

研究公正に関する態度・価値尺度の改変可能性の検討 —HIT-Res (How I Think about Research) の分析をもとに—

景山千愛

京都府立医科大学 医学生命倫理学 人文・社会科学教室 特別研究補助員

要約

近年、各研究機関で研究不正を防止するための教育が行われ、それらへの教育効果の評価の試みも行われつつある。しかし、現時点において、研究公正に関する教育評価の内容や方法にどのような種類があるのかについては、明確なコンセンサスが存在していない。本稿では、研究公正に関する効果を測定する尺度の分類について紹介し、その中でも研究公正に関する態度や価値観に関する尺度の重要性を述べる。そして、態度・価値尺度の一つである HIT-Res (How I Think about Research) を対象として、その内容や開発方法の特徴について述べ、それらの特徴を踏まえて HIT-Res の利用可能性について考察する。

1. はじめに

近年、研究不正への防止策・対応策として、学生や教員、研究者に対する研究倫理教育が、対面形式や e-ラーニング方式など、様々な形態で行われている。しかし、それらの講義や e-ラーニング教材が、どの程度個人の研究倫理への意識に影響を及ぼしているのか、定まった評価は存在していない。また、研究倫理教育の教育効果の評価基準も確立されているとはいえない。

現在のところ、研究公正に関する教育効果を測定する尺度の開発や検証が、複数の既存研究で行われている。また、それら開発の動きに伴って、研究公正の教育効果測定尺度が、教育効果のどの面を測定するものなのか、どのような場面に応じて使用するべきなのかについて、分類も行われ始めている。

本稿では、研究公正に対する測定尺度がどのように分類されるのか既存研究から示す。そして、その中でも態度や価値観を測定する尺度として HIT-Res (How I Think about Research) に着目する。HIT-Res に関する論文から、HIT-Res の測定尺度とし

ての利用可能性や特徴について分析し、それを踏まえて研究公正に関する態度・価値を測定するに際し、より実際の測定尺度の開発という観点から、HIT-Res の持つ課題について述べる。

2. 研究倫理教育の分類と HIT-Res の配置

研究公正教育の効果の測定と一口に言えども、その評価のポイントは様々に存在する。研究公正に関する知識を測定するのか、研究公正に関する態度を測定するのでは、前者と後者の質問項目は異なるであろう。

Antes et al. (2014) は、研究倫理教育において獲得されるべきスキルを、指導目標・方法・評価の対象という点から、4 つに分類している。研究不正の研究を専門とする Antes¹ によると、研究倫理教育においては、「研究倫理の知識 (Knowledge of Research Ethics)」のほか、「倫理的問題解決スキル (Ethical Problem-Solving Skill)」「倫理的感受性スキル (Ethical Sensitivity Skill)」、「態度と価値 (Attitudes and Values)」を対象としており、それぞれの特性に応じて到達目標や評価方法が考えられると述べている (表 1)。例えば、「倫理的感受性スキル」の到達目標は、「研究のデザイン化と実施における倫理的な問題を認識する能力を高める」ことにあり、「現実的なシナリオを提示した後に、シナリオ内の問題点を示すよう参加者に求める自由形式のプロンプトを表示する」評価方法が考えられる一方、「価値と態度」の到達目標は「研究倫理やコンプライアンスに対する建設的な態度を養う」ことであり、評価方法としては「簡潔な記述の後に、記述への同意または不同意を示すために、リッカー型²の尺度の回答を表示する」という形式が考えられるという (Antes et al 2014: 110-111)。

つまり、研究倫理教育の方法は、到達目標に応じて様々であり、それに対応する形で、適切な研究倫理の評価尺度も変わってくるといえる。Antes らは、倫理的問題解決スキルを測定する尺度として EDM (Ethical Decision-Making Measure) と PDR (Professional Decision-Making in Research Measure)²、倫理的感受性を測定する尺

¹ Division of General Medical Sciences, Washington University School of Medicine 所属(当時)。

