九州大学学術情報リポジトリ
Kyushu University Institutional Repository

これからの「再現性問題」の話をしよう

植田，航平
九州大学大学院人間環境学府行動システム専攻

益田，佳卓
九州大学大学院人間環境学府行動システム専攻

佐々木，恭志郎
Faculty of Informatics, Kansai University

山田，祐樹
九州大学基幹教育院自然科学実験系部門

https://hdl.handle.net/2324/6779693

出版情報：The journal of the Institute of Electronics, Information and Communication Engineers. 106 (4), pp.321-325, 2023-04-01. Institute of Electronics, Information and Communication Engineers
バージョン：
権利関係：(C)2023 IEICE

解説

これからの「再現性問題」の話をしよう

Reproducibility : What's the Right Research to Do?

植田航平　益田佳卓　佐々木恭志郎　山田祐樹

Abstract

心理学領域ではその知見に再現されないものが数多く存在することが指摘されてきた．本稿では，この課題に対して心理学コミュニティの内部で議論されてきた話題や対策，オープンサイエンスの潮流を概観し，特に統計的手法に関する話題を中心に議論する．再現性問題は，心理学にとどまらず，人間を研究や測定の対象とするあらゆる分野に関わる話題であると言える．本稿は，中心的な読者である工学や情報学の専門家の研究遂行に有用な情報を提供し，将来，心理学者とこれらの領域の研究者とが協同して学術を盛り上げることを目指す．

キーワード：再現性問題，問題のある研究実践，帰無仮説検定，事前登録，オープンサイエンス

1. 再現性問題

研究の成果が再現されることは，科学の営みにおいて重要な要件の一つであろう[(1)]．「巨人の肩」を築き上げていくためには，科学者コミュニティによる頑健な知見の蓄積が必要不可欠であり，分野を問わず共通する課題である．

しかし，近年，様々な分野において再現性に疑義が呈されており（e.g., 文献(2)），心理学領域においても再現されない知見が数多く存在することが指摘されてきた[(1),(3)]．心理学コミュニティにおいてこの事実に注目が集まり始めた2010年代前半から続く，およそ10年の議論はいまだ途上ではあるものの[(4)]，これまでに「信頼性改革」[(5)]と呼ばれる方法論等の抜本的な見直しや改善策の導入が数多くなされてきた．

本誌の主な読者層である工学系や情報系の研究者・開発者にとっては，心理学はそこまでなじみの深い分野ではないかもしれない．しかし，技術やサービス，プロダクトを実際に使用するユーザ，すなわち人間の認知や印象などに関する知見は工学的研究や開発にも生きると考えられる．また何より，開発した技術や製品の効果を検証する際には人間を相手に行動実験を行う必要があり，そのためには方法の様々な点に配慮することが求められる．心理学は，まさに行動実験やアンケート調査などの研究手法を十八番とする学問である．再現性問題への対応を含め，長い時間をかけて科学的であることを追求する中で磨かれてきた行動計測技術や実験法は，ユーザの反応や声を重視する工学や情報学等の領域における研究実践の際にも大いに参考になると考えられる．そこで本稿は，心理学領域における再現性問題やその対策を概説し，工学や情報学の研究を行う上でも有用だと考えられる情報を提供することを目指す．

再現性問題に関する議論は，研究計画からデータ取得，分析，その他方法論や科学哲学，科学技術社会論的な話題まで多岐にわたる．ただし，本稿は紙幅の制限上，全ての話題を網羅することはできず，主に統計的手法に関する議論をするにとどまった．また，それぞれの数学的，統計学的，哲学的な詳細の議論は引用元の各文献に譲った．

2. 不適切な研究方法

再現性問題に対処すべく行われてきたこれまでの議論の中で，知見の再現性を低下させる要因として，いわゆるネカト（捏造・改ざん・盗用）に代表される研究不正は当然のことながら，それ以外にも多くの問題点がある

