

〔原 著〕

質的研究の質と評価基準について*

能 智 正 博

はじめに

質的研究 qualitative research とは定性的研究とも言われ、日常言語の特性を生かしながら対象を分析して新たな知を生み出していこうとする研究活動の総称である。1980年代以降、質的研究は主に社会科学の領域で広がりを見せてきたが、近年心理学においても関心が高まり、心理学の研究者の間でもそれを試みる人が増加している。心理学に特化した質的研究の概説書や論文集が出版され、心理学研究法の教科書の一章として質的研究が扱われるようになったのも、関心の高まりの一端を示すものであろう（南風原・市川・下山、2001；無藤他、2004；Willig, 2001 上淵他訳、2003）。さらに、2002年にはわが国でも学術誌の性格をもつ「質的心理学研究」が発刊され、2004年には日本質的心理学会が結成されるなど、心理学における質的研究の発表の場、研究者間の交流の場も着実に増加しつつある。

このように質的研究が広がり示すなかで、よりよい質的研究を行っていくためにはどうすればいいのかといった議論も起こっている（例えば、下山・子安、2002）。そうした議論の焦点の1つとなっているのが質的研究の質をどのように判断すればよいかという問題である。自分で質的研究を試みる際にも、どうすればよりいっそう質の高い質的研究に仕上げていくことができるのかどうかという点は、非常に気になるところであろう。量的研究では、あとで述べるように、いくつかの前提のもとで比較的少数の質の基準を定義し、それをもとに評価をしている。それに対して質的研究の場合にはより多数の観点の基準が提案されており、しかも研究者によって用語が異なっていたり強調点が

異なっていたりして、全体像をつかむのは初学者には容易ではないように思われる。

本稿は、そうした一見非常に複雑に見える質的研究の質の基準を大まかに整理して見通しをよくすることを目指とする。必ずしも、質的研究の質に関するすべての基準を網羅しているわけではない。質的研究の質を高めたり評価したりする際のヒントとして使ってもらってもいいし、また、議論を深めていくための一種の叩き台と考えてもらってもよいと思う。以下ではまず、量的な研究の伝統的な評価基準を参考にしながら、それを変奏したものとして質的研究の質を捉えていく。しかしながらそのようにして捉えられた基準は、研究途上で質を高めるために利用することはできても、完成された研究のレポートを第三者が読んでそれを評価する際に役立つとは限らない。そこで、よりすぐれた科学研究の知見とはどういうものであるのかを考えながら、レポートを評価する基準を別の観点から提示してみたい。

伝統的な研究評価の基準

質的研究の質の基準の話に入る前に、伝統的に使われてきた評価基準について少し述べておきたい。伝統的な基準を継承するにせよ批判するにせよ、それから全く切れたところで質的研究の評価が行われているわけではないからだ。一般的に言って、私たちがものごとに対して行う良し悪しの判断は、もっとも単純な形では「真・善・美」という基準をもっている。しかしすべてが研究の評価に使われるわけではなく、伝統的な自然科学の研究においては、このうち「真」がもっとも重要と考えられている。ここで真であることは、認識されたことと現実存在すること、あるいは認識という〈主観〉が現実という〈客観〉に一致することである。研究という文脈でより具体的に言えば、結論的な命題がデータという事実と一致し、データが「現実」に

* 本稿は、伊藤哲司・能智正博・田中共子編『動きながら識る、関わりながら考える ― 心理学における質的研究の実践』（2005、ナカニシヤ出版）の第11章「質的研究の質」（能智担当）に加筆修正を施したものである。

一致していればそれは「真」であり、その研究はとりあえず良いもの・すぐれたものと判断される。なお、“命題”とは、真偽が決定できるような文のことであり、典型的には「AはBである」とか「XがYを引き起こす」などといった形をとる。質的研究にどのくらいこの枠組があてはまるかは議論が分かれるところかもしれないが、とりあえずこの自然科学の枠組を頭に入れておいていただきたいと思います。

