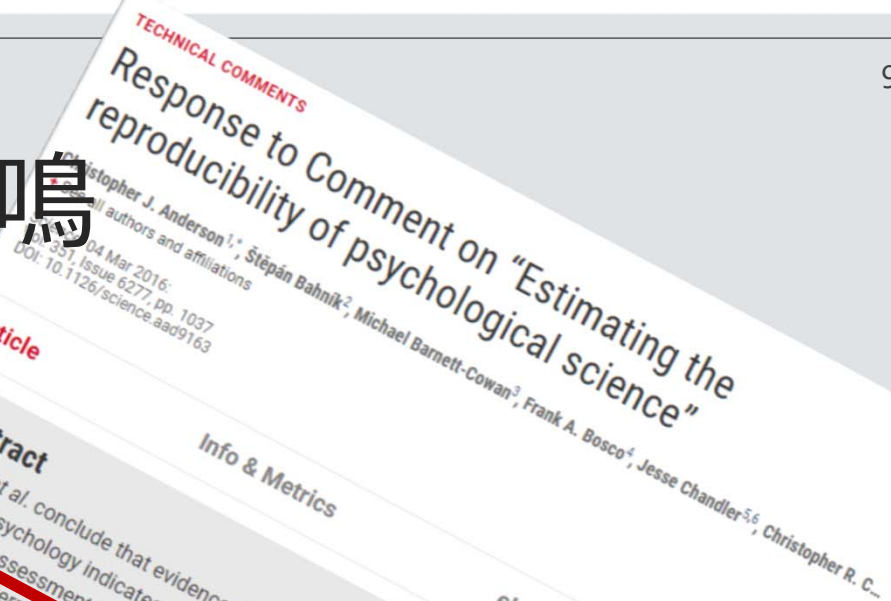
Abstract

Gilbert et al. conclude that evidence from the Open Science Collaboration's Reproducibility Project: Psychology indicates high reproducibility, given the study methodology. Their very optimistic assessment is limited by statistical misconceptions and by causal inferences from selectively interpreted, correlational data. Using the Reproducibility Project: Psychology data, both optimistic and pessimistic conclusions about reproducibility are possible, and neither are yet warranted.

Info & Metrics

eLetters

PDF



あの有名な実験も...!?

<https://www.nytimes.com/2018/07/16/health/psychology-studies-stanford-prison.html>

MIND

Psychology Itself Is Under Scrutiny

Many famous studies of human behavior cannot be reproduced.
Even so, they revealed aspects of our inner lives that feel true.

scrutiny: 監視



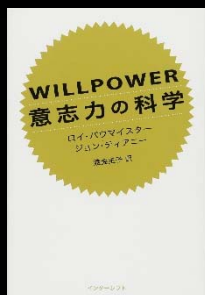
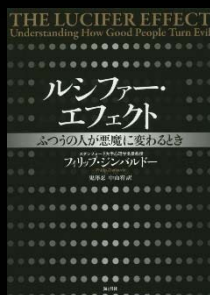
あの有名な実験も...!?

<https://www.nytimes.com/2018/07/16/health/psychology-studies-stanford-prison.html>

再現性に疑問符がついている
著名な研究として、

スタンフォード監獄実験
マシュマロ・テスト
自我消耗 ego depletion

を紹介



心理学の再現性問題

- なぜ危機的状況にあるのか?
- 本質的な問題
 - 「強い理論」が少ない
 - 一方で（だからこそ）脆弱な「理論」が多く提案される
- 本質的ではない「はずの」問題
 - 不正をしたくなる土壌
 - 不正ができてしまう土壌
 - 抗うことが難しい「認知的バイアス」

心理学の再現性問題

- なぜ危機的状況にあるのか？

- **本質的な問題**

- 「強い理論」が（少）ない
 - 一方で（だからこそ）脆弱な「理論」が多く提案される
- 本質的ではない「はずの」問題
 - 不正をしたくなる土壌
 - 不正ができてしまう土壌
 - 抗うことが難しい「認知的バイアス」



早く科学に
なりたーい
問題

強い理論が（少）ない

Eysenckによれば, Newtonの重力理論やEinsteinの相対性理論などのような「強い」理論では, 頑健な実験結果等の根拠に基づく多くの知見が仮説検証の前提となっており, さらに各仮説間には強い相互依存性があるため, ひとつの仮説Aは, 他の仮説Bや前提となる知見による強い制限を受けており, そのため厳密な予測を事前に行うことが可能である。それに対して, 心理学で見られるような「弱い」理論では, 前提となる知見の根拠が乏しく, また知見ないしは仮説間の相互依存性も少ないため, 厳密な事前の予測が難しい。そうした状況の下では, 仮に研究者がデータを見た後にその分析方法を検討すると, 与えられたデータに合致する予測と, それに適した分析を, 知らず知らずのうちに選んでしまう可能性が高まると考えられる。

池田・平石 (2016)

審美的基準の重視

- 結果の一貫性
 - 1論文の中で得られる結果は一貫しているのが美しい
- 物語性
 - 仮説の設定がされ、それがデータによって検証されるのが美しい
- 新奇性
 - 仮説は新奇性を伴っているものが美しい

Giner-Sorolla(2012)