植田航平　益田佳卓　九州大学大学院人間環境学府行動システム専攻
佐々木恭志郎　関西大学総合情報学部総合情報学科
山田祐樹　九州大学基幹教育院自然科学実験系部門
E-mail　yamadayuk@gmail.com
Kohei UEDA, Yoshitaka MASUDA, Nonmembers (Graduate School of Human-Environment Studies, Kyushu University, Fukuoka-shi, 819-0395 Japan), Kyoshiro SASAKI, Nonmember (Faculty of Informatics, Kansai University, Takatsuki-shi, 569-1095 Japan), and Yuki YAMADA, Nonmember (Faculty of Arts and Science, Kyushu University, Fukuoka-shi, 819-0395 Japan).
電子情報通信学会誌　Vol.106 No.4 pp.321-325 2023年4月
©電子情報通信学会 2023

ことが明らかになってきた．その中でも特に重要なものに，問題のある研究実践（QRPs：Questionable Research Practices）が挙げられる．QRPsの悩ましいところは，それを必ずしも意図的に行っているとは限らない点にある．すなわち，適切だと考えて行っていることが実はQRPsに該当するということがあり得るのである．そこで，心理学におけるQRPsに関連する議論を概観し，工学や情報学等の領域においてもより良い研究営為を行うための一助となるような事項を提示したい．

まず，代表的なQRPsとして仮説の事後生成（HARKing：Hypothesizing After the Results are Known[6]）が挙げられる．これは，論文のストーリー構成上の一貫性や審美性（いわゆる，「きれいな結果」）を高めるため等の理由から，結果を確認した後に，さも事前に予測していたものであるかのように仮説を事後的に生成することを指す．例えば，予期していなかった条件間で有意差が見られたとする．その結果を知った後に，あたかも当初からその結果を予測していたかのように序論を書くことはHARKingに当たる．HARKingの問題点は，帰無仮説検定を用いる上での偽陽性の報告可能性が高まってしまうところにある[3]．偽陽性とは，本来は効果がないのに陽性（有意）だと判定してしまう誤りのことを指す．帰無仮説検定は確証的研究の偽陽性率をコントロールするように設計されており，仮説や研究計画・分析手順を事前に定義しておくことが必要になる．しかし，HARKingを含んだ研究報告は，この前提に反するため，偽陽性率をコントロールできなくなってしまうのである．HARKingにおける問題点の背景には，仮説検証型の確証的研究と，仮説を形成するための探索的研究が混同されていることがあると考えられる．データに対して，自由な発想に基づいて事前の計画以外の分析をすること自体は制限されるものではない．しかし，それはあくまでも次の研究の仮説を形成するためのものであり，偽陽性の可能性が高い（言い換えれば，1回のデータにオーバフィットしている可能性がある）結果を基に，断定的に主張を展開することは望ましくないのである[7]．

そのほかにも，多種多様の分析を行い，そのうち「望ましい結果」だけを選択的に報告するチェリーピッキング[8]，本来結果から主張できないはずの解釈を記述する粉飾（スピン，e.g.，文献（9））なども，QRPsとして挙げられる．チェリーピッキングの例としては，2群の比較をしたいときに平均値や中央値といった様々な代表値やデータの変換，検定方法などを用いた思いつく限りの分析をひそかに行い，その中で自身の仮説・予測を支持するような結果だけを論文などで報告し，都合の悪い（e.g.，予測と矛盾するような）結果は伏せることなどが挙げられる．また，論文内での考察や外部へのプレスリリースなどの際に，ある技術による介入効果が統制条件なしで2時点間で有意だったことをもって，あたかも抜群の性能を示したかのように表現するといった具合に結果を過度に誇張・脚色・歪曲・一般化して主張することや，後述するように有意ではなかった結果をもって差や効果が「なかった」と解釈することはスピンに当たり，読者に対して誤った印象を与えてしまう．そのため，得られた結果から言えることを正確に記述，報告することが大切である．