1) 基準の種類とその背景

自然科学を範とする伝統的な実験心理学では、基本的には同じ枠組が使用できる。行動主義の立場に立つ心理学者なら刺激と反応の間の結びつきについて、認知心理学者なら刺激と反応の間にある媒介的な認知過程について、何らかの仮説を立てるところから研究を始めるだろう。そうした仮説は一種の〈主観〉なのだが、それが現実一致しているかどうか、現象を(理想的には実験条件下で)観察することによって検討されることになる。質問紙で人の性格を知ろうとしたときも同様である。質問紙の結果として記述された性格(という主観)が実際の行動に比較されるなどの手続きを通じて〈客観〉との一致が推定されたときに、それは正しい結果とみなされるだろう。こうした考え方の前提には、行動法則や認知、性格など心理学の対象が客観的に存在するという考え方がある。これは、自然科学とも共通する実体論的な仮定—すなわち、事物が人間の認識以前に客観的に存在するという仮定—に対応している。

もっとも典型的な量的研究の過程を単純化して示したのが図1である。研究のターゲットとして仮定されているのが客観世界、すなわち客観的に存在する人間の心理や行動の全般である。そこに見られる属性を記述し、その間の関係を典型的には因果関係として捉えることが研究目標になる。その目標を実現するためには、ターゲット全体を代表するよいサンプルが選択されなければならない。理想的に言えば、少数のサンプルでそのターゲット全体(母集団)を代表させるには、そこから無作為にサンプルを抽出する必要がある。そして、ターゲットに関して作られた仮説的な命題—「AはBに影響している」、「XがYによって生じる」など—をもとに、そのサンプルから行動的なデータが収集され、最終的に、仮説とデータを比較して仮説が採択されたり棄却されたりする。この種の研究における質は、図1に出てくる諸要素、つまり、“客観的世界”、“サ

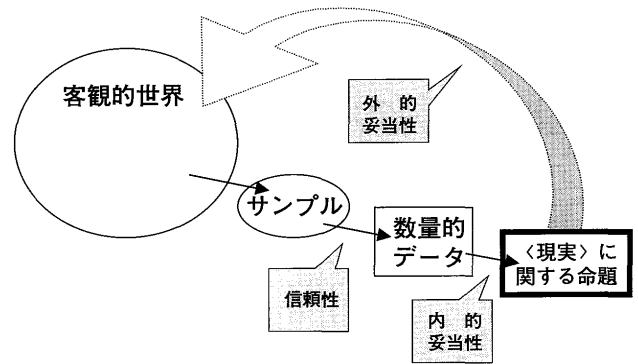


図1 量的研究の過程と関連する評価基準

ンプルとそこから得られたデータ”、“〈現実〉に関する命題”がうまく一致、ないし対応しているかどうかを問うものである。

その一致の確からしさの指標として伝統的な研究で工夫されてきたのが、信頼性reliabilityと妥当性validityというよく知られた概念である(平井、2003)。信頼性とはデータの一貫性のことであり、同じ条件で複数回測定したら同じデータが得られ、最終的な知見においても反復が見られた場合「信頼性が高い」と判断される。図1で言えば、主にサンプルからデータに向かう矢印と関係しており、ここで安定したデータが得られないならばデータ収集の方法などを考え直さなければならない。

妥当性は、ふつう内的妥当性と外的妥当性に分かれるが、内的妥当性は、図で言えば、主として“数量的データ”から“〈現実〉に関する命題”を取り出す過程と関係している。すなわちそれは、測定されたデータに基づいて主張される独立変数・従属変数間の関係がどの程度確からしいかを意味しており、注目している以外の変数の影響で従属変数の値が変動する場合には内的妥当性は低いことになる。最後に外的妥当性だが、これは研究結果としての命題が観察した状況以外にも適用できるかどうかの程度で、「一般化可能性」とも呼ばれる。ここで“〈現実〉に関する命題”が表現する〈現実〉とは特定のサンプルを越えた世界と対応しており、図1では“客観世界”と名づけられた部分である。この外的妥当性を確証するための前提として、上で述べたようなサンプリングの適切さ、つまり想定される“客観世界”をサンプルがどれだけうまく反映しているかが問題になる。