² EDM、PDR とともに、倫理的に問題のある場面のシナリオを提示し、選択肢の中から適切と思われる選択肢を回答者が選ぶ方式。

表 1. 研究倫理の評価ポイントの4分類。Antes et al. 2014 (110=111) を簡略化したもの。

	倫理的問題解決スキル	倫理的感受性スキル	研究倫理の知識	態度・価値
到達目標	研究実施における倫理的な問題解決能力を養う。	研究の設計と実施における倫理的な問題を認識する能力を高める。	研究倫理に関する規定、方針、リソースを特定し、理解する。	研究倫理やコンプライアンスに対する建設的な態度を養う。
根拠	研究者は、倫理的、規制的、対人的側面を含む複雑な問題に直面する。意思決定の質を向上させるために、具体的な戦略を教えることができる。研究者は、問題解決に取り組むためには、倫理的な問題の存在を認識しなければならない。	研究者は、コンプライアンス、プロフェッショナリズム、およびより広範な対人関係の問題に対する感受性を必要とする場合もある。倫理的感受性のスキルは、研究倫理に関する知識、問題解決能力、態度と相互に関連している。	研究者は、研究事業の規則や規制に関する基礎知識を必要とする。この知識は、倫理的感受性と問題解決のための基礎を提供する。	倫理的行動を動機づけるためには、個人は RCR の重要性を理解し、研究の完全性を育成しなければならない。態度は、指導後の行動に影響を与え、動機づけと関与を通じて学習プロセスそのものに影響を与える。態度は価値観やバイアスと密接に関連しており、研究者はそれらやその影響を十分に認識していない場合がある。
想定される評価アプローチ	興味深く、関連性があり、魅力的なシナリオを提示。	現実的なシナリオを提示した後に、シナリオ内の問題点を示すよう参加者に求める自由形式のプロンプトを提示。	多肢選択式の項目で、最良の回答を1つ選ぶか、空欄を埋める。	簡潔な記述の後に、記述への同意または不同意を示すために、リッカート型の尺度の回答を表示。値のステートメントまたはランク順に並べることができる値の名前の提示・
指導法例	ケース分析の筆記、少人数および大人数のグループディスカッション、ロールプレイ、ビデオケース分析、学生が作成したケースライティング、オンライン/ビデオシミュレーション。	ケース分析の筆記、大小のケースディスカッション、ロールプレイ、ビデオケース分析、学生が作成したケースライティング。	朗読、情報提供型の講義、パワーポイントのスライド、質疑応答、小テスト(評定または無評定)、自主学習と研究、学生主導の講義/他の人への指導、個人またはグループのレポート、ワークシート、コンセプトマッピング。	視点に焦点を当てた講義、反省的な作文、討論、ディスカッション、ブログ、奉仕学習、ロールモデル、専門家以外の人(例:コミュニティのメンバー)との交流、自己評価/認識の練習、仲間からのフィードバック、絵や演技などの創造的な練習、他者へのインタビュー、映画、ストーリーテリングなど。

度として TESS (Test for Ethical Sensitivity in Science)³ 等、研究倫理の知識を測定する尺度として REKASA (Research Ethics Knowledge and Analytical Skills Assessment) 等、価値・態度を測定する尺度として HIT-Res 等を紹介している。各尺度の詳細は省くが、この分類は、研究倫理教育が、受講者の知識だけでなく、感受

³ 倫理的に問題のあるシナリオを提示し、回答者が、シナリオにおいてどのような問題があるかを記述する方式。

性や価値観等の内面的要素にも広く効果を及ぼしうることを示唆している。

本稿で検討する尺度である HIT-Res は、Antes らの論文での研究倫理教育の対象・評価等の 4 分類のうち、態度・価値尺度に位置する。態度・価値尺度の特性としては、スキルや知識といった可視的・外面的に評価される対象ではなく、より潜在的・内面的な要素を測定するという点があげられる。Antes らも「態度は、指導後の行動に影響を与え、動機づけと関与を通じて学習プロセスそのものに影響を与える。態度は価値観やバイアスと密接に関連しており、研究者はそれらやその影響を十分に認識していない場合がある」(Antes et al. 2014: 110) と述べているように、態度は人の認識に根差し、倫理的な意思決定や行動の指針となる点で、意思決定や行動、スキルの基底に位置していると考えられる⁴。また、この基底的・潜在的な性質は外形的に測定されるものではない。このため、態度・価値といった内因的要素は、技術や知識のように試験を行えば明確に達成度を評価できるという性質のものではないと考えられる。特に、自己評価的な性質をもつ、自記式の尺度であれば、わざと自分をよく見せるような虚偽的な回答や、事実を反映していない回答がなされる恐れがある。これらの点から、研究公正に関する態度・価値尺度は、他の尺度よりも、開発にあたって検討を要する尺度であると思われる。

