

研究公正に関する自己記述式尺度における質問文の検討

—尺度作成における議論を通して—

景山千愛¹、手良向聡²、吾妻知美³、山脇正永^{1,4}、渡邊洋子^{1,5}、松山琴音^{1,6}、
吉井健悟⁷、今井浩二郎⁸、河原直人^{1,9}、峯村優一^{1,10}、瀬戸山晃一¹

1 京都府立医科大学大学院医学研究科医学生命倫理学人文・社会科学教室

2 京都府立医科大学大学院医学研究科生物統計学教室

3 京都府立医科大学大学院保健看護学研究科

4 東京医科歯科大学臨床医学教育学開発分野

5 新潟大学創生学部

6 日本医科大学医療管理学

7 京都府立医科大学大学院医学研究科生命基礎数理学教室

8 京都府立医科大学大学院医学研究科医療フロンティア展開学

9 九州大学病院 ARO 次世代医療センター

10 群馬バース大学教養部

要約

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）では、研究活動に関する倫理的意思決定を測定する尺度を開発している。本論文では、この開発中の尺度の質問文に着目し、自己記述式尺度が抱える問題点を克服するため、質問文の作成において検討した内容や、この質問文のオリジナリティ、課題について述べる。

1. はじめに

近年、研究不正や不適切な研究活動の防止を目的として、研究倫理教育が、各研究機関で行われている。しかし、実際にそれらの研究倫理教育が受講者の不正行為の抑止策となっているのか、研究不正を行わないような動機付けや意識の変化が行われているかについての検証はほぼなされていない。また、研究倫理教育の教育効果を評価する標準的な基準自体、確立されていない。

このような現状を鑑み、瀬戸山晃一教授を代表とした、日本医療研究開発機構 (AMED) 平成 31 年度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」プロジェクト（以下、瀬戸山班と呼ぶ）では、研究倫理教育の効果測定を担う尺度の開発を行っている¹。

本稿では、瀬戸山班が開発・検証中の尺度を対象に、まずどのような先行研究や既存尺度を参考にしたのかを紹介する。次に、本尺度の質問文において、解答者の忖度をなるべく排除するために、どのような工夫を行ったのか、またそれらの工夫を行うにあたってどのような議論がなされたのかを示す。最後に、より信頼性の高い尺度開発という観点から、本尺度の質問文の持つ課題について述べる。

2. 先行研究：研究倫理に関する倫理的意思決定尺度

Antes et al. (2014) は、研究倫理教育の効果の測定について、その指導目標・方法・評価の対象という点から、4 種類に分類している。Antes らによると、研究倫理教育には、その獲得対称に応じて「研究倫理の知識 (Knowledge of Research Ethics)」、「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)」があり、それぞれの特性に応じて到達目標や評価方法が異なるという。

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。瀬戸山班で開発・検証中の尺度の一つは、Antes らの分類のうち「倫理的問題解決スキル」を測定する尺度にあたり、その到達目標は「研究実施における倫理的な問題解決能力を養う」ことである (Antes et al. 2014: 110=111)。Antes らは、倫理的問題解決スキルを測定する既存の研究倫理測定尺度として、EDM (Ethical Decision-Making Measure) と PDR (Professional Decision-Making in Research Measure) の 2 種類を挙げている。

EDM は Michael Mumford (米国オクラホマ大学、応用社会調査センター所属) を中心に開発された尺度であり、PDR は James DuBois (ワシントン大学セントルイス校、一般医学領域所属) らによって開発された。EDM の尺度の特徴の 1 つは、具体的な研究場面を描いたシナリオと場面を提示し、解答者に次に選ぶべき行動に近い選択肢を 2 つ選ばせることである。1 シナリオにつき 4 つの場面が描かれており、教育前を測定するプレテストに 5 シナリオ (20 場面)、教育後を測定するポストテストに 5 シナリオ (20 場面) が配置されている。