**科学はかくあるべし、的な発想は強い
当然、再現性の検証は軽視されやすい**

心理学の再現性問題

- なぜ危機的状況にあるのか?
- 本質的な問題
 - 「強い理論」が少ない
 - 一方で（だからこそ）脆弱な「理論」が多く提案される

- **本質的ではない「はずの」問題**
 - 不正をしたくなる土壌
 - 不正ができてしまう土壌
 - 抗うことが難しい「認知的バイアス」

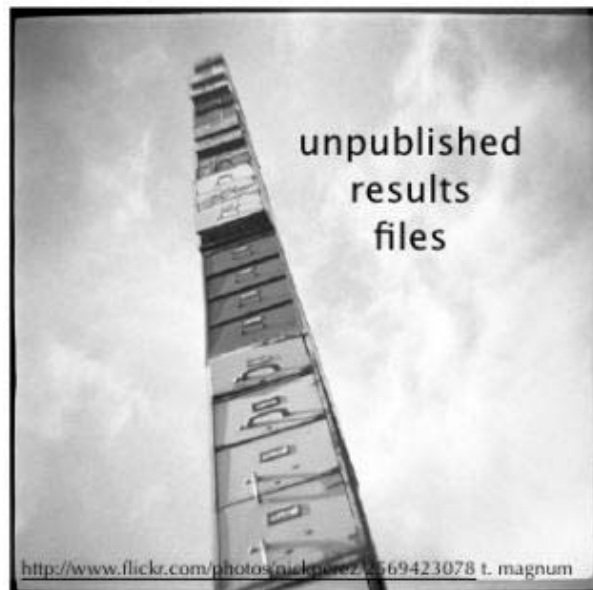


にんげん
だもの
問題

不正をしたくなる土壌

The File-Drawer Problem

「お蔵入り」問題



- Difficult to publish non-replications and replications

- Most journals only publish papers that “make a novel contribution”

- Reviewers/editors tend to hold non-replicating manuscript to higher standard than original.



- Little career incentive to publish a non-replication or a replication

<https://www.slideshare.net/dapthorp/file-drawer-talk>

不正をしたくなる土壌

- 「お蔵入り」問題
 - 審美的基準の重視により、仮説が支持された研究でなければ公刊されにくいので論文にしない→公刊されるのは仮説が支持された研究ばかりになる→仮説が支持されないデータは都合が悪い...という発想に辿り着いてしまう

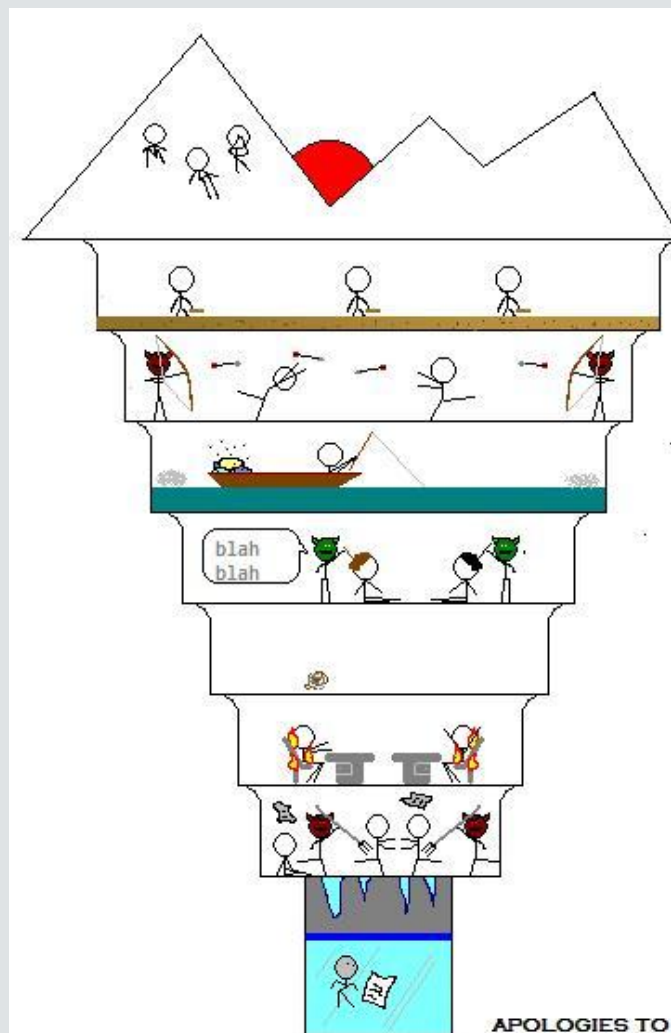
Table 3 Cross-tabulation between statistical results of TESS studies and their publication status (column percentages reported).

Pearson χ^2 test of independence: $\chi^2 (6) = 80.3, P < 0.001$.

