

## 研究公正の諸側面：再現性と研究倫理審査体制

野 内 玲

# 研究公正の諸側面：再現性と研究倫理審査体制

野内 玲\*

## はじめに

近年では、科学研究を進める際の研究の公正性についての取り組みが盛んであり、医学と心理学を中心に研究の再現性問題への関心が国際的にも高まっている。この問題は統計分析や研究データの扱いに関する話題だと思われがちだが、これらを用いない研究分野や研究課題においても無関係ではない。自らがそうした研究を実施しなくとも先行研究として引用する場合もあるし、問題の所在を理解することは重要である。さらに研究公正の問題の一部として、研究成果を客観的かつ透明な仕方でもとめ、他の研究者が適切に理解できる形で発表することの重要性は論をまたないであろう。その際、研究実施の適切性を審査する研究倫理審査委員会も重要な役割を果たしている。

以上の背景を踏まえ、本論では研究活動の公正性を支える考え方を検討するため、研究分野を横断した基礎調査の結果から研究の再現性問題と研究倫理審査に関する分野横断的側面と分野固有の側面とを探索的に分析する。その際、当該問題に関する議論が先行している医歯薬学分野の議論に加え、人文学・社会科学分野の状況も押さえた含意を導くことを目的とする。本論の構成は以下である。1章で研究の再現性問題の見取り図を描く。2章では実施したアンケートの詳細を説明する。3, 4章では研究の再現性と研究倫理審査委員会に関する意識について研究分野比較を行い、その中でもこれらに関する議論を構築する必要があると考えられる人文学・社会科学分野について、医歯薬学分野と比較した結果を検討し、最後に今後の展望を述べる。

## 1. 人文学・社会科学分野における再現性問題

### (1) 研究の再現性問題の背景

本研究の背景には心理学や医学分野での「再現性の危機」とされる近年の動向がある。本節と次節では野内（2022）2章に加筆・修正し、この背景を描く<sup>1)</sup>。医学系研究の再現性問題を指摘した先駆的文献（Ioannidis, 2005）では、医療系研究においてよく引用されるインパクトの高い論文であつても再現性が確認できない場合があることが定量的に分析された。また、心理学分野では社会心理学のトップジャーナルに掲載された超能力論文（Bem, 2011）を端緒に再現性問題が広く認識されるに至ったとされている。統計的検定力が低い手法では、ある効果が実際に生じていてもそれを検出することは難しく（心理学の場合、20～40%程度の検定力しかない場合が多い）、検定力が

---

\* 信州大学医学部助教

低い手法で検証すると報告が増せば増すほど有意ではない報告が含まれる可能性が高くなる。それでも予想通りの結果が出続ける場合、その効果の实在ではなく手続きに問題がある可能性が高いという（統計的有意差の導出に誤りか不正がある）（池田他，2015，5頁）。Bem は結果を見てから仮説を立てるべしという誤った方法論を明言しており，研究手法の考え方自体にも問題があったという（平石・中村，2021，29頁）。また，100件の心理学研究のうち36%の結果しか再現性できなかったという調査結果が示され（Open Science Collaboration，2015），大きな物議を醸した。これらには後述する統計分析やデータの扱いに関わる不正関わっているが，この種の不正を回避するため当該分野では事前に立てた研究計画が適切に遵守されるよう論文投稿前の事前登録といった制度が広まりつつある。

## (2) 人文学・社会科学における再現性問題

次に，人文学・社会科学の研究においてこの問題がどのように関わるかを検討したい。国里は Goodman et. al. (2016) で示された「方法・結果・推論の再現性」を踏まえ，心理学研究における再現性問題の全体像を明瞭に描写している（国里，2020）。ここでは「再現性とは，ある現象が成立する条件を整えることができれば，その現象が繰り返し生じることを指す」という定義が示されている（国里，2020，21頁）。人文学・社会科学において統計学を手法として用いる分野（教育学や経済学）では，心理学分野でのこの定義が同様に当てはまるだろう<sup>2)</sup>。

統計分析を行う際の再現性を損なう不正として HARKing や *p*-hacking と呼ばれる行為がある。HARKing とは探索的にデータを集め，複数の統計分析にかけて有意差が見られたらそのデータと分析結果に基づいて「後付け」で仮説や研究デザインを設定し，論文にまとめる行為である（Hypothesizing After the Results are Known）。その論文は，当初から仮説検証型の研究を計画している，実施した成果として仮説が確かめられた「見栄えの良い」論文となる。一方，*p*-hacking は有意差が出るようにデータを取捨選択する行為である（そのため HARKing と両立する）。いずれの場合も，データが恣意的に選ばれたり，「外れ値」ばかりを含むものであったりという可能性があるため，他の研究者が同じ手順を経ても再現ができないこととなる。

これら行為の説明は心理学や医学など数値データを使う場合が念頭にある。それ以外の研究，例えば人文学・社会科学で標準的な文献解釈等の研究における再現可能性については別の観点がありうるが共通点もある。この点について，保城は社会科学と歴史学が相互に乗り入れた研究を実施する際の問題として，次の行為を指摘していることに着目したい<sup>3)</sup>。