3. HIT-Res のなりたちと開発

前節で、態度や価値観といった人間の基底的性質を測定するには、外面的評価よりも工夫が必要であろうことを述べた。本節では、態度・価値を測定する尺度の一例として「HIT-Res」を取り上げ、その質問項目から、態度や価値観の測定にあたって具体的にどのような工夫がされているのかを確認していく。また、HIT-Res が「HIT」尺度を改変した尺度であるという点に着目し、既存尺度を研究不正用尺度に応用する可能性について検討する。

まず、HIT-Res のなりたちについてみていく。

生命倫理学の専門家である DuBois⁵ らは、既存調査において研究不正の 40 ケース

⁴ 例えばスペンサーらが開発した人材評価であるコンピテンシーの氷山モデルでは、一般的に、人間の能力のうち可視的な部分として「スキル」、「知識」などを、目に見えない潜在的な部分として「価値観」、「動機」などを位置づけている (Spencer & Spencer 2011: 14)。

⁵ Division of General Medical Sciences, Washington University School of Medicine 所属

を検証したところ、自己中心的思考が要因のケースが48%にのぼった (DuBois et al. 2016: 966) ことを踏まえ、多くのノンコンプライアンスの事例は、研究者が認知のゆがみを用いたときに発生しているとの仮説をたてた。このノンコンプライアンスを下支えする認知のゆがみを測定するため、DuBoisらは、既存の認知のゆがみ測定尺度から、研究公正に関する認知のゆがみを測定する尺度を作成することを企画した。具体的には、反社会的行為に関する認知のゆがみを測定する、HIT (The How I Think Questionnaire) を応用することで、研究者の研究公正に関する態度・価値観を測定する HIT-Res を開発したのである。

本来、HITは、回答者の認知のゆがみから、非行との関連を測定するために開発された尺度である。HITは、4つの認知のゆがみ(悪いことを企てる、他人の侮蔑、矮小化/虚偽、自己中心的考え)をもとにした4行動(対立・反抗、物理的な攻撃性、うそつき、窃盗)についての設問を、計45問備えている。回答者は、それら文章に対して、自分がどれだけ同意するかを6段階のリッカート尺度で回答し、回答全体の平均値が高いほど、認知のゆがみが大きいことが示される。

HITの大きな特徴の一つに、自記式尺度につきものである虚偽的な回答や事実と異なる回答に起因する測定内容のゆがみを補正する仕掛けがある。具体的には、社会的に望ましい回答を測定するための変則的回答(Anonymous Response)アイテムと、回答者による誇張されたネガティブさを測定するポジティブフィラー(Positive Filler)アイテムが、ダミー質問として尺度に埋め込まれている。社会的望ましさを測定するAR質問では、「論文はオープンアクセスにするべきだ」(Publications should be open access)等8問から、その回答者がどれだけ自分を社会的に望ましく提示しているかを測定し、ネガティブな質問への同調程度を測定するPF質問では、「過去に、私は断りなく職場から何か持ち出したことがある」(In the past, I took something from work without asking)等7問である。これらのダミー質問への回答傾向から、回答者が社会的望ましさを反映した作為的な回答のゆがみを推測し、どれだけ実際と異なる回答をしているかによって、回答全体のゆがみを補正することができるのである。

先に述べたように、DuBoisら(2014)は、このHITからHIT-Resとして、研究公正に関する価値・態度尺度への改変を行っている。改変にあたっては、オリジナルが窃盗や詐欺に関する設問を行っていたところを、HIT-Resでは利益相反、被験者や