	Null (%)	Mixed (%)	Strong (%)
Not written	64.6	12.2	4.4
Written but not published	14.6	39.0	34.1
Published (non-top-tier)	10.4	37.8	38.4
Published (top-tier)	10.4	11.0	23.1
Total	100.0	100.0	100.0

Franco, Malhotra, & Simonovits (2014)

The Nine Circles of Scientific Hell



Questionable Research Practices (QRPs)

I Limbo	問題のあるやり方を見て見ぬふり
II Overselling	過大に自分の研究を売り込む
III Post-Hoc Storytelling	後づけで話をつくる
IV P-Value Fishing	こっそりと望むp値を手に入れる
V Creative Outliers	外れ値を都合のいいように使う
VI Plagiarism	剽窃・盗作
VII Non-Publication	都合の悪いデータを公開しない
VIII Partial Publication	都合のよいデータだけを選んで公開
IX Falsification	データの捏造

Neuroskeptic(2012)

不正だという自覚の薄い不正



Neuroskeptic(2012)

p-hacking

p 値ハッキングが行なわれがちなのは、当初想定した十分な数のデータをとった時点で分析をしてみると「仮説が支持されたと断言はできないが、まったく支持されなかったと言うのもどうかと思うような微妙な結果」、つまり p 値が 0.05 をわずかに上回るような結果が得られた場合である。こういう結果を報告すると、こんな「アドバイス」をくれる指導教員や先輩がいるかもしれない。

「じゃあ、もうちょっとデータ足してみたら？」

データを足すこと自体に問題があるとはいえないが、その目的が「仮説が支持された」という確証を得るためであれば問題である。アドバイスを得た学生が、データを 1 人分足しては検定をし、また足しては検定し、を繰り返したあげく、ある時点で p 値が 0.05 を下回ったとしよう。そのとき学生はどうするだろう。ここでデータ収集を打ち切ってしまうと、

「追加で 5 人参加してもらったらちゃんと有意になりました！」
「おおそうか、よかったよかった。やっと論文が書けるな！」

拙著『なるほど!心理学研究法』

第9章「研究倫理:研究者として「やってはいけないこと」」

データ（参加者）を1名ずつ足していったって、有意になったらそこでストップ、という方法で「偶然有意になる」結果を得る確率は「案外」高い

「これは、全体としては有意差が見られなくても、参加者のパーソナリティによっては差があったりするんじゃないか？」

という「アドバイス」もありがちな。これもその影響を検討すること自体に問題があるとはいえないが、その目的が「仮説が支持された」という確証を得るためであれば問題である。仮説として想定していたわけではないが、参加者のパーソナリティに関するデータを同時にとっておいて、その影響を加味した分析をすると、

「外向性による違いはなかったんですが、誠実性が高いと差がありました！」
「おおそうか、よかったよかった。やっと論文が書けるな！」

となる。「差を出す」と褒められるのだ。

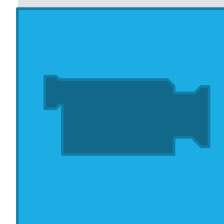
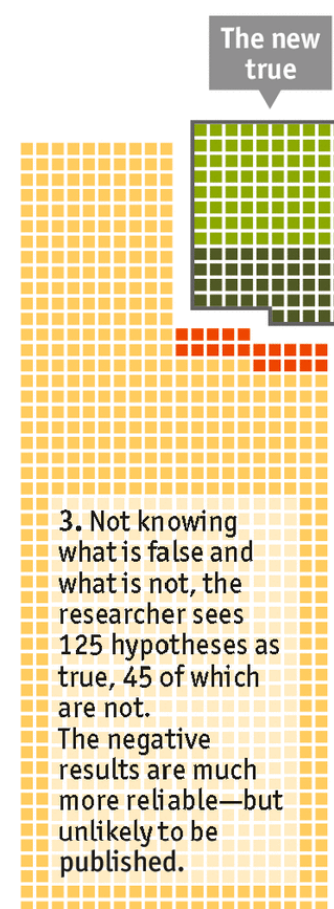
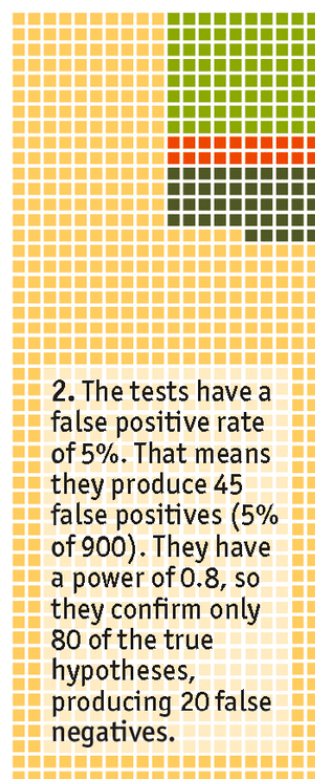
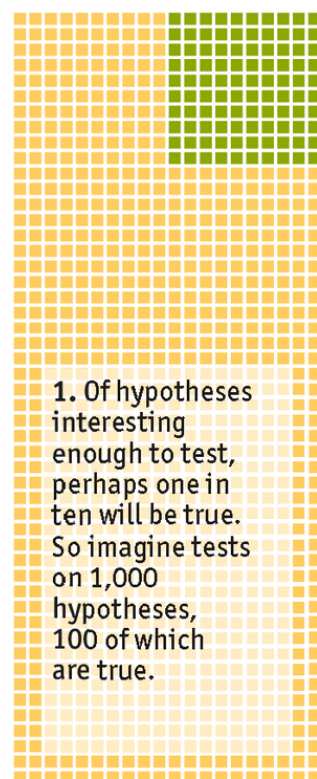
False-Positive (偽陽性)