…この問題には二つのタイプが存在すると考える。一つは，社会学者が既存の歴史研究（二次文献）を参考とする際に生じるもので，自分の理論のためにそれら歴史研究を恣意的に取捨選択して利用する，という行為である。もうひとつのタイプは，研究者（中略）が仮説をあらかじめ用意して，それをもって一次資料を調べる際に起きる。すなわち，自分の説にとって都合の悪い資料は無視し，都合の良い資料のみを証拠として挙げる，という問題である（保城，2015，12頁）（下線強調は筆者による）。

歴史学に限らず、人文学・社会科学における「研究データ」である先行研究や文献資料を取捨選択し、都合の良い資料のみに言及するという構図は HARKing 及び *p*-hacking という行為と同型であり、論述の追跡可能性 (Traceability) を損なう。すなわち、統計を扱わないから再現性問題とは関係がないとする理解は性急である。ただし、研究の探索段階では仮説の構築と検証の明確な区分は困難であり、成果に研究の実施過程をどこまで網羅的に記載するかについては議論が必要である。自身にとって不都合な資料に単に気づいていなかった場合もあり得るだろうが、対立仮説に反論する議論を含め、確証バイアス (Nickerson, 1998) を排除した客観的な議論を構築することが重要である<sup>4)</sup>。

以上を一般化したい。実験を用いない研究において、根拠となる文献・史料や先行研究の解釈 (研究データに相当) から新たな見解を導出する際には、当該データ自体の妥当性に加えて、著者の議論を他の研究者が適切に追うことが可能な形での論述が求められる。研究対象には、社会的構成概念や文献それ自体の解釈、不十分な歴史資料から推測される過去の事実等、論者によって見解が分かれるものも含まれる<sup>5)</sup>。そのため、論述の説得力を示すことが研究の本質となる。その際の論述も、真実を明らかにするというよりも、論述において先行研究等の多様な解釈を相互に競わせ、新奇な解釈へと結び付けていくことに意義が見出される場合もある。もちろん、解釈の帰結に多様性は認めつつも個々の論述は他者から客観的かつ妥当な仕方では解釈できなければならない。根拠となるデータや先行研究が不正に取り扱われたり、それらに基づく分析や論述が不適切であったりすれば、当該研究を再現すること (読者が同じ結果や解釈に達すること) はできない。この意味で、論述の追跡可能性を「再現性」の一形態とみなすことが可能である。

### (3) 研究の再現性問題と研究倫理審査を関連させて追求する意義

研究の再現性問題には多様な側面がある。データと統計の不適切な適用のために再現性が保たれない場合がその代表であるが、研究対象それ自体の特性として、ビッグバン、温暖化問題のような複雑系、進化論という過去から続く一連の出来事など、条件を統制しにくい検証データを取得できず、再現性を評価できない場合も多々ある。これらは不適切な行為による帰結としての再現性問題とは区別が必要である。しかしこれらを研究対象とする論文でも、データから結論への推論の妥当性等、論述の追跡可能性という意味での「再現性」が保たれている必要はある。結論へ恣意的に誘導するためにデータを改ざんしたり、自身の仮説を支持しやすい先行研究を選択したりといった行為は、いうまでもなく不適切な研究行為である。このように研究の再現性問題はデータ不正の問題や研究実施の客観性・透明性、議論の適切性といった、研究公正の規範に関する側面が含まれている。問題として過度に強調するより研究の質を向上させるための契機と考えるべきではないか。

文部科学省の「研究活動における不正行為への対応等に関するガイドライン」に基づき、各大学では構成員に対し、研究不正防止講習会、研究者倫理の e ラーニング等を提供しているが、ガイドラインの策定と e ラーニングによる各研究者への画一的な知識の提供だけでは研究公正の考え方やそれに従った実践について網羅的に習熟することは不可能だと考えられる。これは研究者集団として対応する必要がある問題であり、実際の研究活動の中での指導や経験を通じた習熟の機会を提供

することが重要である。こうした点からも、研究公正の取り組みを再検討することは、大学の研究室等を単位とする教育研究の指導環境や研究風土といった文脈を明らかにすることでもあり、研究活動全体の質を向上することにも繋がりと考えられる。

このように再現性問題への対応を研究公正の一側面と位置付け、研究集団として対応すべき事項だとするならば、各研究分野で研究成果を取りまとめる際に重視している事項を明らかにすることが有用である。その際、医学や心理学分野における研究の大半では、その実施に際して研究倫理審査委員会という第三者の視点での審査が求められていることに注目したい。当該委員会の目的には、後述するように、研究対象者への配慮、オープンデータ、透明性・客観性、利益相反管理といった事項を保証するだけでなく、研究者自身のバイアスを排除し、恣意的な計画変更なしに適切な形で研究を進めていくよう研究者に促すことも含まれており、再現性の確保に貢献しうるものである<sup>6)</sup>。したがって、こうした委員会による第三者的視点が見識が研究者の客観性や透明性の意識を喚起し、研究の再現性及び研究全体の質向上に寄与するか（寄与しうるか）を検討することも重要であろう。

