

Lub, V., 2015, "Validity in Qualitative Evaluation: Linking Purposes, Paradigms, and Perspectives," *International Journal of Qualitative Methods*, 14(5): 1-8.

(V. ルーブ, 2015, 「質的評価における妥当性——目的・パラダイム・視点をつなぐ」)

序文 (pp.1-2)

- 近年, 社会政策やヘルスケアにおける評価を目的として, 質的な手法や情報がますます用いられている。
 - その理由として, 「何が有効か (“what works”)」を探るにあたって実験や質問票にもとづく量的研究がもつ限界に関する理解や認識が進んだこと, 評価における倫理的配慮への要求の高まり, クライアントや患者のウェルビーイングについての洞察, 根拠にもとづく (evidence-based) 介入が人びとの情動や文化, 経験, 習慣とどのように結びつくのかをより綿密に理解することの必要性などがあげられる。
- しかし, ヘルスケアや社会政策の領域の根拠にもとづくという枠組みにおける質的な情報の重要性の高まりは, たんなる学問的な考察の域を超えて, 妥当性の規準 (validity criteria) のより厳密な概念化を要求する。
 - というのも, 量的な効果研究が妥当性基準 (validity standard) を必要とするように, 質的な情報にもとづく政策判断は妥当な研究によって正当化されなければならないという主張が可能だからである。
 - しかし, このような研究の妥当性をどのように定めるかは難しい問題である。質的研究における妥当性は, 方法論に関する文献において多くの論者によって考察されてきたが, 評価との関連性についてはこれまであまり探求されてこなかった。
- 本稿では, 質的研究における妥当性の側面について, それを社会政策における評価の側面と接続させるという明確な目的をもって検討される。
- 評価における評価者とステークホルダーとの関係性, および質的研究の方法論的な特性を考えれば, 評価における質的な情報は3つの異なる目的をもちうる。
 - ①質的な情報は, プログラムそれ自体の道具的な有効性 (instrumental effectiveness) に貢献もしくは焦点を合わせることができる (例: 10 代の妊婦のための支援プログラムにおける意図された効果は現場で観察されているかという問いへの貢献/焦点化)。
 - ②質的研究は, クライアントや対象集団, 実践者にとってある政策もしくはプログラムがもつ意味に焦点を合わせることができる (例: 10 代の人びとは支援プログラムをどのように経験しているかという問いへの焦点化)。

- ③質的評価は、解放のためのアプローチ（＝調査から得られる情報がプログラムに関与する人びとをエンパワーもしくは教育することを目的とするアプローチ）をとることができる（例：10代の妊婦自身は評価のなかで収集された情報によって利益を得ているかという問いの研究）。
- 以下において著者は、社会政策やヘルスケアにおける質的評価の異なる目的が、それぞれ異なる科学的パラダイムや視点と関連づけられうると同時に、適切な（relevant）妥当性の手続きと結びつけられうることを主張する。
 - はじめに、質的な探究と妥当性についてのより一般的な議論について論じる。

質的研究と妥当性の問題（pp.2-3）

- 1970年代ごろからの質的研究の信頼性や客観性に対する批判の高まりは、より厳密な規準や方法論的基準を打ち立てることへの関心を増大させた。
- 妥当性は、この議論において重要な概念である。
 - 科学方法論の実証主義的・合理的（rational）な伝統において、「妥当性」は、研究上の概念の指標もしくは変数が、測定可能かつ当の概念を正確に表すものとなっている程度として定義できる。
 - このような妥当性の定義は、質的で自然主義的（naturalistic）な研究にはあまり有効ではない。結果として質的方法論の文献では、「妥当性」は真正性（authenticity）、適切性（adequacy）、もっともらしさ（plausibility）、中立性（neutrality）といった代替りのラベルを付されてきた。
 - にもかかわらず、学問コミュニティの内部では、質的研究者は何らかの方法でみずからの研究結果が妥当なものであることを論証しなければならないという考えが支配的なようである。そのため、質的な研究成果の妥当性を高めることを目的として、研究の手続きや規準を開発しようとした者もいる。
- おそらくもっとも影響力があるのは、GubaとLincolnの業績¹。
 - 2人の議論は次のような前提から出発する。すなわち、すべての研究は「高い真理値」（high truth value）を備えていなければならないが、「合理的」（もしくは量的）なパラダイムにおける知識の特性は、「自然主義的」（もしくは質的）パラダイムの知識の特性とは異なるという。
 - GubaとLincolnによれば、それぞれのパラダイムは研究の真実性（veracity）を決定するための明確な規準を要求する。合理的なパラダイムでは、内的妥当性、外的妥当性、信頼性、客観性という角度から規準が定式化できるという。一方、自然主

¹ Guba, E. G. and Lincoln, Y. S., 1981, *Effective Evaluation: Improving the Usefulness of Evaluation Results through Responsive and naturalistic Approaches*. San Francisco, CA: Jossey-Bass; Lincoln, Y. S. and Guba, E. G., 1985, *Naturalistic Inquiry*, Thousand Oaks, CA: Sage.