本来は帰無仮説が真なのに
偽として棄却してしまう
(本来検証したい仮説を採
択してしまう) 誤り

Unlikely results

How a small proportion of false positives can prove very misleading

False True False negatives False positives



False-Positive (偽陽性)

Unlikely results

How a small proportion of false positives can prove very misleading

False True False negatives False positives

「何もごまかさなくても首尾よく仮説が支持された」としても、統計的仮説検定の枠組みに従う限り、それが「真」ではない確率はゼロにはならないし、なんと高ければそれは「案外」高い

1. Of hypotheses of interest, perhaps only 10% are true. So imagine tests on 1,000 hypotheses, 100 of which are true.

2. The tests have a false positive rate of 5%. That means they produce 45 false positives (5% of 900). They have a power of 0.8, so they confirm only 80 of the true hypotheses, producing 20 false negatives.

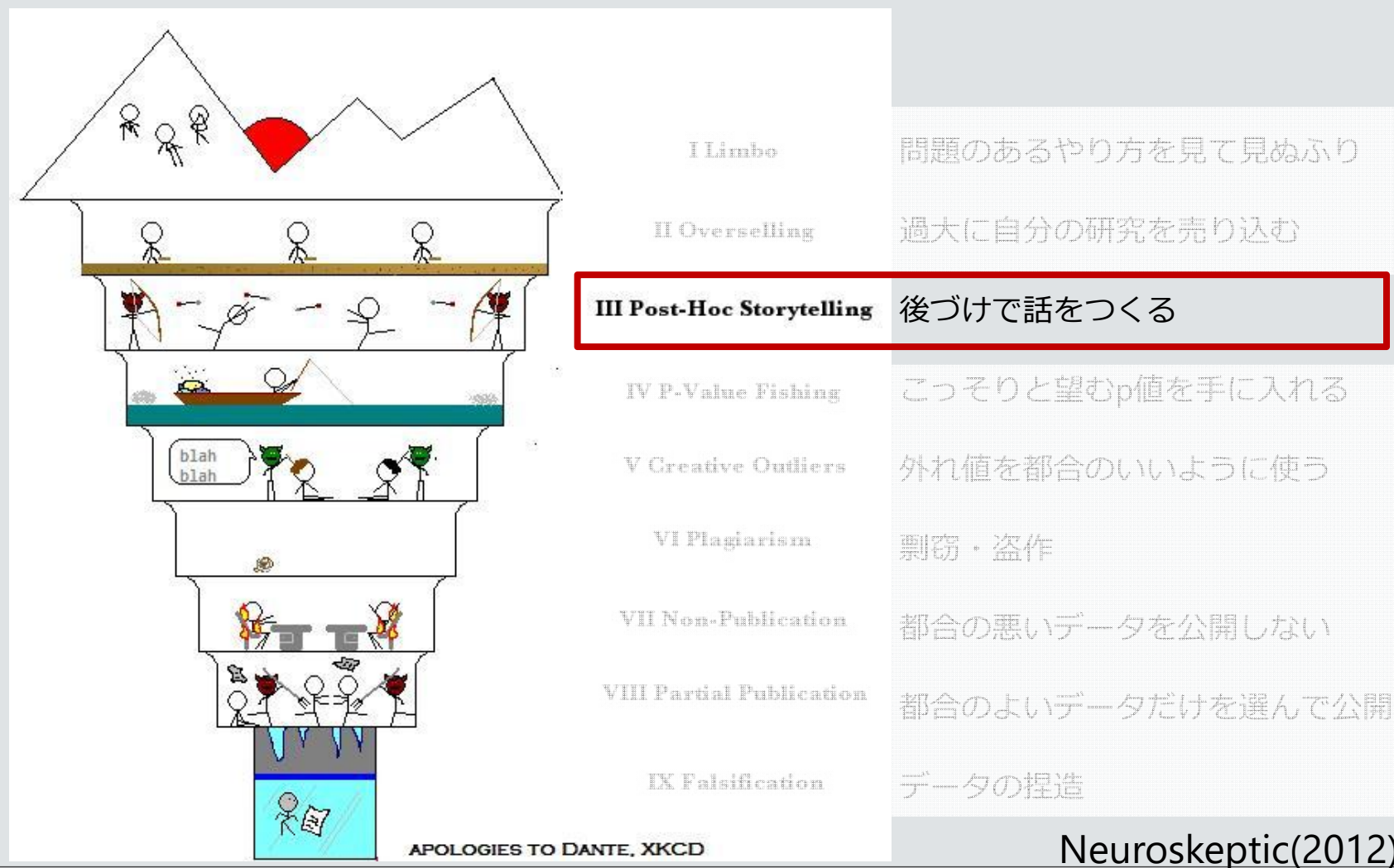
3. No matter what, what researchers see is 125 false positives, 80 true, and 20 false negatives. The result is more likely to be a false positive than a true result.