## 2. Web アンケート調査の実施方法

### (1) 調査票の作成と研究倫理審査

本論の問題意識に関する基礎的調査の実施のため、広島大学高等教育研究開発センター研究倫理委員会にて研究課題「責任ある研究活動を実施するために必要な教育内容」として研究計画、調査票や同意文書等の審査を申請し、2022年2月1日に承認を受けた（審査番号2021審011）。研究の再現性に関する意識は「研究活動の実施に際して重要と思われる項目の認識」という設問で多側面から質問を行い、また「研究倫理審査委員会」については2つの設問を設定した。これらの概要を以下にまとめた。各設問の冒頭にある括弧書きは本論で使用する際の略記表記であり、回答者には示していない。また、各項目は順序尺度で回答を求めた。アンケートの単純集計結果全体は本論に先行する野内（2022）において掲載しており、こちらを参照されたい。

#### 研究活動の実施に際して重要と思われる項目の認識に関する設問（13項目及び自由記述）

[出典の明確性]：参照した先行研究・資料の典拠を明確にする。

[研究対象者の保護]：研究の対象に関する倫理面での配慮を考慮する。

[研究上のルールの明確性]：研究を実施する上で必要なルール（法律、指針等）を明確にする。

[データの管理]：研究データ\*を適切に管理・保管する。

\* 実験や観測、シミュレーション結果、アンケート調査結果、フィールドワークの記録、貴重な古文書・資料、インタビューで得られた内容（テキスト、音声）、撮影した映像・写真等)

[データの収集方法]：研究データの収集方法を明確にする。

[データの分析過程]：研究データの分析過程を明確にする。

[先行研究の再現性]：先行研究の再現性\*\*を確認する。

\*\* 同じデータと同じ方法で再解析する，新たなデータと同じ方法で追試する，新たなデータと違う方法で追試する等，条件を変えてもある現象や結果が生じること

[自身の先行研究の再現性]：自身の研究結果の再現性を確認する。

[研究ツールの知識やスキル]：研究に関連する機器，ソフトウェア，プログラム等を使用するための知識やスキルを習得する。

[情報共有]：共同研究者間での情報共有を心がける。

[議論の構造の明確化]：成果発表の際，議論の構造を明確にする。

[論証や解釈の飛躍の回避]：成果発表の際，論証や解釈に飛躍がないようにする。

[政府やアカデミアの動向]：研究分野に関連する，政府や学術団体の国内外の動向を知る。

#### 研究倫理審査委員会での審査が有用な項目（9項目）

[研究計画全体]，[アンケートの設問内容]，[インタビューの質問内容]，[研究対象者の選定基準]，[研究対象者の数]，[使用する統計手法を含む研究全体のデザイン]，[利益相反の有無]，[個人情報・プライバシーの管理]，[取得したデータの管理方法]

#### 研究倫理審査委員会の意義（7項目）

[客観性や透明]：外部の有識者が審査に加わっている。審査のプロセスが明確になっている。

[公平性]：審査に関する利益相反の有無が検討されている。審査結果に対して意見を述べる機会がある。審査結果が適切に開示されている。

[平等性・等質性]：審査委員の構成・専門性は適切である。審査にかかる時間は適切である。

[機会]：委員会の開催数，スケジュール等は適切である。

[効率性・迅速性]：申請をするためのシステムの操作は容易である。申請書を用意するための様式（記入例）がある。審査の必要性や内容を判断するための手引き（フローチャートなど）が提供されている。

[審査や規程の厳密性・実質性]：研究内容に沿った有益な意見を得られる。

[審査の必要性]：研究者の自律に基づいて研究を実施すべきである。

## (2) Web アンケートへの案内

アンケートは SurveyMonkey を使用した Web フォームで実施した。アンケートへの案内は、BMC 経営管理センターを通じて国内大学院の研究科長宛にハガキにて送付し、研究科に所属する研究者・大学院生にアンケート URL を展開してもらえるよう依頼した（送付先リストは『全国大学一覽』（文部科学省）をもとに作成）。また、回答者を大学院生以上に限定したのは、研究室内で研究指導を受け学会への参加・発表といった機会が出てくる世代だからである。アンケートへの誘導は初回の案内に加え、1回のリマインダーを送付した（研究機関担当者等からの連絡によりアンケート展開が確認できた場合や、初回案内時に改組等の情報をもらった研究科宛には一部送付を省略した）。

### (3) Web アンケートの回答結果

アンケート回答者の研究分野と身分・職階を表1と2にまとめた。回答者の研究分野は以前に日本学術振興会科学研究費助成事業で使用されていた分類で回収し、4分野に分類した。(欠損データの小数点の調整から割合の合計は100%にならない)。初回案内ハガキ郵送は1,923件(2022年2月17日)、その後のリマインダー送付は1,906件であった(2022年3月16日)。回答締切は2022年3月31日で、その期間内に Web 上で参加同意のあった回答は1,714件であった。ここには内容を確認するためのアクセスや同意後の拒否も含まれ、実質的な回答は1,172件、そのうち設問全てに回答があったのは1,023件であった。なお、以下の分析において、専門分野間での比較を行う際には人文学・社会科学・理工学・医歯薬学の4分類を用い、詳細が不明な「その他」の回答は除外した。

表1 回答者の研究分野

分野	件数	%
人文学	78	7.6
社会科学	104	10.2
理工学	456	44.7
医歯薬学	293	28.6
その他	92	9.0
合計	1023	100