上述のような研究の計画や仮説，結果の解釈に関わる側面だけでなく，統計的分析においても留意すべき点が存在する．不適切な統計的手法として代表的なものに，p値ハッキング（p-hacking）がある．p値ハッキングとは，帰無仮説検定において「有意な結果」が得られるよう，データの取得や分析時の手順，データの除外基準などを恣意的に操作することによって，p値を低めようとする行為である[8]．例えば，帰無仮説検定はその特性上，サンプルサイズが大きくなればなるほど有意になりやすい傾向があるが，この特性を悪用し，何度も中間分析を行い「有意な結果」が得られるまでデータを追加したり（N増し），たまたま有意になった時点でデータ収集を打ち切ったりすること（任意停止）は，代表的なp値ハッキングである．そのほかにも，外れ値の基準を事後的に決定するなどして望まないデータを削除したり（外れ値活用），事前の計画や仮説の段階で想定していなかった共変量を含めた分析を行って得られた「有意な結果」をあたかも計画していたかのように報告すること，分析に至るまでのデータの加工の方法を事後に恣意的に決定することなどもp値ハッキングである．このように，かつては是とされてきた研究営為も含め，p値ハッキングに該当する行為は多岐にわたる[注1]．

また，p値ハッキングとはやや異なるが，いわゆる「有意傾向」（marginally significant）という記述にも注意が必要である．現在の帰無仮説検定は，ネイマン・ピアソン流の流れをくんでいるが[10]，その枠組みでは事前に有意水準を決定する必要がある．そして，p値が有意水準を下回るか否かで，効果の有無を2値的に判断する．心理学をはじめ，多くの分野では有意水準を5%（または1%）に事前に設定することが多いが，p値が10%未満であった場合に「ほぼ有意な結果である」というような解釈をすることは，事前に決めたはずの水準，いわばゴールポストを恣意的にずらしていることになり，問題がある[11],[12]．一方で，p値が非常に小さく，

（注1） ここで重要なのは，上述のHARKingの議論のとおり，仮説や事前の計画と異なる分析を行った際は，そのことを明記し，二次的・探索的な分析として取り扱うことである．もちろん，偽陽性の問題などはある上に，あくまで探索的な知見であるため，より強い主張をするためには改めて仮説を生成し，再びデータを取得した上で分析，議論することが求められる．これは，後述の査読付き事前登録の論文でも同様である[21]．

有意水準を大きく下回ったときに「高度に有意」などと記述することも，ネイマン・ピアソン流の仮説検定の考え方にのっとれば，同様に適切ではない[11]．

更に，有意ではなかった結果についての解釈も注意が必要である．一般に帰無仮説検定では差がないことは示すことができない．（帰無仮説を採択できない．）そのため，有意な結果ではなかったからといって，「差がない・効果がない」と結論することは妥当ではない．そのようなことを示したい場合は，非劣性検定または同等性検定[13]やベイズファクター[14]などを用いる必要がある．

ここで紹介しているものは，基本的に偽陽性や結果の誤った解釈を導くものである．工学や情報学等の研究ではデバイスやツールの効果検証を行うことも多いが，その場面においては，上記の行為はその効果について読者に誤った印象を与えるものである．つまり，効果がでたらめなものを世に提供することにつながりかねないため，こうした行為がいかに危ういかを理解頂けると思う．

3. 提案されている対処法

それでは，これらの問題に対してどのように対策をすればよいのだろうか．前述のような QRPs の多くが生じ得るのは，データの取り方から統計的手法，結果の解釈に至るまで研究者が持つ自由度の多さに起因することが指摘されてきた[15]～[17]．心理学においては，このような研究者が恣意的に決定できる範囲を狭めることで QRPs が生じる余地を減らすアプローチが提案され，実践されてきた．