2) 質的研究への適用の困難さ

〈客観〉ないし〈現実〉との一致が重視される量的研究の評価基準について述べてきたが、質的研究を評価する場合これら基準をそのまま使うことには議論が多い。研究者によっては、これらを適用することは意味がないとすら考えている(例えば、Wolcott, 1990)。その理由は、質的研究の認識論的・存在論的な特徴にもよるし、方法論的・手続き的な特徴にも関係している。なお、「認識論」、「存在論」などについては、本稿では直接言及しないで考察を進めていくので深入りはしないが、興味のある方は、例えば、Lincoln & Guba (1985) や能智 (2005) などを参照していただきたい。

ともあれ、どういう点において適用がむずかしいのか、3つの基準のそれぞれについて考えておこう。信頼性について言えば、質的研究では研究対象者からいつでも同じような安定したデータが得られるかどうか、得られることに意味があるかどうか、といった問題がしばしば指摘される。質的研究のデータ収集においては自然場面における観察や面接が多用されるわけだが、そうした場面における言動は表面的にはかなり大きく変動するのがふつうである。例えば、自然場面での観察では、基本的に状況は1回限りのものであり、すべての環境刺激が全く同じでそれを受け止める対象者の意味づけも全く同じなどということはいえない。また、観察者が異なれば、同じフィールドでも行動の現れは異なることが考えられる。にもかかわらず全く同じ行動が一貫して見出されるとしたら、それはむしろ普通でない事態であって、その理由を問いかけていく必要が出てくるだろう。また、無理に行動的なデータのレベルにおいて反復や安定を求めようとする、データ収集の状況を過度に制限したり状況をコントロールしようとするなど、質的研究の利点を殺してしまうことにもなりかねない。

データから「〈現実〉に関する命題」を構成する過程、つまり内的妥当性に関わる過程においても、質的研究の場合には独特な手続きをとるために、その評価はそれほど単純ではない。質的研究では、データの性格上、データの表面的・形式的な属性—例えば、フィールドノーツの字の大きさや濃さといった属性—ではなくその向こう側にあると想定される「意味」が多かれ少なかれ問題とされる。その「意味」も、一般的で辞書的なものよりもむしろ文脈のなかで浮かび上がってくる独特の意味を読みとり、解釈していかなければならないことが多い(能智, 2004)。「対象に関する命題」の

構成は、そうした「読み」と切り離せないのだが、そうした場合、直接収集されるデータと最終的な結果の間には、量的研究以上に複雑なデータの「加工」や「操作」、すなわち解釈が必要とされるだろう。また、構成される命題はたいてい数式のような単純な形で表現されるものではなく、要因間の複雑な関係が言葉で記述される。したがって、量的研究の評価のように、データのこの部分を検討しておけば内的妥当性を確認できるというような、単純なやり方は通用しにくい。

さらに外的妥当性に関して言えば、質的研究では「客観的」に存在する母集団を代表する研究対象(サンプル)を選択しづらいことが多いという点に注意すべきだろう。ランダムサンプリングできるほど明確に母集団が定義できないこともあるし、また、定義されていてもそうした母集団に属する人を数多く研究協力者とすることがむずかしい場合もある。例をあげておこう。知的障害をもつ人の心理や行動を研究しようとするとき、「知的障害」を定義するという問題が当然起こってくる。まず、一般社会で行われている定義と知的障害をもつ人自身による定義は、必ずしも一致しないかもしれない。法律的な定義を便宜的に用いるにしても、プライバシーの問題もあってその定義にあてはまる人びと全体が特定できるわけではないし、ましてやそこからランダムに選んで研究に協力してもらうことはほとんど不可能である。サンプリングの手続きを根拠にして、外的妥当性を保証することは、質的研究では非常に困難と言わなければならない。

そもそも質的研究では、客観的な世界がまずあってそこから価値中立的にデータが取り出されるという仮定がしばしば疑問視される(例えば、Lincoln & Guba, 1985)。その場合、「知的障害」の例もそうだが、一般に客観的に存在すると考えられている対象も、実際には人々の間で行われている習慣や力関係に基づく定義によって構成されたものかもしれない。構成のしかたはその個人が属する文化や社会などによっても異なるし、個人の観点によっても違うことがある。また、そのように多様な文化や観点が交錯する世界を、研究者だけが客観的に観察できるわけでもない。研究者もまた、ある特定の文化のものの見方から外に出られはしないからである。外に出たと思っても、やはりその視点は別の文化の中にある。