表2 回答者の身分・職階

身分・職階	件数	%
大学院生(修士課程)	243	23.7
大学院生(博士課程)	153	14.9
PD・OD・研究生	11	1.1
助教	97	9.5
講師	55	5.4
准教授	157	15.3
教授	298	29.1
その他(名誉・客員教授等)	9	0.9
合計	1023	100

## 3. 研究の再現性に関する分析

### (1) 研究の再現性に関する大学院生と教員の意識の差

研究の再現性の意識について、各研究分野の大学院生と教員(ポスドク含)を比較するため、これら2群を対象としてマン・ホイットニーのU検定で分析し(有意水準5%)、標準化された検定統計量から効果量<sup>7)</sup>を算出した(表3)。以降も含め、分析にはIBM SPSS ver. 26を用いた。

分析の結果、[出典の明確性]( $Z = 2.032$ ,  $p = 0.042$ ,  $r = 0.06$ ), [先行研究の再現性]( $Z = -4.052$ ,  $p = 0.000$ ,  $r = 0.13$ ), [自身の先行研究の再現性]( $Z = 2.088$ ,  $p = 0.037$ ,  $r = 0.07$ ), [知識とスキル]( $Z = -3.770$ ,  $p = 0.000$ ,  $r = 0.12$ ), 「飛躍の回避」( $Z = 2.083$ ,

表3 大学院生と研究者の再現性意識の差

項目	Z	p	r
出典の明確性	2.032	.042*	.06
研究対象の保護	0.382	.703	-
研究上のルールの明確性	-1.632	.103	-
研究データの管理	-0.812	.417	-
研究データの収集方法	0.615	.538	-
研究データの分析過程	-0.043	.966	-
先行研究の再現性	-4.052	.000*	.13
自身の先行研究の再現性	-2.088	.037*	.07
研究ツールの知識やスキル	-3.770	.000*	.12
情報共有	1.820	.069	-
議論の構造の明確化	-0.509	.611	-
論証や解釈の飛躍の回避	2.083	.037*	.07
政府やアカデミアの動向	-0.230	.818	-

\*  $p < .05$

$p = 0.037$ ,  $r = 0.07$ ) に関して効果量は極めて小さかったものの有意な差が見られた。これらの中でも [先行研究の再現性] と [自身の研究の再現性] に焦点を当て、回答者の身分・職階ごとの違いをクラスカル・ウォリスの順位和検定を用いて分析した (有意水準5%)。その結果, [先行研究の再現性] について有意差がみられ ( $p = 0.001$ ), 分布を確認すると, 講師から教授の職階において低い評価への回答が多く見られた。また [自身の研究の再現性] の項目について, 有意差はなかったものの, 分布から「その他 (名誉教授等)」の職階においてのみ低い評価になっていた。

再現性問題は近年, 大学や学会の中でシンポジウムや教育プログラムの形で情報提供がなされる機会が増えており, 大学院生がその他の研究公正の内容と共に強く意識していることは考えられる。一方で, 今回の回答結果から即座に教員クラスの研究の再現性への意識が乏しいとは見做せないが, 大学院生との違いの理由は次のように考えられる。まず, 研究以外の業務が多くある研究者にとって, 研究活動の中で他者の再現性の確認に割くりリソースが少なかったり, 優先度が低かったりする (再現性を確認する論文それ自体はあまり評価されない)。もしくは, 研究者同士の信頼性及びリスペクトとして, そこに疑いを持たない。後者の点については, [自身の研究の再現性] について職階による違いが見られなく, それぞれが責任をもった研究発表しているという自負があると考えられることから推測できる。すなわち, 自身が適切に成果発表をしており, 当然, 他の研究者もそのようにしているだろうという互いの信頼関係のもと, 他人の成果を疑う根拠がないのである。

## (2) 研究の再現性に関する研究分野間の差

次に, 研究の再現性に関する意識について, クラスカル・ウォリスの順位和検定を用いて研究分野間で回答の分布を比較した (有意水準5%)。研究分野間の比較において研究者のみを対象とした場合と全ての回答者の場合とで検定結果に差が見られなかった。そのため各研究分野での指導者-被指導者間の関係性や意識の相違は同型であると見做し, 研究分野4分類間での比較結果のみを示す (表4)。

研究分野間全体で, [データの分析過程]「議論の構造」「飛躍の回避」の各項目には差がなかった。これらの項目は研究成果をまとめる際の基本事項としてどの研究分野でも重要視されていると見做せる。一方で残りの項目では研究分野間でその分布に有意な差が見られた。そこでボンフェローニ調整によって修正した結果から, より詳細にどのペアで差があったのかを確認した。

[出典の明確性] では, 人文学と社会科学はどちらも理工学・医歯薬学との間で差があった (人文学・社会科学での重要度が高い)。先行研究の内容や解釈, 記述に深く関わって自身の研究を進

表4 研究分野別の再現性への意識

項目	$r$
出典の明確性	.000*
研究対象の保護	.000*
研究上のルールの明確性	.000*
研究データの管理	.000*
研究データの収集方法	.014*
研究データの分析過程	.136
先行研究の再現性確認	.000*
自身の先行研究の再現性確認	.000*
研究ツールの知識やスキル	.000*
情報共有	.005*
議論の構造の明確化	.568
論証や解釈の飛躍の回避	.143
政府やアカデミアの動向	.002*