代表的な手法が，研究の事前登録（pre-registration，プレレジ）である[18]．プレレジとは，実験や調査を実施する前にその研究の目的や仮説，手続き，分析手法といった具体的内容を第三者機関（e.g., OSF, https://osf.io/）に登録することを指す．登録内容はその日時とともに記録され変更できなくなるため，プレレジした内容とは異なる予測や手続きによる実験，解析の存在を明確にすることができる．プレレジされた研究は登録内容にのっとって行われるため，データの取得後など研究の中途における研究者自由度を低減することにつながり，HARKing や p 値ハッキングを防ぐことが期待される．なお，プレレジの具体的な流れについては，文献(19)が詳しい．ただし，プレレジのみで QRPs を全て防ぐことの限界も指摘されており[17], [20]，近年ではプレレジをより発展させた査読付き事前登録（RRs : Registered Reports，レジレポ[21]）を受け付ける雑誌も増えてきている．レジレポは，その名のとおり事前登録の内容に対しても査読を受け，その後データを取得して結果や考察が記載された原稿に対して改めて査読がなされるという出版スタイルである．第一段階の査読をパスした研究は原則的採択（in-principle acceptance）となるため，その後の結果が仮説を支持するものであるかどうかにかかわらず原則受理され，出版される．そのため，HARKing をはじめとする QRPs をそもそも行う必要や動機付けがなくなるのである．

また，昨今，オープンサイエンスの潮流が世界中を席巻している[22]．前述のプレレジやレジレポもその流れの一つである．加えて，調査・実験で用いた刺激や結果のデータ，分析スクリプト等を公開することで，第三者による検証や刺激の再利用などを行えるようにするオープン化も推奨されるようになってきた[23]．オープンサイエンス自体は科学の進歩や展開に合わせてこれまでもこれからも広がっていく流れであると考えられるが[24]，それだけでなく，再現性問題への対処とも密接に関わっている．例えば，OSF 上に研究の計画や手続き，仮説などをプレレジすれば，誰でもその記録を確認することができるため，事後の検証が容易になる[注2]．また，このようなオープンサイエンスの流れの中に再現性問題の対処法を位置付けることで，それぞれの取組みにグラデーションを付けて捉えることも可能かもしれない．研究遂行上，様々な理由で情報等を公開できないというケースも考えられるが，可能な範囲でオープンサイエンスを実践していくことは事後の検証や追試，後続研究につながる一歩となるのである．特に，工学分野であれば開発に伴う特許や知的財産権等の折衝が求められることもあるだろう．この辺りについては議論の途上であり，様々な分野・領域の科学者や社会の構成員が引き続き議論に参画することが重要だと言えよう．

これらの対策に加え，統計的な側面についても議論がなされてきた．本稿ではサンプルサイズと検定力，効果量の三つについて取り上げる．帰無仮説検定はこれら 3 要素に有意水準を加えた四つの要素から成る．すなわち，検定を適切に用いるためにはこれらの理解が必要である．

かつては確証的研究であるにもかかわらず，何名を対象に実験を行うのかが曖昧なままデータ収集が行われるケースも度々見受けられた．一方で，帰無仮説検定については事前にサンプルサイズを決めることが原則的には望ましく，近年ではこのことも強く推奨されるようになった．（サンプルサイズ設計または例数設計，その方法を含め文献(25)に詳しい．）その方法の一つに有意水準や後述の検定力，効果量から必要なサンプルサイズを決定する検定力分析がある．先述のとおり，サンプルサイズを事後的，恣意的に決定することは QRPs に当た

（注 2） なお，OSF へプレレジする内容は公開に猶予期間（embargo）を設けることも可能であり，最終的な研究成果が公表されるまで第三者に情報が漏れないようにすることもできる．これは，上述の特許や知財とオープンサイエンスに関する議論とも一部関連するだろう．