実験動物の保護、研究不正など研究関連の事柄に変更された（「誰もが利益相反を抱えているが、たいしたことではない」、「自分自身の研究にとって必要であれば、自分の学生に調査を完ぺきにするよう要求するだろう」、「私は自分の知識を他人と共有するときに幸せを感じる (PF 質問)」、「論文はオープンアクセスにするべきだ (PF 質問)」、「過去に、私は断りもなく仕事から抜け出したことがある (AR 質問)」などである。

DuBois らは試作尺度をもとにパイロットテストを行い⁶、因子分析を行うことで各設問文の妥当性を検証した。最終的な因子は、4つの認知のゆがみと PF 質問、AF 質問の3因子となった⁷。カテゴリーに応じたアイテム数は、悪いことを企てる (8 アイテム)、他人の侮辱 (9 アイテム)、矮小化／虚偽 (8 アイテム)、自己中心的考え (8 アイテム)、AR 質問 (6 アイテム)、PF 質問 (6 アイテム)、の45アイテムである (DuBois et al 2014: 977)。

また、パイロットテストにおいて、参加者に同時に実施した他尺度と HIT-Res との構成概念妥当性⁸が検証された。この結果、HIT-Res は道徳的非関与の傾向性尺度と強く相関しており、シニシズム尺度と正の相関があるなど、道徳的非関与やシニシズムと、研究公正に関する意識の低さに相関があることが示された。その一方、ナルシシズム尺度との相関は明白ではなかった。また、社会的望ましさを測定する AR スコアはマーロー＝クラウンの社会的望ましさの尺度と強く相関しており、HIT-Res において補助的に社会的望ましさを測定できることが明らかになった。併存的基準妥当性

⁶ 調査サンプルとして、キャリアステージや年齢、研究分野の多様な、NIH に資金提供された 300 人の研究者がリクルートされた。回答時間はおよそ 75–120 分であり、100 ドルの謝礼が支払われた。測定する概念をどの程度適切に反映しているかを検証するため、参加者たちは HIT-Res のほかに、道徳的非関与の傾向性尺度 (PMD) グローバルシニシズム尺度、ナルシシズム尺度を用いた。また、併存的妥当性 (外部の基準とどの程度関連があるかに関する妥当性) を調べるために PDR を用いた。さらに、回答の社会的望ましさによるコントロールを調べるために、マーロー＝クラウンの社会的望ましさの尺度を用いた。このほか、ジェンダーや年齢、研究歴や研究分野、母語など、プロフィールに関する質問もなされた。

⁷ ただし、HIT の開発においては、確認的因子分析で、6 因子と 3 因子の、2 つの異なるモデル (6 因子モデルは、4 つの認知のゆがみそれぞれと AR スケール、PF アイテムを単独の因子として扱う。3 因子モデルは、4 つの認知のゆがみをすべて 1 因子として扱う) で適合性を検証し、6 因子モデルが支持された。一方、HIT-Res でも確認的因子分析を行ったところ、6 因子モデルは受け入れがたく、3 因子モデルに適合した。HIT オリジナルは多様な因子によって決定された一方、HIT-Res は認知のゆがみが 1 因子として現れている。これは、オリジナルの HIT のデータと一貫していないことを示すわけではない。

⁸ ある評価法が、測定しようとする概念や特性をどれだけ適切に反映しているか検証するもの。

については、HIT-Res は PDR と負の相関があり、研究に関する倫理的意思決定において低いスコアを持つ者は、傾向として研究公正に関する意識が低いことが示された。

このように、HIT-Res の尺度としての利用可能性は、おおむね支持されているといえる。

4. HIT-Res の利用可能性

前節では、HIT から HIT-Res への改変事例を見てきた。本節では、既存尺度から応用的に研究公正に関する尺度を作る点、研究公正に関する態度・価値を図る尺度である点を対象にして、HIT-Res の利用可能性について述べる。

まず、HIT-Res の利点として、一般的な尺度作成と比べ、作成コストや手間が大幅に省かれる点がある。本来、尺度作成のプロセスは、構想段階、尺度素案を作成する段階、パイロットテストを行い尺度の妥当性を検証する段階と、コスト、時間、手間を費やす必要がある。それに対し既存尺度を改変する尺度開発においては、オリジナルの尺度をもとに質問文を改変するため、本来であれば尺度素案作成過程に多くの検討時間を費やすところを、短縮することができるであろう。