\*  $p < .05$

める研究と、実験で新たなデータを取得することを主題とする研究の差があったと考えられる。[研究対象の保護]では、理工学で他分野よりも重要度が低かった。この結果は当該分野においては人を全く対象としないか、医工連携研究以外では該当する研究が少ないという研究分野の特性から想定でき、回答者が適切に回答をした証左とも考えられる。次に、[ルールの明確性]は医歯薬学が社会科学・理工学に比して重要度が高く、[データの管理]は医歯薬学がどの分野に比しても重要度が高かった。この結果は、医歯薬学では倫理指針で人由来の資料・情報の厳格な管理が要求されていることが根拠だと考えられる。[データの収集方法]は理工学において他研究分野との差が見られた。[先行研究の再現性確認][自身の研究の再現性確認][研究ツールの知識とスキル]は理工学・医歯薬学で重要度が高く、人文学・社会科学間で差がなかった。再現性の項目について、設問の中で「研究データ」「再現性」という用語の意味を補足した。しかし、人文学・社会科学系でも数値データを扱う分野や心理学等の再現性への意識が高まっている分野を除き、回答者が自身の研究に関連が薄い設問だとみなした可能性はある。人文学・社会科学分野において[出典の明確性]への意識が他分野と比べて高く、「議論の構造」「飛躍の回避」には他分野との差がなかったことを考慮すると、当該分野において研究の公正性に関する意識が乏しいとは考えられない。また前述した論述の追跡可能性という観点は、むしろこれらの項目に収まるものとして考えることも可能である。すなわち、これらの側面を踏まえて改めて当該分野の特性に即した形で再現性という概念の定義を定式化する必要があると考えられる。

#### 4. 研究倫理審査委員会による審査とその意義に関する分析

##### (1) 人文学・社会科学系の研究における研究倫理審査の背景

人を対象とする研究を実施する際に人権保護等の観点から事前に研究倫理審査委員会での審査が必要である根拠は、個人情報保護法といった一般的な法律と、生命科学等に関する学術研究における法律・指針である。こうした倫理審査が不要な研究分野（全く人等を対象としない場合）と医学系の研究倫理指針が適用される一部の研究を除き、後述するように人文学・社会科学分野でも医学系研究とは全く別の側面での倫理的配慮が必要な場合は多々あると考えられるが、指針等の形での標準的なルールはない。そうした現状の中、国内の多くの人文学・社会科学系の大学・部局では自律的に規程を定め、その運用を行なっているようである（野内、2021）。

人文学・社会科学での研究倫理審査に係る基本は、「人を対象とする生命科学・医学系研究に関する倫理指針（2022年3月10日一部改正）」に共通すると考えられる。同指針には、1) 研究の社会的・学術的意義、2) 研究分野の特性に応じた科学的合理性の確保、3) 研究参加による利益と不利益のバランスの考量、4) 倫理審査委員会での審査の義務、5) インフォームド・コンセントの取得、6) 社会的に弱い立場にある者への配慮、7) 個人情報等の適切な管理、8) 研究の質及び透明性の確保という基本方針がある。人文学・社会科学系の研究では身体への直接的な侵襲は生じないと思われるが、過去のトラウマ経験を有する人や社会的マイノリティにある方への調査では参加者の心理面での負担を考慮し、調査の実施設計を検討する必要がある（調査への参加を知られるだけでそ

の人や地域の事情を暴露しうる。参加者募集の案内方法や調査の実施場所、調査項目の記述にも配慮する必要がある)。また、授業内でアンケートを実施する場合も、教員は生徒・学生から見て強い立場にあるため、教員側から明示的・非明示的の圧力がないように回答環境を設定する必要がある。その他、研究者にはメディア出演、団体の役職就任といった機会もありえ、経済的なもの以外の利益相反開示も必須である。医学系の研究倫理審査委員経験者らによる座談会では、前述した倫理指針の目的の8)を踏まえ、「研究倫理審査委員会は、データの透明性や客観性に関して、研究実施の中身への意見をしても良い」ことを強調している(青松他, 2014)。論文査読と異なり倫理審査委員会は研究の意義や設定した仮説等へ意見すべきではないが、研究の技術面への意見は研究対象者の保護や成果の公正性にも寄与しうる(例えば、統計分析の観点から過剰なサンプル数の設定を指摘することは参加者の不必要な参加を回避し、参加者の不利益のバランスを取ることになる)。

このように、人文学・社会科学の研究の実施に際しては、人権保護以外でも考慮すべき事項は多くある。「研究倫理指針がないから審査の申請は不要」ではなく、「倫理指針がなくとも審査の申請を検討する」こと、それこそが研究者の倫理的責任である。自身の理解と研究分野での慣行が常に標準的なものである保証はどこにもない。また多くの研究倫理審査委員会では、規程で対象としていない研究内容であっても、必要に応じて審査を受け付けているであろう<sup>8)</sup>。