現代では自然科学の分野ですら、データは常に研究者のもっている「理論」を通してはじめて見えてくるものとされることが多い(村上, 1979)が、人文・社

社会科学においてはなおさら研究者の観点が収集されるデータに影響するだろう (Willig, 2001 上淵他訳、2003)。だとしたら、厳密には図1で描いたような構造自体がなりたらず、この構造を前提とした質的判断基準も、根拠を失ってしまうことになるかもしれない。質的研究の質的判断基準を問い直すことが求められるゆえんである。

質的研究の質を高めるための視点

以上述べたように、信頼性・妥当性といった評価基準を、そのまま厳密に質的研究に適用しようとするのはいささか無理がありそうに見える。しかしながら、質的研究が量的研究と全く無縁のことをやっているかと言えそうとも言えない。いずれも、「研究」と名付けられて営為を広い意味での「科学」の名のもとで行っているからには、素朴に観察しているだけでは見えてこないことを、データをもとにしてより確かな言葉にしていこうとする点で、何らかの共通点をもっているはずである(能智、2005)。

そうした共通性を考慮し、図1を修正して質的研究にも適用できる形にしたのが図2である。質的研究の場合にも、特定の対象(「事例」)からデータ(「質的データ」)を収集することになるが、その対象は必ずしも母集団から抽出されたサンプルとは限らないし、上で述べたようにそもそも母集団なるものをはじめから想定することが適切でない場合すらある。ただ、そのデータから事例や事例に関わる何ごとかかについて、研究結果として言えること、すなわち“〈現実〉に関する命題”を作らねばならないという点は、質的研究でも同

じである。ここで〈現実〉とは、直接のデータを超えた何かではあるが、それはデータの背後にある特定事例の全体像や多くの事例の共通特性かもしれないし、事例を越えた抽象的な何かが想定されている場合もある。以下では、図2のような過程のなかで質的研究を評価するのに利用可能な視点と、その視点に即して質的研究の質を高める方法について整理していきたい。

1) 質の高いデータが収集・使用されているか

質の高い研究を行うための条件の1つとして、直接の研究対象から「質の高い」データを得てそれを分析に使用することが必要である。これは量的研究で「信頼性」と呼ばれている基準とゆるやかに対応している。もっとも、量的研究ではデータとそこから得られる推論の安定性・一貫性が重視されていたのに対し、質的研究では若干ニュアンスが異なる。質的研究における評価基準を早くから提唱していた例えば Lincoln & Guba (1985) は、「信頼性」という言葉をやめて「依拠可能性 (dependability)」という用語を提案した。この用語を利用させてもらおうとしたら、ここでのポイントは、収集されたデータがその後の分析、つまり、仮説生成やモデル構築などといった作業のために、dependable (頼りになる、依拠できる) かどうかということである。頼りになるデータというのは、研究者の想像やでっち上げではなく信用できるものであり、そこから新たな仮説やモデルを見出すことができるほどの豊かな内容をもっているということにほかならない。ここではとりあえず、この基準が満たされるための2つの条件について述べておこう。

フィールドとの関係 こうした意味におけるデータの依存可能性は、少なくとも部分的には、データソースであるフィールドや研究対象者と研究者との関係によって決まってくる。文化人類学のフィールドワークでは、対象となる地域に一定期間以上住んでその文化を直接体験しているかどうか、信頼できる情報提供者を得てその人と良好な関係を結びながらデータ収集ができていitかどうか、といった点がデータ収集の要件の1つである (Langness & Frank, 1988 米山・小林訳、1993; 佐藤郁哉、1992)。これは心理学の質的研究においてもある程度あてはまるであろう。質的研究では一般に「参加観察」と言って、対象者をある条件のもとで客観的に観察するのではなく、フィールド(現場)における関わりながらの観察が重視される。ゆえに、その関わりの質がデータの質に関係することは十分考え