義的なパラダイムでは、「信憑性 (credibility)」「適合性 (fittingness)」「確認可能性 (conformability)」といった規準を用いるほうがよいとする (のちに「信憑性」「転用可能性 (transferability)」「確実性 (dependability)」と再定義される)。

- Guba と Lincoln が概念化した定評ある研究手続きとは以下のようなもの。
 - 消極的事例の選択：外れ値 (消極的事例) を意識的に探し出し説明することでデータの解釈を拡張させる。
 - 同業者への報告 (peer debriefing)：研究プロセスについての外部評価の一形態であり、同業者から研究の手続き、意義、解釈、結論について質問を受ける。
 - フィールドにおける長期にわたる参加と観察：研究対象を適切に描写するために十分な期間にわたって研究を遂行する。
 - 監査証跡 (audit trails)：研究者は研究のプロセスやそこでの選択を入念かつ年月の順にしたがって記録する。
 - メンバーチェック：収集したデータや用意されたカテゴリー、解釈、結論などに対するインフォーマントもしくは参加者からの体系的なフィードバックを受ける。
- 依然、多くの質的研究者は Guba と Lincoln の規準を方法論的な標準とみなしている。

質的研究における妥当性基準への批判 (pp3-4)

- 妥当性をめぐる議論を進めることに努める向きがある反面、質的研究のために前もって定められた規準を設けることの望ましさを否定する論者もいる。
 - Sandelowski と Baaroso (2002) は、質的手法の認識論的な幅が統一されたひと揃いの規準によって代表させるにはあまりに広すぎるという視点から、プロジェクトの質は研究ごとに別々に確定されなければならないと主張する。
 - Rolf (2006) は、質的研究が単一の科学的パラダイムを当てにできず、それゆえ質的な規準についての合意に到達することを目指すあらゆる試みはほとんど見込みがないと指摘する。総体として「質的研究」と記述できるような質的な理論もしくは方法論の領域についての共通理解は存在しないという。
 - Hammersley (2007) は、質的アプローチのなかには、知識の産出が研究の唯一の目的であるとの考えを明確に否定し、政治的な「アクション」を主張するものがあることを指摘し、このようなアプローチが研究の質を評価するうえでの従来とは異なる規準をもたらすと強調する。同時に、普遍的な規準を定式化することの難しさは、研究における価値前提の差異にも由来すると指摘する。社会的領域 (social domain) における研究は、そこに重要な差異をもたらさうる価値前提によって枠づけられており、研究の基礎をなす前提についての共有がなされていないほど、その研究の適切さを擁護することは難しくなり、妥当性の規準に関する合意に到達する

ことも困難になるという。ただし、「ガイドライン」の形式をとる規準は質的研究のより厳密な評価に役立つとも述べる。

- このほか、妥当性の規準をめぐる議論が質的研究の倫理の側面に注意を向けていないと指摘する論者もいる。質的研究の妥当性は、研究者の注意深さや共感、慎重さ、感受性といった角度から定式化されるべきだと主張される。

質的研究における妥当性のモデル：目的・パラダイム・視点をつなぐ（pp.4-6）

- 質的研究における妥当性の問題に対する立場は、主としてどの科学的パラダイムを支持するのかに依存しており、それに応じて前もって定められた規準の望ましさを否定する論者がいる。しかし同時に、異なるパラダイムは異なる規準を要求するのだと主張することもできる。
- Creswell と Miller は、研究者による研究手続きの選択は、研究者の前提となるパラダイム（paradigm assumption）と、研究者がみずからの研究の正当性を証明するために用いるレンズ（lens）という 2 つによって主に規定されると述べる²。
 - Creswell と Miller はこの 2 つを軸に、適切な妥当性確認の手続き（validity procedure）を識別するための枠組みを構築した。
 - この枠組みによれば、以下 3 つの競合するパラダイムが、質的研究に対する認識論的な立場を形成する主要なものとされる。
 - ①ポスト実証主義：質的研究は、量的研究と同様、体系的であると同時に厳密な手法によって構成されなければならないとみなす。このパラダイム内部では、量的研究コミュニティにおける厳格な方法論的作法に相当するものを質的研究においても探ろうとするという。
 - ②構成主義（constructivism）：現実についてのより多元的、解釈主義的、文脈依存的な視点をとる（すなわち時間や空間や状況に対して敏感になる）とされる。このパラダイム内部の手続きでは、「外的妥当性」ではなく転用可能性を用いるなど）妥当性の代わりとなる語彙が探られるという。
 - ③批判的パラダイム：近代社会の構造的不平等や権力関係への批判として出てきた視点であり、社会から取り残された集団のエンパワメントに取り組む質的研究者のあいだで受け入れられてきたという。研究の妥当性はステークホルダーや参加者とともに継続的に批判・交渉されるべきであり、研究者はみずからが開示する知識の性質について反省的かつ包み隠さない態度をとるべきだとされるという。

² Creswell, J. W. and Miller, D. L., 2000, "Determining Validity in Qualitative Inquiry," *Theory into Practice*, 39: 124-130.