表1 様々な p 値、検定力、仮説が正しい確率において、有意な結果が偽陽性である確率

統計 検定力 (%)	有意となった結果が偽陽性である 確率 (%)		
	$p=.05$	$p=.01$	$p=.001$
仮説のうち 80% が正しい場合			
20	5.9	1.2	0.1
50	2.4	0.5	0.05
80	1.5	0.3	0.03
仮説のうち 50% が正しい場合			
20	20	4.8	0.5
50	9.1	2	0.2
80	5.9	1.2	0.1
仮説のうち 10% が正しい場合			
20	69.2	31	4.3
50	47.4	15.3	1.8
80	36	10.1	1.1

註：Sterne & Smith (2001), Table 3 の一部を改変して引用。

不正だという自覚の薄い不正



Neuroskeptic(2012)

HARKing

- **Hypothesizing After the Results are Known** 仮説の後付け
 - 出た結果に合わせて仮説を作り替える
- 心理学研究で対象とする場面には、実験的に操作したり（できたり）、そもそも主たる関心としていた独立変数以外にも、従属変数に影響しそうな剰余変数が（無数に）想定可能
- それらをあれこれとっておいて、事後分析で影響をもつことが分かった変数を、あたかも当初から主たる関心のひとつとしていたかのようにストーリーを再構築する
 - 探索的に検討したのに、あたかも仮説検証だったかのように...

HARKing

HARKingの弊害は様々に論じられているが、おそらく最大の問題点は、それが第一種の過誤を増大させてしまう点にあると思われる。仮に、得られたデータに対して、様々な共変量の投入などの分析を複数行ったとしよう。そして多重比較の修正を無視すれば、その分析のひとつで有意差が得られたとしよう。ここでHARKingを用いず、当初の理論的予測に忠実に従った報告を行うならば、多重比較問題の無視が明らかとなってしまうため、報告の信頼性は損われ、論文の出版につながらないかもしれない。ところがここでHARKingを使えば、当該の有意差が得られた検定が、あたかも当初から予測されていた唯一の分析であるかのように扱うことができるため、多重

比較の問題を「隠蔽」することが可能になる。さらに、心理学の弱い理論のもとでは、事前に厳密な理論的予測を立てることが難しいという事実を考えると、こうして「隠蔽」された多重比較を、論文報告のみから見抜くことは非常に難しいと思われる。このようにHARKingには、言わば「見かけの研究者の自由度」を低め、第一種の過誤の発生確率を高める危険性がある。

池田・平石(2016)

人間の認知的バイアス

- 人間の判断と意思決定は合理的選択理論とは異なった方法で行われている
 - Confirmation bias 確証バイアス
 - 仮説や信念を検証する際にそれを支持する情報ばかりを集め、反証する情報を無視または集めようとしない認知的バイアス
- False Positive**
- Hindsight bias 後知恵バイアス
 - 物事が起きてからそれが予測可能だったと考える認知的バイアス→**HARKing**



にんげん
だもの
問題

悪循環がもたらす再現性危機



池田・平石 (2016)

学会の対応

表2 頑健性確保のため各ジャーナル、団体が制定した新ガイドラインの例

	Society for Personality and Social Psychology (Funder et al., 2014)	Psychonomic Society (The Psychonomic Society, 2012)	Psychological Science (Eich, 2014)
統計報告	<ul style="list-style-type: none"> 事前の検定力推定とサンプルサイズ決定方法の報告 効果量と95%信頼区間の報告 	<ul style="list-style-type: none"> 事前の検定力推定とサンプルサイズ決定方法の報告 効果量と95%信頼区間の報告と重視 	<ul style="list-style-type: none"> サンプルサイズ決定方法の報告 効果量と95%信頼区間の報告とメタアナリシスの奨励
QRPsの禁止	<ul style="list-style-type: none"> 修正なしに同一データに対して複数の検定を行うことの禁止 データ分析の結果を見てデータ収集を止めることの禁止 データ分析の結果を見て観察、指標、項目、条件、参加者、行った実験などを除外して報告することの禁止 	<ul style="list-style-type: none"> 修正なしに同一データに対して複数の検定を行うことの禁止 データ分析の結果を見てデータ収集を止めることの禁止 データ分析の結果を見て観察、指標、項目、条件、参加者、行った実験などを除外して報告することの禁止 測定したすべての従属変数の報告 共変量分析は事前に計画し、探索的であると示す 	<ul style="list-style-type: none"> データ収集停止ルールの報告 測定したすべての独立変数、指標の報告 当該論文の研究目的のために分析したすべての従属変数の報告 観察を分析から除外した場合、理由とともにそれを示す
その他	<ul style="list-style-type: none"> 付録での教示等の正確な報告 データ公開 追試の奨励 柔軟で多様な研究評価軸の維持 	<ul style="list-style-type: none"> 帰無仮説が棄却されなかった時は第二種の過誤を疑う ネガティブな結果報告の奨励 ベイズ統計など帰無仮説検定以外の統計手法も受け入れる 	<ul style="list-style-type: none"> 論文の方法と結果の字数制限を撤廃 データ公開の奨励 マテリアル公開の奨励 事前登録の奨励

池田・平石 (2016)