¹ プロジェクトの詳細は HP を参照のこと。 <https://amed.setoyama.jp/index.html>

PDR で用いられる言葉は EDM よりも易しく、プレテストとポストテストそれぞれにつき 4 シナリオ 16 場面、選択肢もそれぞれ 6 問ずつと簡素化されている。

EDM と PDR の大きな違いに、質問文がある。EDM では、シナリオに続き場面が提示された後、「What should Abrams do?」² (Mumford, et al. 2006: 344) といったように、選択肢から「取るべき行動」2 つを選ぶように質問文で促される。一方、PDR では「あなたが実際にこのような困難な状況に置かれた場合にとるであろう選択に最も近いもの 2 つを予測しなさい」(DuBois, et al. 2016: 402) とある。

ここまで、倫理的意思決定尺度の特徴を概観してきた。次章では、主に倫理的意思決定尺度の「質問文」に着目し、当プロジェクトで開発されてきた尺度がどのようにこれら既存尺度の利点を取り入れ、さらなる工夫を行ったかを説明する。

3. 本尺度における開発と検討課題——質問文に着目して——

本章では、実際に瀬戸山班がどのような尺度を作成し、作成にあたりどのような議論を重ね、改良を行ったか、質問文にフォーカスして述べる。

まず、研究員達が尺度の設問案の作成に取り掛かり始めたころの全体会議では、以下のような質問文案が提示されている。

(注：ジレンマ的状况を描いたシナリオを受けて)

「上記のような A 准教授の状況にあなたが置かれた場合、あなたが取るだろう行動に最も近い選択肢を 2 つ選んで教えてください。」³

このように、本尺度では、「最も良い」選択肢を問う EDM ではなく、「あなたが最も取りそうな」選択肢を選ばせる PDR やスキル尺度を踏襲することを試みた。その理由として、本尺度は、解答者が実際にとるであろう行動に最も近い解答を選ばせることを目的としていることが挙げられる。解答者に「望ましい」解答を選ばせても、実際にそのシナリオのような状況に置かれた場合、解答者が「望ましい」行動をとるとは限らないためである。

ただし、質問文において「望ましい」選択肢ではなく、現実「あなたが最も取り

² Abrams はこのシナリオの主人公

³ 峯村優一研究員 (当時) 作成のシナリオより抜粋

そんな」選択肢を選ぶように指示したとしても、解答者はより「望ましい」選択肢を選んでしまう可能性がある。従来、質問紙調査においては、解答者が意識的・無意識的によい選択肢を選んでしまうという、「社会的望ましきバイアス」⁴の存在が指摘されてきた。

社会的望ましきバイアスの存在を加味すると、質問文においてどのような選択肢を選ばせるのか明示するだけでは、解答者の実際の行動に近い選択肢を選ばせることは難しい。このため、質問文の指示内容に依存しないかたちで、解答者の実情に近い選択肢を選ばせる工夫が必要となる。質問文について検討を重ねる中で、各シナリオにつき以下の質問文を用いることが提案された。

「① あなたがこのような状況に置かれた場合、最も取りそうな行動を2つ選んでください。

②あなたの周りの人（上司、同僚等）がこのような状況に置かれた場合、最も取りそうな行動を2つ選んでください。」（第18回全体会議議事録より）

同一シナリオに対して質問文を2つ設けており、1つ目の質問文は「あなたが最も取りそうな行動」、2つ目の質問文は、「あなたの周りの人（上司、同僚等）が最も取りそうな行動」を問う。

「あなた自身」だけでなく「あなたの周りの人（上司、同僚等）」について問うこととしたのは、解答者の社会的望ましきバイアスに影響されずに、組織に対する認識を明らかにすることを可能とするためである。人は自分自身には甘い評価をしがちであるとしても、「周りの人」の行動は冷静に評価できると考えられる⁵。

上記のように、本尺度において同一シナリオに2種類の質問文を設定することが確定したのち、2020年度末に、計3回の査読会議が開かれた。査読委員会は、医学研究者や研究公正の専門家、心理学者ら9名で組織されており、本尺度について多面的な検討が行われた⁶。査読委員会の中で、査読委員からこの質問文について、以下のよ