ここで、研究の再現性と研究倫理審査委員会の関係も改めて考えたい。研究データの扱いに関しては、人文・社会科学系でも国際的に進められている「オープンサイエンス」の方向性を検討する必要があるし、研究倫理審査委員会が研究者側と研究実施に際した情報共有をすることで研究の公正性や客観性の意識が担保される。しかし、あらゆる研究において倫理審査を要するものでも、同委員会が再現性の責任を負うものでもない。研究の再現性を考慮する一義的な主体は、当該研究を実施した研究者に他ならない。ここで検討したいのは、研究の再現性に寄与する複数の要素の中で、あくまで一つの機能としての研究倫理審査委員会が果たしうる役割である。以下の分析では、医学系では指針に基づき倫理審査が課せられている中で研究者が倫理委員会にどんな意義を認めているかを確認し、それを下地として医学系以外の研究分野における研究倫理審査のあり方を模索する。

## (2) 研究倫理審査委員会への審査経験

まず、回答者が所属する大学・部局において、研究倫理審査の体制の有無を把握しているかをまとめた(表5)。大学院生と研究者で分けたのは、倫理審査に関する教育指導の状況を概観するためである。ここでは「不明」の回答比率の少なさが重要である。人文・社会科学系の回答者は、医歯薬学系と同程度に、審査体制の有無を把握している。研究者(教員)は倫理審査委員会の委員を任命されることもあり、研究者自

表5 研究倫理委員会の体制把握と身分

分野	身分	有	無	不明	計
人文学	院生	16	1	5	22
	研究者	44	4	8	56
社会科学	院生	22	0	7	29
	研究者	63	5	7	75
理工学	院生	89	15	87	191
	研究者	218	5	42	265
医歯薬学	院生	85	1	9	95
	研究者	187	2	9	198
その他	院生	44	4	11	59
	研究者	30	1	2	33
合計	院生	256	21	119	396
	研究者	542	17	68	627
		798	38	187	1023

身は委員会の設置等を把握していても、本人たちの研究には必要ないために大学院生に研究倫理審査の指導をしない場合は考えられるし、研究室単位で共同研究をする場合、教員が審査手続きを済ませて学生には審査委員会の詳細を知らせないという可能性もある。人文学・社会科学においては全く人を対象とせず、文献解釈的なタイプの研究だけをする分野も多くある。そのように当該分野に適用される明確な研究倫理指針が設定されていない中で所属機関の研究倫理審査を把握している状況は好ましいものと見做せる。

次に、研究倫理審査の申請経験の有無を分野別に集計した(表6)。申請経験がない理由は、「審査について指導を受けたことがなく、必要性を検討したことがない」が135件、「審査の必要が全くない研究内容(文献研究等)だから」が277件、「研究内容に倫理審査に関する事項が含まれていそうだが、そのような指導や求めを受けたことがない」が40件、「学内の規定やルールに照らして倫理審査の必要がないことを確認できた」が111件であった(複数回答可)。

これら理由について、客観的には倫理審査が必要だと思われるが本人の認識違いで不要だと判断している者がいることも否定できない。また、申請経験がない理由を「現在準備中」とした自由記述もあり、潜在的な研究倫理審査の申請経験者も申請経験のない者に含まれている。

以上から、領域特有の指針がない中でも人文学・社会科学系の研究倫理審査の体制が構築されつつあることが窺えた。しかし、研究者がその必要性をどこまで把握しているかは不明である。当該分野での研究実態も多様であり、実際に審査を要する研究者・学生に必要な情報の周知徹底がどの程度できているかをさらに調査・検討する必要があると考えられる。

### (3) 研究倫理審査委員会での審査が有用な項目

人文学・社会科学・医歯薬学の倫理審査経験者に限定し、「倫理審査委員会での審査をすることが有用な項目」に関する認識をクラスカル・ウォリスの順位和検定で分析した(有意水準5%) (表7)。その結果、[研究計画全体]、[研究対象の選定基準]、[研究対象の数]、[統計手法を含む研究全体のデザイン]において研究分野間で有意な差が見られた。ボンフェローニ調整によって修正した結果から、より詳細にどのペアに差があったのかを確認した。いずれの項目においても社会科学と医歯薬学のペア間に有意な差があった。

研究倫理審査において、アンケートとインタビューの内容の確認、利益相反の申告、個人情報保

表6 研究倫理審査委員会の申請経験

分野	ある	ない	計
人文学	22	56	78
社会科学	44	60	104
理工学	102	354	456
医歯薬学	223	70	293
その他	47	45	92
合計	438	585	1,023

表7 有用な審査項目の認識の差

項目	<i>p</i>
研究計画全体	.029*
アンケートの設問内容	.115
インタビューの質問内容	.118
研究対象者の選定基準	.006*
研究対象者の数	.001*
統計手法を含む研究全体のデザイン	.000*
利益相反の有無	.940
個人情報・プライバシーの管理	.967
取得したデータの管理方法	.852

*n* = 289 (人文22, 社会科学44, 医歯薬学223),

\* *p* < .05

護のための匿名化，取得したデータの管理方法については研究分野に関係なく，いわゆる研究参加者への配慮に関する基本的な審査項目として理解されていることが予想される。しかし研究デザインや対象者の数，統計手法の選択といった研究の科学性に関係し，それぞれの研究内容により近い項目については，社会科学系と医学系とで理解が異なっていた。