学会の対応

表2 頑健性確保のため各ジャーナル、団体が制定した新ガイドラインの例

	Society for Personality and Social Psychology (Funder et al., 2014)	Psychonomic Society (The Psychonomic Society, 2012)	Psychological Science (Eich, 2014)
統計報告	<ul style="list-style-type: none"> 事前の検定力推定とサンプルサイズ決定方法の報告 効果量と95%信頼区間の報告 	<ul style="list-style-type: none"> 事前の検定力推定とサンプルサイズ決定方法の報告 効果量と95%信頼区間の報告と重視 	<ul style="list-style-type: none"> サンプルサイズ決定方法の報告 効果量と95%信頼区間の報告とメタアナリシスの奨励
QRPsの禁止	<ul style="list-style-type: none"> 修正なしに同一データに対して複数の検定を行うことの禁止 データ分析の結果を見てデータ収集を止めることの禁止 データ分析の結果を見て観察、指標、項目、条件、参加者、行った実験などを除外して報告することの禁止 	<ul style="list-style-type: none"> 修正なしに同一データに対して複数の検定を行うことの禁止 データ分析の結果を見てデータ収集を止めることの禁止 データ分析の結果を見て観察、指標、項目、条件、参加者、行った実験などを除外して報告することの禁止 測定したすべての従属変数の報告 共変量分析は事前に計画し、探索的であると示す 	<ul style="list-style-type: none"> データ収集停止ルールの報告 測定したすべての独立変数、指標の報告 当該論文の研究目的のために分析したすべての従属変数の報告 観察を分析から除外した場合、理由とともにそれを示す
その他	<ul style="list-style-type: none"> <u>付録での教示等の正確な報告</u> データ公開 <u>追試の奨励</u> 柔軟で多様な研究評価軸の維持 	<ul style="list-style-type: none"> 帰無仮説が棄却されなかった時は第二種の過誤を疑う ネガティブな結果報告の奨励 ベイズ統計など帰無仮説検定以外の統計手法も受け入れる 	<ul style="list-style-type: none"> 論文の方法と結果の字数制限を撤廃 データ公開の奨励 <u>マテリアル公開の奨励</u> 事前登録の奨励

池田・平石 (2016)

学会の対応(APS)



Registered Replication Reports

Multi-lab, high-quality replications of important experiments in psychological science along with comments by the authors of the original studies.

追試研究の奨励

RESEARCH ARTICLE

Estimating the reproducibility of psychological science

Open Science Collaboration^{*,†}

† See all authors and affiliations

Science 28 Aug 2015;
Vol. 349, Issue 6251, aac4716
DOI: 10.1126/science.aac4716

Article

Figures & Data

Info & Metrics

eLetters

PDF

You are currently viewing the abstract.

View Full Text

Empirically analyzing empirical evidence

One of the central goals in any scientific endeavor is to understand causality. Experiments that seek to demonstrate a cause/effect relation most often manipulate the postulated causal factor. Aarts *et al.* describe the replication of 100 experiments reported in papers published in 2008 in three high-ranking psychology journals. Assessing whether the replication and the original experiment yielded the same result according to several criteria, they find that about one-third to one-half of the original findings were also observed in the replication study.

Science, this issue 10.1126/science.aac4716

In total, 100 replications were completed by 270 contributing authors.