る．やはり重要な点は，帰無仮説検定を使用する際は原則的に適切なサンプルサイズを事前に決定し（研究報告時には決定理由について明示），その計画にのっとって研究を遂行することである[注3]．

次に，検定力とは，帰無仮説が誤っているときにその帰無仮説を棄却できる確率である[25]．サンプルサイズを事前に設計する目的の一つが，適切な検定力を確保することである．なぜなら，検定力が十分でなければ，本来は有意な結果が得られるはずであったものが帰無仮説を棄却できず，本来主張できたはずの成果を取り逃がしかねないからである．一方で，余りに大き過ぎるサンプルサイズも問題であるため[25]，適切な検定力を確保することが必要であり，そのための手法の一つが前述のサンプルサイズ設計なのである．

効果量とは，「『効果がない』という帰無仮説の状態に対して，対立仮説がどれくらい異なっているかを量的に示すもの」[25]である．分かりやすく言えば，差や効果がどれほどの大きさであるかということを，他の研究と比較しやすくするための量的指標である．（詳細については文献(10)等を参照されたい．）近年では検定を行った際には，p 値だけではなく効果量も併記することが推奨されてきているが，この背景には帰無仮説検定における2値判断への批判がある[27]．一般に有意性の判断は差や効果がある・ないの2択である．しかし，工学や情報学などの応用場面を見据えた研究では，統計的に有意であるだけではなく，実際の場面において意味・意義を持つ効果の大きさなのか（practical significance[28]）に基づいた判断が必要なこともあるだろう．ここで，p 値は差や効果の大きさを表す指標ではない．そのため，帰無仮説が棄却されたとしても，それは差や効果が0ではないということが示されただけで，どれほどの効果があるかは判断できない．その欠点を補うために，p 値だけではなく効果量も記載することで効果の大きさを量的に解釈することが可能になるのである．

このように，帰無仮説検定に関わる統計量や量的指標は多岐にわたる．これらを正確に理解し，事前にサンプルサイズ設計を行った上で研究を遂行し，論文内にも適切に各種統計量を記載して，論文の読者に対して豊富な情報を提供することが，再現性を高めることにもつながるのである．そもそも，帰無仮説検定の結果のみに依存して科学的結論の全てを下すことには慎重になる必要がある（e.g., 文献(29)）．様々な指標や知見を総合して結論する姿勢が再現性問題に限らず重要であろう．

（注3） 検定力分析によるサンプルサイズ設計を実際に行うと，多くの場合，1群当り6〜10名といった慣習的に決定していたサンプルサイズよりも，かなり多くの参加者を募ることが必要になる．（場合によっては桁が二つ変わることもある．）所属機関の立地など，様々な事情でそれだけの規模のデータを集められない場合については，その理由をプレレジや論文等に明記した上で，可能な範囲でデータ収集を行い，それに応じた適切な結果の解釈をすることが望ましい．また，サンプルサイズをどのように決定し，正当化するかについては文献(26)に詳しい．

4. 今後の方向性

本稿では，心理学における再現性問題とその対応について概観し，人に関する工学的・情報学的研究を行う上で参考になると思われる情報について述べてきた．ここまでの議論を受け，我々は人間についての工学的研究における，工学者や情報学者と心理学者（ないし心理統計学者）との協働を提案したい．

心理学者は10年以上にわたり，より一層心理学の知見の信頼性を高めるため，様々な議論と取組みを行ってきた．それゆえに，人間を対象とする研究を行う際の方法論や統計的手法についての知識が積極的に議論・更新されてきたと言えるだろう．これらの取組みは，いまだ道半ばであり，再現性問題が完全に解決したとは言い難い状況である[4]．しかしながら，着実に方法論は洗練されてきており，少なくともそれらに対する意識が高まっていることは事実である．行動実験について真剣に向き合う中で培われてきたこれらのノウハウは，心理学領域だけではなく，近隣の応用領域の研究にも生かす道があるのではないだろうか．