⁴ 社会的望ましきバイアスとは、実際よりも望ましい解答をしてしまうこと。他人を欺こうとする意図的な行為ではなく、本人が自分をどう認識（あるいは誤認）しているかの反映と考えられている（Strainer et al. 2015=2016: 102）。

⁵ このアイデアは瀬戸山晃一教授によるものである。

⁶ 査読委員会のメンバーについては謝辞参照のこと。

うな指摘がなされた。

査読委員 A⁷：現在の質問紙で、自分の考えと周りの人の考えを分けて聞いているが、その合致度のずれを検証するという方法はあると思う。また、現在は「あなたの周りの人（上司、同僚等）がこのような状況に置かれた場合、最も取りそうな行動を2つ選んでください。」としているが、上司と同僚は分けて聞か、どちらかに特定したほうがよい。

査読委員 B：上司と同僚の位置づけは、特に医学部の場合異なる。「あなたは、どう行動すると思うか」で聞くのと、「あなたの周りの人はどう行動すると思いますか」とで聞くのでは、かなり異なる解答が出てくることが予測される。

（2021年3月19日 第2回査読委員会 検討内容より）

査読委員 C：認知領域の問題では、「あなたの周りの人」について問う必要はないのではないか。また、上司から圧力をかけられているような内容のシナリオがあるが、その場合に混乱するのでは。

研究代表者：「あなたの組織の大半の人は」という聞き方はいかがか。

査読委員 A：一つの手だと思う。「あなたの周りの人」について、例えば「上司」など特定すると、解答者の周りの人が限定される。また、学生さんが答えにくくなる。

（2021年3月19日 第2回査読委員会 検討内容より）

ここでは、主に2つの指摘がなされている。まず、選択肢について「あなた」と「あなたの周り」で分けて聞くことで、それぞれの解答のずれを把握することができ、何らかの検証に使用できる可能性があるということである。

もう一つは、医学部文化の特有性を鑑みた場合の質問文の問題についてである。現在の質問文では「あなたの周りの人（上司、同僚等）」が取りそうな行動を問うている。しかし、一般の研究組織よりも上下関係の強い医学部においては、上司と同僚の立場は質的に異なるということである。また、解答者が学生の場合、研究者と異なり「上司」や「同僚」が存在していない場合もあるため、解答しにくいという指摘もされていた。

⁷ 査読委員の先生方の了承を得ていないため、本原稿では匿名化する。

4. 本尺度の質問文の可能性と課題

前章での本尺度の質問文に関する議論を踏まえて、本尺度について利用可能性と課題について述べる。

まず、査読委員からのコメントにあった「あなた」と「あなたの周り」について解答やスコアの差があったとして、それをどのように解釈すべきかという課題が存在する。例えば、「あなたが最も取りそうな行動」と「あなたの周りの人（上司、同僚等）が最も取りそうな行動」について、それぞれ選択肢やスコアが同程度である場合は、解答者自身と解答者の周囲について、倫理的意思決定のレベルが同程度であることが推察できる。しかし、例えば「あなたが最も取りそうな行動」の方で得られたスコアが、「あなたの周りの人（上司、同僚等）が最も取りそうな行動」についてのスコアよりも高いあるいは低い場合、その結果をどのように解釈すべきなのか。一案として、「あなたが最も取りそうな行動」と「あなたの周りの人（上司、同僚等）が最も取りそうな行動」のスコアの差から、解答者が研究組織の慣習に染まっているかどうかの手がかりとすることが考えられる。教育学の分野においては、公式的に教授される内容だけでなく、ある知識や意識が非明示的・無意識的に伝達される現象について「隠れたカリキュラム」と呼ばれ、医学教育の分野においても応用がなされてきた（板井 2012 等）。隠れたカリキュラムの効果は、公式的なカリキュラムよりも影響が大きいと言われており（大生 2011）、研究公正に関する組織的なポリシーが設定されていたとしても、ポリシーを軽視した行動・ポリシーと乖離した行動が上司から部下に継承されることで、非公式的に非倫理的な組織環境が生まれている可能性がある⁸。