また、HIT-Res の場合、HIT の設問内容については手を加えず、オリジナルをそのまま研究公正に関する内容に置き換えることで、オリジナルの尺度開発において工夫が凝らされた部分である AR 質問や PF 質問を生かすことに成功している。このように、尺度作成において手間がかかる点をそのまま応用できるのが、尺度改変の持ち味といえるであろう。

検討が必要な点としては、改変した尺度であっても、オリジナルの尺度の持つ特性・尺度構成などに拘束されるという点がある。具体的には、HIT-Res の場合、オリジナルの尺度である HIT の構成に影響を受けるために、実際に測定する対象をうまくとらえることができなくなる可能性がある。HIT の尺度としての構成は、回答者の認知のゆがみの程度について、4つの認知のゆがみ（悪いことを企てる、他人の侮蔑、矮小化／虚偽、自己中心的考え）をもとにして、4行動（対立・反抗、物理的な攻撃性、うそつき、窃盗）に自分がどれだけ同意できるか問うものである。しかし、この尺度を改変する場合、4つの認知のゆがみに基づく4行動という点を固定して質問文を作らなければならない。つまり、研究不正に関する質問を4カテゴリーに限定して行わなければならないのである。HIT-Res では各項目を研究者向けに、それぞれ「利益相反」、

「被験者と実験動物の保護」、「研究不正」、「責任ある研究」に関するルールの違反とされているが、これら4つのカテゴリーはそれぞれ問題の幅や次元という点で同列に並べることができるものなのか、また、これら4つのカテゴリーに含まれていない研究コンプライアンスの問題があるのではないかについては、議論が必要であろう。

ただし、少なくとも DuBois らのパイロットテストにおいては HIT-Res の妥当性は支持されており、また実際に HIT-Res を活用した事例も存在している。DuBois らは、PI プログラム (The Professionalism and Integrity in Research Program)⁹ という、研究のルールや規制に違反した研究者のニーズに合わせた矯正プログラムにおいて、PDR と HIT-Res を被験者に回答してもらい、その結果について分析している (DuBois et al. 2016)。PI プログラムは3日にわたって行われ、プログラム参加者たちは、1日目のプログラム開講前と、3日目のプログラム終了後に、PDR と HIT-Res を回答し、データを得た。さらに、2014年に行った HIT-Res のパイロットテストで得た700人の研究者の RCR の知識尺度、ナルシシズム尺度、シニシズム尺度、道徳的非関与の回答データをサンプルとして分析を行った。

つまり2016年の調査の特徴は、研究におけるルール違反をした人を対象として、研究公正ワークショップの効果を検証したことである。この研究デザインは、研究公正に関する態度・価値観の変化を観察するだけでなく、HIT-Res という尺度そのものの利用可能性を検証するのにも適した設定となっている。

参加者たちのワークショップ前後のスコアを比較すると、ワークショップ後、HIT-Res のスコアは明らかに減少し (平均 3.0 から 2.6 へ)、PDR のスコアは増加している (13.2 から 14.7 へ)。興味深いことに、HIT-Res の AR スコアも減少している (平均 4.8 から 4.5 へ)。ワークショップ後に社会的望ましさに関するスコアが減少しているということから、参加者たちがワークショップ後に、自分を偽らない、より率直な態度となっているといえるのである。

HIT-Res はその尺度構成について検証の余地があるように思われるが、ワークショップにおける HIT-Res の検証から、実際の研修プログラムにおいて、その利用可能性に

⁹ PI プログラムは、2013年から2015年の間、米国の24機関の39研究者をトレーニングした。違反のうち最も一般的なのが、ラボでの監督の欠如、インフォームド・コンセントや雇用での違反、盗用、動物実験の取り扱いに関する違反である。違反の理由は、コンプライアンスの優先度の低さ、ルールの不十分な理解、非効率なコミュニケーションなどである (DuBois et al 2016)。

大きな問題はないであろうことが示されている。また、HIT-Res がオリジナルの尺度 HIT を改変することで、尺度作成に必要な時間やコストをカットできたことは、実際のワークショップ等での利用をより簡便にしているという点で、この尺度の利用可能性をさらに高めているといえよう。