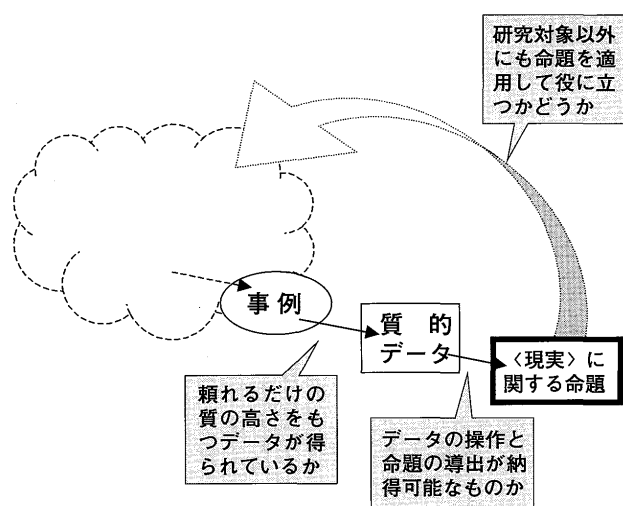


図2 質的研究の過程と関連する評価の視点

られる。また、研究対象者が社会的に劣位にあるとされる人々の場合には、研究者のアプローチの仕方によっては拒否的な態度が示されて、データ収集自体が困難になることすらある（能智、2001）。

したがって、質的研究の質を高めていくために前提となる作業の1つは、ある程度の時間をかけて研究フィールドを熟知し、研究対象者と良好な関係（ラポール）を構築しつつデータ収集をすることだと言える。もちろん、関わりが深くなりすぎるとかえって見えにくくなる事柄もあるかもしれない。しかし、少なくとも敵対的な関係が豊かなデータを生み出すことはまずあり得ないだろう。いずれにせよ、得られるデータの内容や質は、研究者と対象者との関係の関数なのだということは心に留めておくべきである。研究のレポートにおいては、そうした点がわかるように、研究者と研究フィールド、および研究者と対象者との関わりが記述されていなくてはならない。完成したレポートを評価する場合には、評価者はしばしば、そこで記述されている関わりの程度に注目することがある。

「厚い記述」 また、収集されたデータが、新たな仮説の生成やモデルの構築を支えるためにdependableであるかどうかは、それが「厚い記述（thick description）」になっているかどうかという点にも関連する。この言葉は、哲学者の G. Ryle の考え方を文化人類学者の C. Geertz が紹介して有名になったものであり、単に記述量の多さを意味しているわけではない（Geertz, 1973, 吉田他訳, 1987）。例えば、観察対象の A という少年は片方のまぶたがけいれんし、B という少年は別の少年に目配せをしたとする。この2つは、「片目をつぶった」という外見的行動のみの「薄い記述」をする限りでは、まったく区別できない。けれども、その行動がどういう状況で生じたか、少年がその後どのように行動したか等をともに記述していくことで、明らかに異なる意味をもつものとして区別される。後者のような意味的な詳細がわかるような記述を、Geertz は「厚い記述」と呼ぶのである。Denzin (1989) によれば、観察された行為を時間的な流れや空間的な脈絡とともに、その関連性も含めて記述する、といった心がけのなかで、記述データは「厚み」を増していくとされる。

心理学の質的研究においてデータを収集する際にも厚い記述は重要である。例えば、研究の様々な段階でそうした厚い記述のデータに戻ることによって、対象を常に多面的に見ていくことが可能になるだろう。フィールドワークにおけるデータ収集初期には、自分

の当初の関心だけに縛られることなく、対象の様々な側面を網羅的に記述することが多い（箕浦、1999）。インタビューを通じてのデータ収集においても、問いに対する被面接者の答えを要約的に書くのではなく、逐語的なトランスクリプトを作ることが基本であり、加えて、被面接者が示すノンバーバルな行動（服装、仕草等）や面接者の働きかけ方も含めて記述する。さらに、単に記述の量ではなく記述の厚さを生み出すには、記述されている事象を他のさまざまな事象と関係づけながらデータを蓄積していくことが求められる。