#### (4) 研究倫理審査委員会の意義

次に，同じく人文学・社会科学分野と医歯薬学分野において，研究倫理審査委員会の存在意義への認識の違いを分析した（表8）。その際，審査経験の無い回答者は「どちらでもない」を「分からない」と同義の選択肢として回答していることが考えられ，倫理審査経験者の回答に限定した。分析にはクラスカル・ウォリスの順位和検定を用い，回答の分布を比較した（有意水準5%）。その結果，[効率性・迅速性]を除くすべての項目において有意差が得られた。ペアごとの比較をボンフェローニ調整によって修正した結果から，より詳細にどのペアに差があったのかを確認したところ，差があったのはすべて社会科学と医歯薬学のペア間においてであった。分布を確認すると，いずれの項目においても社会科学では医歯薬学よりも低い評価で回答していた。

研究倫理審査の申請後，研究者は審査コメントに応じて内容の修正等を行い，承認を待つ必要がある。その意味で自由に研究を開始できないという歯痒さは分野間で共通していると考えられる。その他の項目について，各審査項目が設定されている理由や審査の必要性に関して申請者側と審査委員会側で共通理解がなければ，研究者が審査委員に任命されても委員間で提出すべき意見の方針が定まらない。その結果，委員は倫理審査委員会の審査範囲を超えた意見を提示したり，申請者に過度な修正対応を要求したりする。申請者はその結果を受けて倫理審査委員会での審査を研究実施の非本質的な手順とみなし，倫理審査委員会の意義に否定的な評価をすることは考えられる。

## 5. まとめと今後の展望

本調査では，限定的な回答者からの結果とはいえ，研究の再現性や研究倫理審査という研究活動の公正性や質の担保に関わる事項に関して探索的な洞察を与えた。研究の再現性問題は統計分析について主に議論がなされているが，研究成果を客観的に解釈できるかという問題として換言すれば，数値データを扱わない研究においても適用可能であると考えられる。しかしながら，「研究データ」や「再現性」という語で含意されるものが研究分野において多様であるため，本調査は十分な定義のもとに実施できていたとは言い難い。その適切な定義を検討した上で再調査する必要がある。

表8 倫理審査委員会の意義の認識の差

項目	<i>p</i>
客観性や透明	.000*
公平性	.007*
平等性・等質性	.022*
機会	.006*
効率性・迅速性	.219
審査や規程の厳密性・実質性	.018*
審査の必要性	.004*

*n* = 289 (人文22, 社会科学44, 医歯薬学223),  
\**p* < .05

次に研究倫理審査について、人文学・社会科学一般という括りで見れば、現状の体制に大きく不備があるとか不十分であるといった結果は得られなかった。しかしながら、社会学や心理学といった、人文学・社会科学系でも倫理審査の必要性の高い分野が全体の意識を底上げしている可能性はありうる。そのため、人文学・社会科学分野をより細分化した形で調査を実施することが必要であろう。

最後に、社会科学系と医歯薬学系で研究倫理審査委員会の審査項目の意義に異なる認識を持っている原因は、本調査で直接的には分析できなかった。しかし、医歯薬学系では他分野に比べて研究実施上のルールと科学性に重きをおいていたという分析結果を踏まえれば、指針等の形で明文化された標準的文書の有無がその原因の可能性として考えられる。そもそも研究倫理審査委員会は研究の実施を妨げるために存在しているのではなく、研究機関と研究者が責任を持って研究を実施できるようにするための仕組みである。そのため、研究倫理審査委員会の設置目的や、審査上のポイント（標準の実施手順やマニュアルの構築）など、分野固有の事情を考慮した議論の蓄積が重要であると考えられる。それがなければ、大学・部局・委員会ごとに蝸壺化した審査体制が生まれ、学会や学術誌で成果を発表する際に混乱する可能性もある。また、科学技術・イノベーション基本法（令和3年4月）では人文学・社会科学も新しく対象に含まれるようになり、他研究領域との学際的共同研究の機会や外部資金も増えている。したがって、所属機関や学協会の研究倫理審査体制が時宜や他分野の動向に即した内容になっているかを適宜確認することが重要であろう。

以上、本研究で課題として見出された人文学・社会科学系分野における再現性の含意、同分野における研究倫理指針が持つ意義については今後の調査によって明らかにしたい。

## 【謝辞】

本研究は RIHE 公募型研究 B03003, JSPS 科研費 JP 21K00029の助成を受けたものです。

(URL: [https://rihe.hiroshima-u.ac.jp/research\\_activities/international-joint/summary/](https://rihe.hiroshima-u.ac.jp/research_activities/international-joint/summary/))

## 【注】

- 1) 本稿のはじめに、1章、2章の一部は、発表済みの資料（野内，2022）に加筆・修正して一連の議論を構成した。発行元は当該資料のシリーズを「速報性を重視し暫定的にまとめて発信することを目的として企画された」とし、「査読無し」の「報告書」に位置付けており、これを基に大幅に追記・修正をした原稿を査読付き学会誌・研究誌等に投稿することは許容されている。
- 2) 例えば、『経済セミナー』（日本評論社）では2022年6・7月号 No. 726において特集「経済学と再現性問題」が組まれており、この問題に関する認識が拡大していることが窺える。
- 3) これと同内容の声明として、アメリカ歴史学会の専門職行動基準書がある。川邊咲子、亀田堯宙、後藤真（翻訳）（2021）AHA 専門職行動基準書（2019年改定版），<https://rchr.info/aha->

statement-on-standards-of-professional-conduct/ <2022年7月21日アクセス>。

- 4) つまり、研究におけるアブダクションのプロセスをどのように捉え、それと HARKing の関係性をどのように考えるかということである。この点については、井頭昌彦氏（一橋大学）との意見交換および氏が発表予定の文献の草稿を拝読する中で着想を得たものである。
- 5) 人文学における研究内容の特性を踏まえた再現性を検討した研究として Peels (2019) がある。
- 6) 1 - (1) にて言及した「事前登録」といった制度は、ジャーナル側での同様の「機能」だと言える。
- 7) ここでの効果量の計算は水本 (2011) の解説に基づく。
- 8) 戸ヶ里他 (2022) は、人文・社会科学系学術誌での人を対象とする研究倫理への対応は心理・教育系で若干見られるが、多くは研究機関側にその保証を一定程度委ねていると解釈している。