5. 課題と展望

前節での HIT-Res の利用可能性の議論を踏まえて、今後、研究公正に関する価値・態度尺度を作成する場合に課題となる点について議論する。

まず、前節で述べたように、尺度改変はオリジナルの尺度の形式に則って行われるものである。このため、オリジナルの尺度構成に準じた時に、測定対象を捕捉できるかについての検討が必要である。HIT ではオリジナルでの下位尺度である「対立・反抗」、「物理的な攻撃性」、「うそつき」、「窃盗」が、HIT-Res ではそれぞれ「利益相反」、「被験者と実験動物の保護」、「研究不正」、「責任ある研究」に置き換えられた。このように、オリジナルの尺度構成を保持したまま、内容をどのように研究公正向きに置き換えるかについて検討が必要である。逆に言えば、何を測定したいかによってオリジナルの尺度としてどの尺度を選択するかも決まってくるといえる。

次に、研究公正に関する態度・価値尺度の構成概念について述べる。HIT-Res の構成概念妥当性の検証において、これまでは道徳的非関与やナルシズムに関する尺度との関連が検証されてきた。確かに、人の態度や価値観といったものは、人の内面に属するものであるから、同じく内面的要素を測定する尺度との関連を検証するのは妥当であろう。しかし、公正な研究活動は、研究者個人の態度や価値観だけでなく、研究者をめぐる組織環境と結びついていると考えられる。研究公正を尊重する組織風土であるか、上司や同輩が研究不正を行っていないかどうか、それを黙認していないかどうかといった組織風土は、研究者個人の研究公正に関する態度や価値観に影響を及ぼすと思われる。特に日本であれば、上下文化や同調圧力といった文化的背景が存在するだけに、環境的要素には配慮するべきであると思われる¹⁰。今後、研究公正に関する態度・価値尺度の検証においては、内面的要素を測定する尺度のみならず、研究環境に関する尺度も併せて問うことで、環境的要素と態度とのかかわりについて検証す

¹⁰ Antes らも、論文の目的上、環境的要素を測定する尺度についてはほとんど言及していないが、研究公正における環境的要素の重要性については認めている (Antes et al 2014: 115)。

ることが必要であろう。

また、研究公正に関する態度・価値観と、各研究者が所属する研究組織における風土の関連について新たな知見を得ることで、新たな尺度開発において知見を役立てることも考えられる。特に、組織風土と態度・価値観の関連を踏まえた教育評価尺度であれば、研究不正の原因探索時に行われる研究者の内因（出世欲、焦り等）と外因（雇用制度、組織風土等）という二分法を超えるものとなることが予測される。既存研究での尺度分類を超えた、オリジナリティある尺度開発のために、今後の調査・検証が期待される。

謝辞

本研究は、国立研究開発法人日本医療研究開発機構（AMED）の助成のもとに行われた。開示すべき潜在的利益相反はない。

また、第二期 AMED 研究公正高度化モデル開発支援事業「学際的アプローチによる研究倫理教育のモデル評価プログラムの開発と検証」（瀬戸山班）の研究代表者の瀬戸山晃一先生、研究分担者の渡邊洋子先生、今井浩二郎先生、河原直人先生からご助言をいただいた。記して感謝を申し上げる。

参考文献

- Antes, L. Alison., James M. DuBois, 2014, Aligning objectives and assessment in responsible conduct of research instruction, *JOURNAL OF MICROBIOLOGY & BIOLOGY EDUCATION*, 15 (12):108-116.
- DuBois, M James., John T. Chibnall, and John Gibbs, 2016, Compliance Disengagement in Research: Development and Validation of a New Measure. *Science and Engineering Ethics*, 22:965–988
- DuBois, M James., John T. Chibnall, Raymond Tait, and Jillon S. Vander Wal, 2018, The Professionalism and Integrity in Research Program: Description and Preliminary Outcomes. *Academic Medicine*, 93 (4): 586-592
- Spencer M Lyle., Signe M. Spencer, 1993, *Competence at Work: Models for superior performance*, John Wiley & Sons. (=梅津祐良、成田攻、横山哲夫、2011、『コンピテンシー・マネジメントの展開 監訳版』、生産性出版).