データのレベルにおける記述の厚さは、研究のトレーニングにおいて習得すべきものだが、研究の途上においても絶えずチェックしていかなければならない。自分で自分のフィールドノートやトランスクリプトを読み返してみることも必要だろうし、また、指導してくれる先生や同専攻の友人・先輩にデータを読んでもらったりすることによってもある程度チェックすることができるだろう。一方、研究結果のレポートの場合には、そこにデータ全体が含まれることはないにしても、読み手は、直接引用されるデータやそのレポートそれ自体の記述から、記述の厚さを推測することができる。そしてこれもまた、レポートを評価する際の1つの視点になるだろう。

2) データの処理と命題の導出は適切か

どんなに対象の素朴な記述を重視した質的研究でも、集められたデータそのものを結果として投げ出すだけでは研究とは言えない。データの選択や縮約や配列が加えられ、抽象化や推論がなされて“〈現実〉に関する命題”が構成され、そこではじめて研究としての1つの条件が満たされることになるだろう。その際の手続きに問題がないかどうか、結果の記述がデータと矛盾していないかどうか等といった点も、質的研究の質を考えていく上では重要である。データとそこからの推論の関連について Lincoln & Guba (1985) は、数量的研究で用いられてきた「内的妥当性」の概念を使わず、質的研究のために「信用性（credibility）」という用語を提案している。日本語で「信用性」と言ってしまうとピンとこないかもしれないが、要するにこれは、レポートに記述されているカテゴリーやカテゴリー間の関係がその背景にあるデータを確かに（credibleに）反映しているかどうかという点にかかわっている。

データをもとにして何らかの命題を導く際に見られる誤りは、質的研究であろうと量的研究であろうと、典

型的には2つのタイプがある(佐藤達哉, 2002)。1つは統計的検定で「第一の過誤」と呼ばれているもので、実際には関係や差がないのにあるとみなす、いわゆる“慌て者の誤り”である。すなわち、データのなかに明確な証拠が見られないのに思いこみでそこにXというカテゴリーがあると結論づけたり、あるいは、偽相関にすぎないのにXとYの間に因果関係を見てしまった場合などがそうである。もう1つは「第二の過誤」と呼ばれており、これは逆に、関係や差が存在するのに存在しないとみなしてしまう、“ぼんやり者の誤り”である。レポートに引用されたデータを読んでいて、「こういう概念を使えばデータをうまく説明できるのではないか」とか、「こういう関係がデータには現れていそうだ」と感じるとしたら、その研究結果には「第二の過誤」が含まれている可能性がある。

これらの過誤を回避するためにいくつかの方法が提案されており、そうした方法の適用が見られない研究に対しては、「信用性」が疑わしいという評価が下されるかもしれない。研究者は、研究の過程においていずれかの方法で自ら導き出した命題を「検証」し、さらにその検証の方法をレポートに明記しておく必要があるだろう。以下にその方法のいくつかを紹介しておく。

仮説を別のデータで試す 質的分析の過程で概念間・要因間の関係が仮説的に見出された場合、その関係をデータが確かにその関係を支持しているかどうか、データに戻って検討しておかなければならない。これは、量的研究とはまた異なるかたちの「検証」にあたる。そのための手続きの1つが“分析的帰納法”であり、次のような手続きで進められる(Robinson, 1951)。まず説明すべき現象を暫定的に定義し、データに即してその現象について「概念間・要因間の関係」という形で説明仮説を作る。次にその仮説が別のデータと比較され、データに適しない場合には仮説を修正するか、あるいは説明すべき現象を再定義してそのデータの事例を現象から除外する。その後さらに、次のデータが仮説に照らし合わされ、同様な手順で仮説が検討される。この手続きが反復されることで、仮説がより確かなものになったり、あるいは精緻なものになったりする。

例えば、データのある部分を根拠に、事象Aが事象Bの生起条件であるという仮説が浮かび上がってきたとしよう。まず行うべき作業は、他のデータにおいて事象Bが認められるときAが先行して生起しているかを確かめることである。あるいは、Aが生起していな

