

社会心理学における 再現可能性問題の概要

竹澤正哲（北海道大学）

September 7, 2011



TILBURG UNIVERSITY

Understanding Society

Press Release 7 September 2011

Prof. Diederik Stapel suspended

The Executive Board of Tilburg University has suspended Prof. D.A. Stapel from his duties with immediate effect. Dr Stapel, who is a Professor of Cognitive Social Psychology and Dean of the Tilburg School of Social and Behavioral Sciences, has committed a serious breach of scientific integrity by using fictitious data in his publications.

Artikel

[terug naar 't overzicht](#)

PhD defense of student prof Stapel postponed

A thesis defense at Tilburg University has been postponed because the PhD student was supervised by professor Theo Verhallen and fraudulent former professor Diederik Stapel. The PhD defense was scheduled for October 7.

September 15, 2011

univers.online - a student magazine of UvT

November 28, 2012

Levelt Committeeによる調査報告書の概要

捏造データが含まれていた論文の総数は55本

+ 10人分の博士論文

このうち、2人の院生が執筆した博士論文は全てのデータがStapelの手によって捏造されたものであり、その1人は既にassistant professorとして就職



June - July, 2012



Dirk Smeesters
Erasmus University



Lawrence Sanna
University of Michigan

1. 捏造と改ざん

2. 追試

Journal of Personality and Social Psychology

© 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

Keywords: psi, parapsychology, ESP, precognition, retrocausation

Effect size of 0.50; skepticism about psi; issues of replication, and theories of psi are also

Keywords: psi, parapsychology, ESP, precognition, retrocausation

確率的事象としての実験

- 1つのデータポイントが確率的事象であるのと同様に、1つの実験もまた確率的事象に過ぎない
- 決定的な実験なるものは存在しえない — いかなる実験にも偽陽性の可能性が存在する
- 科学では、追試という名の集合知によってこの問題の解決が要請されている
- Bem騒動は、社会心理学においてこの制度が機能していないことをスキャンダラスな形で例示した

裏から読んで 超能力ってどうなのでしょう？

心理学

安田女子大学心理学部 講師
平石 界 (ひらいし かい)



平石界 (2013) 超能力ってどうなのでしょう？
心理学ワールド2013年4月号
http://www.psych.or.jp/publication/world_pdf/61/61-32.pdf

老人プライミングの追試を巡って

- Doyen, Klein, Pichon, & Cleeremans (2012) : 二重盲検法の下では老人プライミング (Bargh, 1996) は再現できない事、実験者効果の可能性が強い事を示す
- Psychological Today (ブログ) 上で, Barghが反撃 → PLOS ONEのeditorをも巻き込んだ論争勃発
- 追試という営みが正常に機能した例であるが、教科書に掲載される著名な研究ですら容易には再現できないことを鮮烈に示した

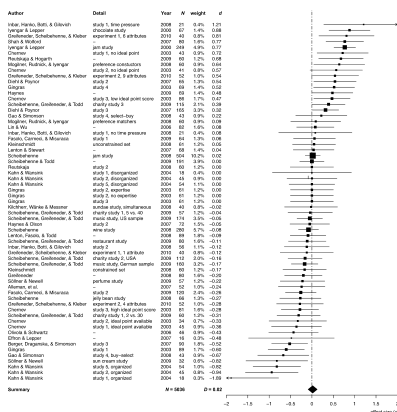
Choice overload (Iyengar et al. 2000)



選択肢の数が多いほど、選択後の満足度が低下する現象

コロンビア白熱教室 (NHK) で有名なSheena Iyengarによって見出された

Scheibehenne et al. (2010)によるメタ分析



$D_{mean} = 0.02$,
 $[CI_{95}] = -0.09 \sim 0.12$

ただし効果量のばらつきは、サンプリングエラーだけでは説明できない=何らかの特定の条件が効果を生み出している可能性を示唆

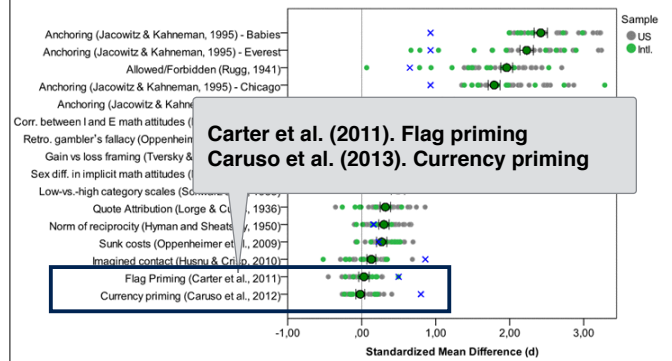
- だが追試の失敗 = 現象の否定ではない
 - 完全な追試は困難だから
 - 自然科学では実験のスキルを身につけるだけで長年の訓練が必要とされる
- 同様の理由により、メタ分析も最終兵器とは言えない
- 安易な追試 = 精度の低い研究による**数の暴力**を防ぐための試みが生まれつつある

Many Labs Replication Project

Richard A. Klein (Florida), Kate A. Ratliff (Florida), Brian A. Nosek (Virginia)が主導

12カ国、36の研究グループが協力し、同一のプロトコルを用いて、16の研究を追試 (n=6,344)

Many Labs Replication Project (2013)



Reproducibility Project

Brian A. Nosek (Virginia)が主導

トップジャーナル3誌に2008年に掲載された全論文を、150人の研究者が追試するプロジェクト

- *Journal of Personality and Social Psychology*
- *Psychological Science*
- *Journal of Exp. Psychology: Learning, Memory, and Cognition*