## 【参考文献】

- 青松棟吉・大谷尚・西城卓也 (2014) 「座談会：医学教育研究における研究倫理」『医学教育』45 (4), 249-274頁。
- 池田功毅・樋口匡貴・平石界・藤島喜嗣・三浦麻子 (2015) 「心理学研究は信頼できるか？－再現可能性をめぐる」『再ナビ！ブックレット Vol. 1』ちとせプレス。
- 加藤司 (2018) 「『パーソナリティ研究』の新たな挑戦－追試研究と事前登録研究の掲載について」『パーソナリティ研究』27 (2), 99-124頁。
- 国里愛彦 (2020) 「再現可能な心理学研究入門」『専修人間科学論集』心理学篇 Vol. 10 (1), 21-33頁。
- 戸ヶ里泰典他 (2022) 「国内人文・社会科学系学協会機関誌における人を対象とする研究倫理への対応の実態－学協会機関誌のウェブサイト上の情報提供内容より－」『放送大学研究年報』39, 1-6頁。
- 野内玲 (2021) 「大学における研究倫理・研究公正を考える－特に人文・社会系領域を中心として－」『広島大学高等教育研究開発センター編ディスカッションペーパーシリーズ』No.17。
- 野内玲 (2022) 「人文学・社会科学分野の研究の再現性に関する基礎調査」*Advancement of Higher Education Research: RIHE Monograph Series* Vol. 8, 広島大学高等教育研究開発センター。
- 平石界・中村大輝 (2021) 「心理学における再現性危機の10年」『科学哲学』54 (2), 27-50頁。
- 保城広至 (2015) 『歴史から理論を創造する方法 社会科学と歴史学を統合する』勁草書房。
- 水本篤・竹内理 (2011) 「効果量と検定力分析入門－統計的検定を正しく使うために－」『2010年度部会報告論集「より良い外国語教育のための方法」』外国語教育メディア学会関西支部メソドロジー研究部会, 47-73頁。(https://www.mizumot.com/method/mizumoto-takeuchi.pdf)
- Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, 100, 407-425.

- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science* 349 (6251): aac4716. doi:10.1126/science.aac4716.
- Goodman, S.N., Fanelli, D., & Ioannidis, J.P. (2016). What does research reproducibility mean? *Science Translational Medicine*, 8(341), 341ps12. doi:10.1126/scitranslmed.aaf5027.
- Ioannidis, J.P.A. (2005). Contradicted and Initially Stronger Effects in Highly Cited Clinical Research. *JAMA* 294(2), 218-228. doi:10.1001/jama.294.2.218.
- Nickerson, R. (1998). Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology*, 2, 175-220. doi:10.1037/1089-2680.2.2.175.
- Peels, R. (2019). Replicability and replication in the humanities. *Research Integrity and Peer Review* 4, 2. doi:10.1186/s41073-018-0060-4.

# Aspects of Research Integrity Reproducibility and Research Ethics Review Boards

Rei NOUCHI \*

The purpose of this paper is to report on the result of a cross-field survey of the reproducibility of research and the system of research ethics review boards. In medicine and psychology, the reproducibility problem, which includes a form of research misconduct, is addressed through approaches such as compliance with research protocols or appropriate use of statistical analysis. On the other hand, in research that does not use numerical data or statistical analysis, reproducibility issues can arise in the form of arbitrary arguments toward an intended conclusion. In this regard, there is not sufficient research and sharing of efforts in the humanities and social sciences in general.

To clarify these points, this paper presents an exploratory analysis of the results of a web-based questionnaire survey to examine the training and guidance systems in place for researchers and graduate students to promote research integrity, and how research ethics review boards can contribute to the management of research objectivity and transparency in terms of research ethics and its impact on research activities and integrity. The results were analyzed in an exploratory manner.

The results are as follows. First, it is necessary to reformulate the recognition of the issue of reproducibility of research in a manner specific to each area of humanities and social science. Second, in terms of ethical review processes in the fields of humanities and social sciences in general, the results of this study did not reveal any significant deficiencies or inadequacies in the current system. However, it is possible that the high need for ethical review in certain fields within the humanities and social sciences, such as sociology and psychology, may be elevating overall awareness. Finally, this study was unable to directly analyze the reason for the different perceptions of the significance of research ethical review committees in the fields of social sciences and medicine. The results do indicate, however, that the field of medicine places more emphasis on rules and it is possible that the presence of standardized guidelines may be a contributing factor.

---

\* Assistant Professor, School of Medicine, Shinshu University