- Cresswell と Miller は、上記のような前提となるパラダイムにくわえて、質的研究の妥当性を評価する際の 3 つの異なる視点＝レンズ（①研究者自身、②研究参加者、③外部のレビュアーもしくは読者）に沿って、9 つの異なる妥当性確認の手続きを同定する（次ページの表 1）。
 - メンバーチェック、監査証跡、長期にわたる参加、同業者への報告、無効とする証拠（disconfirming evidence）＝消極的な事例選択については、Guba と Lincoln が論じたものと同じ。
 - 三角測量（triangulation）とは、研究者がみずからのカテゴリーや結論の基礎を異なる情報源に置くという妥当性確認の手続き。
 - 反省性（reflexivity）とは、研究者がみずからの個人的な価値や信念を、それがどれほど研究結果に影響を与えたのかが明確になるようなやり方で報告書に明示する程度を表す。
 - 厚い記述（thick description）とは、研究の背景や参加者、主題についての詳細な記述を含むことであり、その目的は、読者を研究対象である世界やその主要な特徴に引き込むような出来事の記述を生み出すことにある。
 - 共同研究（collaboration）とは、参加者が共同研究者として（もしくはより形式的ではない関係で）研究に携わるべきという規準。
- 著者によれば、Cresswell と Miller の枠組みは、質的評価における妥当性の新たなモデルの基礎を提供する。
 - 先述のように、質的評価は 3 つの異なる目的（道具的な有効性／クライアントらにとっての意味／解放）をもちうる。著者は、これら評価の目的を Cresswell と Miller が区別した前提となるパラダイムと関連づけることができると述べる。
 - 第一の目的はポスト実証主義に対応する。ポスト実証主義の世界観では、ある特定の社会プログラム／政策を、その独立した効果を（「道具」として）評価することのできる別個の実体とみなす。
 - 第二の目的は構成主義に対応する。政策／プログラムに関わる人びとによって構成される政策／プログラムの実施や機能の多元的な現実を明らかにすることが目指される。
 - 第三の目的は批判的パラダイムに対応する。クライアントや対象集団の教育や社会進出にくわえて、評価に関わる研究者と回答者との協同が強調される。
 - 評価の目的とパラダイムを関連づけることで、質的評価のための適切な妥当性の規準の新たなモデルをつくりだすことができると著者はいう（次ページの表 2）。
 - 政策もしくはプログラムの道具的な有効性に焦点化する質的評価の場合、三角測量、メンバーチェック、監査証跡の規準が重要。これらの規準は、政策やプログラムの道具的な有効性を評価する際の歪みとなる可能性のある誤った（因果）推論やバイアスを回避もしくは発見するのにもっとも適している。

- 外部のレビュアーの視点からは、示された知見がデータによって支持されるか、プログラムの有効性についての結論は十分な根拠をもっているかを確認することのできるよう、監査証跡を行うことが重要（例：10 代の人びとが学校にとどまることを保証するという意図された効果を本当に果たしているのか、を問う）。
- なお、著者による新しい評価のモデルは、Cresswell らの枠組みを補足するいくつかの手続きを含む。
 - ほかのタイプの研究や既存研究から得られた知見と評価結果とを対比させることで、三角測量はよりよいものとなる。
 - 公正な取引（fair dealing）は、多くの異なる視点が評価において網羅されていることを保証する（例：10 代の妊婦だけでなく、専門家や家族構成員など、彼女らの周囲にいる関連する集団をインタビューする）。

結論と考察（pp.6-7）

- 本稿で提示された枠組みは、質的評価のチェックリストの役割を果たしうる。
 - しかし、この枠組みの主たる有用性は、（理論的な）参照点としての役割を果たしうる点にある。この枠組みの重要な特徴は、パラダイムもしくは認識論的な意味で「どれか一方の肩をもつ」ことを回避し、質的評価の異なる目的、パラダイム、視点を考慮に入れるプラグマティックなアプローチを提供していることにある。
 - このようにして、本稿のモデルは、柔軟性を厳密さと結びつけるだけでなく、パラダイムの違いを尊重すると同時に「共通のガイドライン」に対する要求に答えることができる。