Reproducibility Project: Psychology

Public 45

Contributors: Christopher Jon Anderson, Joanna Anderson, Marcel A.L.M. van Assen, Peter Raymond Attridge, Angela Attwood, Jordan Axt, Molly Babel, Štěpán Bahník, Erica Baranski, Michael Barnett-Cowan, Elizabeth Bartmess, Jennifer Beer, Raoul Bell, Heather Bentley, Don van den Bergh, Leah Beyan, Bobby den Bezemer, Denny Borsboom, Jannik Bosch, Frank Bosco, Sara Bowman, Mark Brandt, Erin L. Braswell, Hilmar Brohmer, Benjamin T. Brown, Kristina Brown, Jovita Brüning, Ann Calhoun-Sauls, Shannon Callahan, Elizabeth Chagnon, Jesse J. Chandler, Christopher R. Chartier, Felix Cheung, Phuonguyen Chu, Linda Cillessen, Russ Clay, Hayley Cleary, Mark Cloud, Michael Cohn, Johanna Cohoon, Simon Columbus, Giulio Costantini, Leslie Cramblert Alvarez, Edward Cremata, Jan Crusius, Jamie DeCoster, Michelle DeGaetano, Nicolás Della Penna, Marie Deserno, Olivia Devitt, Laura Dewitte, Philip DiGiacomo, David Dobolyi, Geneva T. Dodson, Brent Donnellan, Ryan Donohue, Roel van Dooren, Shiny van Doorn, Rebecca A. Dore, Angela Rachael Dorrough, Annie te Dorsthorst, Anna Dreber Almenberg, Michelle Dugas, Elizabeth Dunn, Kayleigh Eassey, Sylvia Eboigbe, Jesse Eggleston, Jo Embrey, Sacha Eppkamp, Tim Errington, Vivien Estel, Frank J. Farach, Jenelle Feather, Anna Fedor, Belén Fernández-Castilla, Susann Fiedler, James G. Field, Tanika Fitneva, Taru Flagan, Amanda Forest, Eskil Forsell, Joshua D. Foster, Michael C. Frank, Rebecca S. Frazier, Heather Fuchs, Philip Gable, Jeff Galak, Elisa Maria Galliani, Juhap Gampa, Sara García, Douglas Gazarian, Elizabeth Gilbert, Roger Giner-Sorolla, Andreas Glöckner, Lars Goeliner, Jin X. Goh, Rebecca M. Goldberg, Stephen D. Goldering, Patrick T. Goodbourn, Shauna Gordon-McKeon, Bryan Gorges, Jessie Gorges, Justin Goss, Jesse Graham, Jeremy R. Gray, C.H.J. Hartgerink, Fred Hasselmann, Timothy B. Hayes, Emma Heikensten, Felix Henninger, Grace Hicks, John Hodson, Taylor Holubar, Gea Hoogendoorn, Marije van der Hulst, Denise Humphries, Cathy O. Y. Hung, Ithell Immanuel, Vanessa C. Irisk, Georg Jahn, Frank Jäkel, Marc Jekel, Magnus Johannesson, David J. Johnson, Kate Johnson-Grey, Larissa Johnson, William Johnston, Al Jonas, Jennifer Joy-Gaba, Heather Kappes, Kim Kelso, Mallory C. Kidwell, Seung Kyung Kim, Matthew Kirkhart, Bennett Kleinberg, Goran Knezevic, Franziska Maria Kolorz, Robert Wilhelm Krause, Job Krijnen, Tim Kuhlmann, Yoram Kevin Kunkels, Megan Kyic, Calvin Lai, Amir Laique, Daniel Lakens, Kristin Lane, Bethany Lassettter, Lili Lazarevic, Stéphanie P. LeBel, Key Jung Lee, Minha Lee, Kristi Lemm, Carmel Levitan, Melissa Lewis, Lin Lin, Stephanie Lin, Matthias Lippold, Darren Loureiro, Daniel Luman, Ili Luteijn, Sean Mackinnon, Heather N. Mainard, Denise Marigold, Dan Martin, Tyler Martinez, E.J. Mascampo, Josh Mataric, Psy D., CAHIMS, Maya Mathur, Michael May, Kateri McRae, Todd McElroy, Nicole Mechin, Pranjai Mehta, Johannes Meixner, Alissa Melinger, Jeremy K. Miller, Mallorie M. Smith, Katherine Moore, Marcus Möschl, Matt Motyl, Stephanie Muller, Marcus Munafò, Alisa Raquel Muñoz, Koen Hje Neijenhuis, Taylor Nervi, Gerdalf Nicolas, Gustav Nilsson, Brian A. Nosek, Catherine Olsson, Oileen Osborne, Lutz Oskamp, Misha Pavel, Olivia Kathleen Perna, Dr Cyril Pernet, Marco Perugini, R Nathan Pigotone, Michael Pitts, Franziska Piessow, Jason M. Prenoveau, Ate Ratliff, David Reinhard, Frank Renkewitz, Daan van Renswoude, Ashley A. Ricker, Anastasia Rigney, Hedderik van Rijn, Andrew M. Rivers, Mark Roebke, Abraham M. Rutchick, Robert S. Ryan, Onur Sahin, Anondah Saide, Gillian Sandstrom, David Santos, Rebecca Saxe, René Schlegelmich, Kathleen Schmidt, Sabine Scholz, Arissa Seibel, Dylan Selterman, Samuel Shaki, William B Simpson, H. Colleen Sinclair, Jeanine Skorinko, Agnieszka Slowik, Joel S. Snyder, Courtney K. Soderberg, Arina Sonnetimer, Nicholas Brant Spencer, Jeffrey R. Spies, Angela D. Staples, Sara Steegen, Mia Steinberg, Stefan Stieger, Nina Strohminger, Gavin Brent Sullivan, Thomas Talhelm, Megan Tapia, Manuela Thomae, Sarah Thomas, Helen Tibboel, Pia Tio, Frits Traets, Steve Tsang, Francis Tuerlinckx, Alexa Tullett, Paul Turchan, Jolif vanpaemel, Alejandro Vásquez Echeverría, Anna van't Veer, Natalia Vélez, Mathijs van de Ven, Marieke Vermue, Mark Verschoor, Michelangelo Vianello, Martin Voracek, Iina Vuori, Eric-Jan Wagenmakers, Joanneke Weerdmeester, Ashlee Welsh, Erin Westgate, Joeri Wissink, Michael Wood, Andy, Emily Wright, Sining Wu, Marcel Zeelenberg.