こうした組織的に広がる「隠れたカリキュラム」を明らかにするため、その組織に所属して間もない人を「隠れたカリキュラム」に染まっていない人として着眼し、「あなた」と「あなたの周り」のスコア差を分析するという方法が考えられる。例えば、その組織に所属して間もない人について「あなた」より「あなたの周り」のほうが高得点の場合、よい隠れたカリキュラムに染まっていない可能性があり、逆に「あなたの周り」より「あなた」のほうが高得点であった場合、悪しき隠れたカリキュラムに染まっていない可能性が読み取れる⁹。

⁸ 例えば、研究や教育における指導者が、学生の前で公式的ポリシーは理想論に過ぎないとほめかすような行動などが考えられる。

⁹ ただし、「社会的望ましきバイアス」の影響を考慮する必要がある。

最後に、この質問文の課題について述べる。まず、上述したように、自己記述式の質問紙においては、質問文を工夫したとしても「社会的望ましきバイアス」が入り込みやすいという制約が存在する。このため、2つの質問文のスコアの差によって、本当に組織環境を把握できるのかが不明である。

一方、本尺度の質問紙で考えられる利点として、解答者のリファレンスポイントが固定されやすいという点が考えられる。解答者は、自分の所属する組織を参照点とすることで、「あなたの周りの人」と比較して自分について解答することになる。このため、同一組織内での得点差の比較を行う場合、ある程度の客観性は担保されると考えられる。この場合、同程度の資質を持つ研究者であっても、その研究者が倫理的に優れた環境に置かれていた場合は得点が相対的に低く、一方倫理的に劣った環境に置かれていた場合は得点が高くなってしまふと考えられるため、組織間での検証には注意が必要である。

本尺度の課題は、自己記述式質問紙の持つ限界に起因していた。一方、本尺度の質問文は、自己記述式質問紙につきものである「社会的望ましきバイアス」をある程度克服できる可能性を備えている。ただし、この可能性は、現時点ではあくまで推測であるため、実際に本尺度を用いた実証を行う必要がある。

謝辞

本研究は、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」の成果の一部である。

査読委員の先生方は、岸太一先生（京都橘大学）、楠見孝先生（京都大学）、大門貴志先生（兵庫医科大学）、中村征樹先生（大阪大学）、錦織宏先生（名古屋大学）、野内玲先生（信州大学）、松浦晃洋先生（藤田医科大学）、山下紀子先生（国立がん研究センター）、山本洋一先生（大阪大学）である（順不同。ご所属は2021年3月現在）。記してお礼申し上げます。

利益相反

本研究は、国立研究開発法人日本医療研究開発機構（AMED）研究公正高度化モデル開発支援事業（平成31年度～令和3年度）の助成による研究開発プロジェクト「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」のもとに

行われた。開示すべき潜在的利益相反はない。

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.
- 板井孝孝郎, 2012, 「プロフェッショナリズム教育と, その実践の根底にあるもの —— 『隠れたカリキュラム hidden curriculum』」, 『日本内科学会雑誌』 101 (1): 201-205.
- Mumford, Michael D., Lynn D. Devenport, Ryan P. Brown, Shane Connelly, Stephen T. Murphy, Jason H. Hill, and Alison L. Antes, 2006, “Validation of Ethical Decision Making Measures”, *Ethics and Behavior*, 16 (4), 319-345.
- 大生定義, 2011, 「医学教育とプロフェッショナリズム」『日本医科大学医学会雑誌』 7 (3): 121-128.
- Strainer L. David, Geoffrey R. Norman, John Cairney, 2015, “Health Measurement Scales: A Practical Guide to Their Development and Use” , Fifth Edition, Oxford University Press. (= 木原雅子, 加治正行, 木原正博訳, 2016. 『医学的測定尺度の理論と応用——妥当性、信頼性から G 理論、項目反応理論まで』 メディカル・サイエンス・インターナショナル).