Open Science Framework

以上2つのプロジェクトは、Center for Open Scienceが構築したOpen Science Framework (wikiや、バージョン管理システムなどからなるweb環境) 上で展開されている



各種財団の支援を受け、2013年に Brian Nosek (Virginia) によって設立されたNPO

APSによる取り組み：事前登録制による追試

*Perspective on Psychological Science*誌は、事前登録制に基づく追試論文のカテゴリを新設 (*Registered Replication Reports*)

実験費用に対する資金援助も提供

3. False-Positive Psychology

False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant

Joseph P. Simmons¹, Leif D. Nelson², and Uri Simonsohn¹
¹The Wharton School, University of Pennsylvania, and ²Haas School of Business, University of California, Berkeley

False-positive psychology : 基本的なロジック

- $p = .05$ の検定で第1種の過誤が生じる確率は**5%**
- 2つの独立した検定で「第1種の過誤が最低でも1回以上生じる確率」は**9.75%** ($1 - 0.95^2$)
- 10の独立した検定の場合は**40%** ($1 - 0.95^{10}$)

(1) 複数の独立条件

- 自尊心を高めることがもたらす効果を検討。複数の操作が考えられるが、どれが本当に自尊心を高めるか事前には不明
- 統制条件の他に、自尊心を高める条件を複数設定→仮説通りの結果が得られた条件のみを報告
- 独立条件の数が3つの場合でシミュレーションを実施する。条件間で差がないにもかかわらず有意差が生じる確率は、5% → **12.6%**に上昇
- 42%の研究者が行っている (John, Loewenstein & Prelec, 2012)

(2) 複数の従属変数

- ある操作がストレスに与える影響を検証。だがストレスの指標は複数考えられるため、複数の従属変数を測定
- 最低でもどれか1つの従属変数で仮説通りの結果が得られたら報告
- 従属変数が2つの場合、5% → **9.5%**に上昇
- 78%が行っている (John et al., 2012)

(3) 共変量の統制

- 分散分析で有意な結果が得られない場合、関連しそうな変数を共変量として投入し(共)分散分析を実施し、有意な結果が得られたら報告
- 共変量が1つ、条件との交互作用効果1つを投入する場合、5% → **11.7%**に上昇

- (1) 複数の独立条件 **12.6%↑**
- (2) 複数の従属変数 **9.5%↑**
- (3) 共変量の統制 **11.7%↑**
- (4) ある操作X **7.7%↑**

(1)~(4)すべての手法を組み合わせると
第1種の過誤の発生確率は5% → **60.7%**に上昇

探索的研究と偽陽性のトレードオフ

- (1)~(3)の慣行は、しばしば探索的研究の名の下に正当化されている
- 特に社会心理学のように理論的な構成概念と操作的に定義された変数の間に一意の対応関係が定まらない領域では、独立条件や従属変数を事前に絞り込むことは困難

採択されつつある解決方法

- 行なったことを包み隠さず報告せよ — 測定したすべての変数、すべての条件、すべての分析を
- その上で結果の判断は読者と査読者の手に委ねよ — 偽陽性の可能性が高いか否かの判断を

Psychological Science

- MethodsとResultsセクションは字数制限から外す

Personality and Social Psychology Bulletin

- 実験で測定された全質問項目と全条件に関する情報の提出を義務化

(4) ある操作Xとは？

有意な結果が得られなかったら
n=10のサンプルを追加する

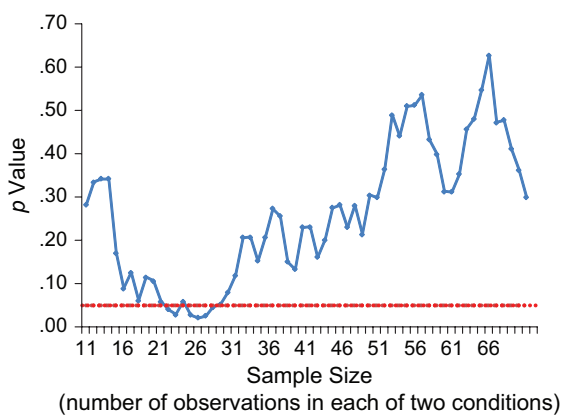
5% → **7.7%**

まずn=10のデータを取り、
有意になるまでn=1ずつデータを追加していく

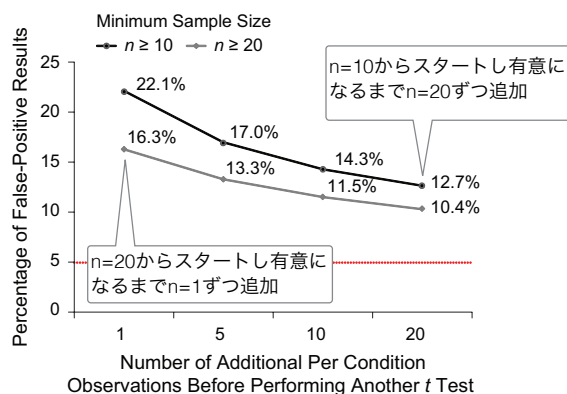
5% → **22.1%**

72%の研究者が行なっている (John et al., 2012)

Simmons et al. (2011): t検定のシミュレーション



Simmons et al. (2011): t検定のシミュレーション



最後に

再現可能性問題から生まれた制度的変化

- 追試を促進する動きの広まり
 - Open Science Framework
 - Perspectives on Psychological Science
- 探索的研究の名の下に用いられてきた慣行を、禁止ではなく開示によって第三者の評価に委ねる方針の広まり
 - Psychologica Science
 - Personality and Social Psychology Bulletin
- これらの変化への反対はもはや正当化できない

We need statistical thinking,
not statistical rituals

Gerd Gigerenzer (1998)