そこで，今後は工学や情報学，心理学といった領域にとらわれず，三者やあるいはそれ以外の領域の研究者も含め，「人間を対象とする研究領域」全ての研究者が協力していく体制を作り上げることが必要だろう．（特に基礎研究を行うことが多いような）心理学者は，工学者や情報学者と組むことで知見の応用や現実問題への解決・支援といった視点を自身の研究に取り入れることができ，他方，工学者や情報学者は心理学の基礎的知識や手法を取り組むことで最新の心理学的な知見に基づいた研究を行うことが可能になるだろう．再現性の危機をある種の契機として，心理学と関連領域が手を取り合うことで応用的な研究と基礎的な研究を車輪の両輪として回し，それぞれの分野が一層の発展を遂げることに期待したい．

文　　献

（1） Open Science Collaboration, “Estimating the reproducibility of psychological science,” Science, vol. 349, aac4716, Aug. 2015.

（2） M. Wadman, “NIH mulls rules for validating key results,” Nature, vol. 500, pp. 14-16, July 2013.

（3） 池田功毅，平石　界，“心理学における再現可能性危機：問題の構造と解決策，” 心理学評論，vol. 59, no. 1, pp. 3-14, July 2016.

（4） 平石　界，中村大輝，“心理学における再現性危機の10年—危機は克服されたのか，克服され得るのか—，” 科学哲学，vol. 54, no. 2, pp. 27-50, March 2022.

（5） S. Vazire, “Implications of the credibility revolution for productivity, creativity, and progress,” Perspectives on Psychological Science, vol. 13, no. 4, pp. 411-417, July 2018.

（6） N.L. Kerr, “HARKing : Hypothesizing after the results are known,” Personality and Social Psychology Review, vol. 2, no. 3, pp. 196-217. Aug. 1998.