複数ラボが協力して同じ手続きでデータを収集する体制（多数の研究者の参加とそれをサポートするシステム）が必須

まとめ

早く科学に
なりたーい
問題

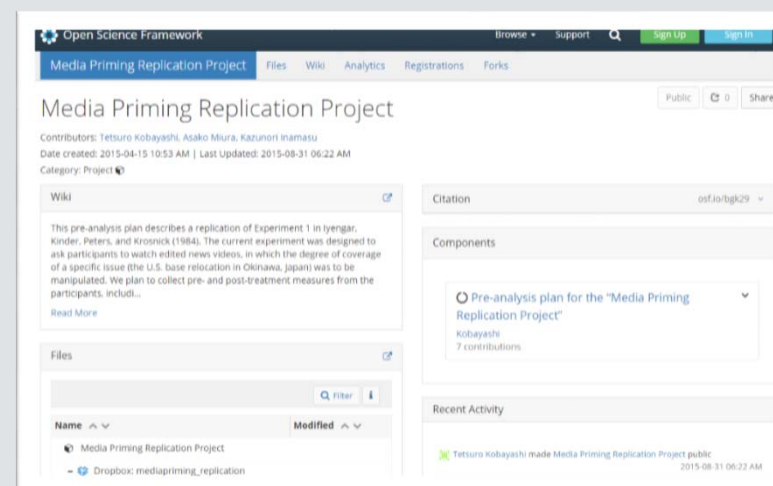
にんげん
だもの
問題

- 心理学の再現性が問われている
- 再現性の低さを産み出している要因は、心理学が「科学的手法」を標榜することそのものと、それによってもたらされる、実に人間らしい振る舞いの両方にある
- しかし解決のために様々な動きがあり、そのうちの一つが追試研究の奨励である
- シチズンサイエンスがそこに寄与できる可能性がありそう

Preregistration 事前登録

透明性と開放性の確保によって、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たとしても、そこに QRPs が含まれており、またそれが出版バイアスをくぐり抜けるために糊塗されていたとしたら意味がない。第9章で詳しく述べたとおり、QRPs は研究に着手した後、データ収集から分析へと進めて行く過程で発生しやすいのだから、そこで不正が行なわれないためには、「正」の研究手続きをあらかじめ定めておき、これを第三者に向けて公開することが効果的だと考えられる。これが**事前登録制度（pre-registration system）**であり、これを出版する側のシステムに取り入れ、なおかつこの時点で審査をする試みが始まっている。これが事前審査つき事前登録制度である。

事前審査つきの事前登録制度が画期的なのは、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点である。つまり、研究の目的（論文の序論部分にあたる）、取得予定のサンプルサイズ（実験参加者や調査回答者の数）、研究計画、統計分析手法など（論文の方法部分にあたる）を投稿し、その内容が審査される。そして、審査を通過した研究は、その内容をタイムスタンプ（登録日時）とともに事前に登録することによって、その時点以降の改編が不可能になるよう「凍結」される。その後実際にデータが収集され、得られた結果がどのようなものであろうと、その報告がそのジャーナルに掲載される。



Open Science Frameworkを利用したプレレジ
（事後にKobayashi, Miura, & Inamasu(2017)として公開）



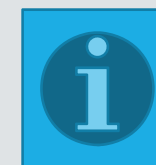
Nosek et al.(2018)

Preregistration 事前登録



事前審査付きの事前登録制度が画期的なのは、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点である。つまり、研究の目的（論文の序論部分にあたる）、取得予定のサンプルサイズ（実験参加者や調査回答者の数）、研究計画、統計分析手法など（論文の方法部分にあたる）を投稿し、その内容が審査される。そして、審査を通過した研究は、その内容をタイムスタンプ（登録日時）とともに事前に登録することによって、その時点以降の改編が不可能になるよう「凍結」される。その後実際にデータが収集され、得られた結果がどのようなものであろうと、その報告がそのジャーナルに掲載される。

<https://cos.io/rr/>



学会の対応(BPS)

We're offering Registered Reports
across all eleven of our academic
journals

13 July 2018

You can now improve the replicability of your work by publishing a
Registered Report with the British Psychological Society



日本での取り組み

心理学評論

Vol. 59 No. 1

2016

特集「心理学の再現可能性:我々はどこから来たのか 我々は何者が 我々はどこへ行くのか」

特集号の概要

巻頭言

目次(原著論文) ▼

目次(コメント論文) ▼



原著論文

- [池田功毅・平石界 心理学における再現可能性危機:問題の構造、現状と解決策](#)
- [山田祐樹 認知心理学における再現可能性の認知心理学](#)
- [森口佑介 発達科学が発達科学であるために:発達研究における再現性と頑健性](#)
- [鮫島和行 システム神経科学における再現可能性](#)
- [澤幸祐・栗原彬 動物心理学における再現可能性の問題](#)
- [大久保街亜 帰無仮説検定と再現可能性](#)
- [小塩真司 心理尺度構成における再検査信頼性係数の評価―「心理学研究」に掲載された文献のメタ分析から―](#)
- [藤島喜嗣・樋口匡貴 社会心理学における“p-hacking”の実践例](#)
- [渡邊芳之 心理学のデータと再現可能性](#)

コメント論文

- [小島康生 人間の観察研究における再現可能性の問題](#)
- [松田一希 フィールド研究の再現性とは何か?](#)
- [平井啓 心理学研究におけるリサーチデザインの理想](#)
- [三中信宏 統計学の現場は一枚岩ではない](#)
- [武田美亜 再現可能性の問題から始める心理学研究の「バックヤードツアー」](#)
- [東島仁 研究公正から見た再現可能性問題](#)
- [佐倉統 科学的方法の多元性を擁護する](#)

<http://team1mile.com/sjpr59-1/>

学会の対応（日本）

...?

学振の取り組み?

課題設定による先導的人文学・社会科学研究推進事業
人文学・社会科学データインフラストラクチャー構築プログラム