（7） D. Bishop, “Rein in the four horsemen of irreproducibility,” Nature, vol. 568, p. 435, April 2019.
（8） J.P. Simmons, L.D. Nelson, and U. Simonsohn, “False-positive psychology : Undisclosed flexibility in data collection and analysis allows presenting anything as significant,” Psychological Science, vol. 22, no. 11, pp. 1359-1366, Oct. 2011.
（9） I. Boutron, S. Dutton, P. Ravaud, and D.G. Altman, “Reporting and interpretation of randomized controlled trials with statistically nonsignificant results for primary outcomes,” JAMA, vol. 303, no. 20, pp. 2058-2064, May 2010.
（10） 大久保街亜，岡田謙介，伝えるための心理統計―効果量・信頼区間・検定力，勁草書房，東京，2012.
（11） R. Dar, R.C. Serlin, and H. Omer, “Misuse of statistical tests in three decades of psychotherapy research,” Journal of Consulting and Clinical Psychology, vol. 62, no. 1, pp. 75-82, Feb. 1994.
（12） W.M. Otte, C.H. Vinkers, P.C. Habets, D.G.P. van IJzendoorn, and J.K. Tijdink, “Analysis of 567, 758 randomized controlled trials published over 30 years reveals trends in phrases used to discuss results that do not reach statistical significance,” PLOS Biology, vol. 20, no. 2, e3001562, Feb. 2022.
（13） D. Lakens, A.M. Scheel, and P.M. Isager, “Equivalence testing for psychological research : A tutorial,” Advances in Methods and Practices in Psychological Science, vol. 1, no. 2, pp. 259-269, June 2018.
（14） 岡田謙介，“ベイズファクターによる心理学的仮説・モデルの評価，” 心理学評論，vol. 61, no. 1, pp. 101-115, July 2018.
（15） A. Gelman and E. Loken, “The garden of forking paths : Why multiple comparisons can be a problem, even when there is no “fishing expedition” or “p-hacking” and the research hypothesis was posited ahead of time,” 2013. http://www.stat.columbia.edu/~gelman/research/unpublished/p_hacking.pdf
（16） 山田祐樹，“自由を棄てて透明な心理学を掴む，” 心理学ワールド，vol. 83, pp. 34-35, Oct. 2018.
（17） A. Ikeda, H. Xu, N. Fuji, S. Zhu, and Y. Yamada, “Questionable research practices following pre-registration,” 心理学評論，vol. 62, no. 3, pp. 281-295, Feb. 2020.
（18） B.A. Nosek, C.R. Ebersole, A.C. DeHaven, and D.T. Mellor, “The preregistration revolution,” Proc. National Academy of Sciences, vol. 115, no. 11, pp. 2600-2606, March 2018.
（19） 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，“実証的研究の事前登録の現状と実践，” 心理学研究，vol. 92, no. 3, pp. 188-196, Aug. 2021.
（20） Y. Yamada, “How to crack pre-registration : Toward transparent and open science,” Frontiers in Psychology, vol. 9, 1831, Sept. 2018.
（21） C.D. Chambers and L. Tzavella, “The past, present and future of registered reports,” Nature Human Behavior, vol. 6, pp. 29-42, Nov. 2022.
（22） UNESCO, “UNESCO Recommendation on Open Science,” 2021. https://unesdoc.unesco.org/ark:/48223/pf0000379949.locale=en
（23） 武藤拓之，“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，vol. 92, no. 9, pp. 800-805, Sept. 2022.
（24） 有田正規，“オープンサイエンスの源流からこれまで，” 科学，vol. 92, no. 8, pp. 698-702, Aug. 2022.
（25） 村井潤一郎，橋本貴充，心理学のためのサンプルサイズ設計入門，講談社，東京，2017.
（26） D. Lakens, “Sample size justification,” Collabra : Psychology, vol. 8, no. 1, March 2022.
（27） R.L. Wasserstein and N.A. Lazar, “The ASA statement on p-values : context, process, and purpose,” The American Statistician, vol. 70, no. 2, pp. 129-133, June 2016.
（28） R. Torkar, C.A. Furia, R. Feldt, F.G.O. Neto, L. Gren, P. Lenberg, and N.A. Ernst, “A method to assess and argue for practical significance in software engineering,” IEEE Trans. Softw. Eng., vol. 48, no. 6, pp. 2053-2065, June 2022.
（29） 三中信宏，“再現可能性ははたして必要なのか：p 値問題から垣間見る科学研究の多様性，” 計量生物学，vol. 38, no. 2, pp. 117-125, March 2018.

（2022 年 10 月 31 日受付）

植田（うえだ） 航平（こうへい）

2021 島根大・人間科学卒，同年から九大大学院人間環境学府修士課程に在籍．認知心理学を軸に，航空分野におけるヒューマンエラー防止，低減のための研究を目指す傍ら，再現性問題やオープンサイエンスをはじめとするメタ科学的観点から見る心理学にも関心を抱いている．

益田（ますだ） 佳卓（よしたか）

2021 新潟大・人文卒，同年 4 月から九大大学院人間環境学府修士課程在籍．心理学の信頼性向上のための取組みやオープンサイエンスを軸に，時間知覚や科学哲学，統計学に関心を抱いている．最近は社会と科学の関わり方についても興味が広がってきた．

佐々木（ささき） 恭志郎（きょうしろう）

2016 九大大学院人間環境学府博士後期課程了．博士（心理学）．関西大・総合情報・准教授．感情，身体性，心理的時間・空間，心理的所有感など幅広い研究テーマを扱っている．近年はメタサイエンス研究にも従事している．

山田（やまだ） 祐樹（ゆうき）

2008 九大大学院人間環境学府博士後期課程了，博士（心理学）．九大・基幹教育院・准教授．ヒトの知覚・認知メカニズムの解明を主目的として研究しており，研究行為自体を支えるメカニズムにも興味を持っている．