いときにBが起こったりしていないかどうかを確かめてもよい。矛盾するデータがあれば仮説を変更したり、介在する別の要因を探索したりすることになる。それまでに見出されたと考えられる概念間の関係とは異なるデータ(ネガティブ・データ)を積極的に探索してみることも、結果として現れてきた命題の検証のために行っておく価値がある。ただし、ネガティブ・データが見出された場合、それまでの仮説が全く無価値であることを意味するわけではない。それはむしろ、仮説の適用範囲を限定したり、そのデータをも説明できるようにより包括的な仮説を構築したりするよい機会である。

別の視点から見直す データから導き出された関係に関する命題を検証する方法としては、その他に、トライアングレーション(triangulation)をあげることができる(Denzin, 1978)。これはもともと「三角測量法」という意味で、ある地点を2つの別の地点から見て三角形を作るという手続きを用い、数多くの三角形の組み合わせによって周辺の位置関係や形状等を正確に測定する方法である。質的研究では、ある視点において得られたデータやそこから見出されたことに対して、別の視点から見ても同じようなものが得られるかどうかチェックすることを指している。「別の視点」にはいくつかのタイプがある。例えば情報源に関して、面接データから導き出された命題が観察データにも認められるか確かめていく場合がある。また、まれには質的研究からの知見を量的研究の方法で確認していくこともトライアングレーションと呼ばれる。複数の視点から得られた複数の命題に矛盾がなければ、その命題はより確からしいものとみなされるだろう。

トライアングレーションの一部と見なすこともできるが、他者の視点を導入して仮説の確かめを行うことがある。Lincoln & Guba (1985)は、なかでもメンバーチェックと呼ばれる方法を重視している。これは、研究対象者やその関係者などに、データやその解釈、および最終的な研究結果を示して、それが納得のいくものであるかどうかを検討してもらうという方法である。また、データ収集と分析の過程をなるべく詳細に記録しておき、同僚や友人、あるいは他の専門家に示して、彼らの目でデータ処理や推論の根拠などについてチェックを行うというやり方もとられる。いずれの方法をとるにしても、他者の目から見て、研究者によるデータの分析過程とその結果に大きな問題点が見いだされなかった場合には、その研究の「信用性」がより

高いものとなる。

もっとも、ここで注意しなければならないのは、別の観点のもとで導き出された結果が研究者の導き出した結果と異なってしまったとしても、研究者の側の誤りとばかりは言えないという点である。どんな視点から見ても同じ結果になるのが理想と考えるのは、むしろ、人間の認識とは独立に唯一絶対の現実があるとする実体論的前提への逆戻りであるとも言える (Seale, 1999)。視点や関心が異なれば〈現実〉の見えが違ってくるという立場に立つのであれば、トライアンギュレーションやメンバーチェックで違った解釈が現れてくるのも決して目くじらを立てるべきことではない。そうした場合、研究者が自らの分析過程を反省することはもちろん大切ではあるが、それと同時に、どういう理由でそのような違いが生じたのかを説明する包括的なモデルなり仮説なりを構築することを試みてもよいだろう。もしそうした説明ができない場合には、やはり認識の歪みや推論の誤りなども含めて、研究者の分析上の問題点を想定していかなければならない。

3) 結果を利用することができるか

伝統的な量的研究において、信頼性や内的妥当性とならんで重要視されている研究の質が「外的妥当性」である。それは上で述べたように一般化可能性のことであり、これはデータから母集団全体にあてはまるような平均像を描き出すことで達成されるのだが、質的研究では必ずしも平均像を描くところに目標が置かれるわけではない (Schofield, 2002)。例えば Lincoln & Guba (1985) は、「外的妥当性」の代わりに「転用可能性 (transferability)」という言葉を用いている。この transfer という言葉にしても、サンプルから母集団へといういっそう大きな集団への、つまり「上」への転用だけを意味するものではなく、比喩的に言うと、「横」への、ときによっては「下」への transfer ができるかどうかということも、研究の質の高さと関係している。以下では、特定のデータから得られた命題をそのデータ以外の何事かの理解や洞察などに利用できるかどうかという意味で、「外的妥当性」におおまかに対応するような質的研究の質をとらえてみたい。

自分と自分の周囲への「転用」 レポートに記述されている対象がリアリティをもった存在に感じられ、その対象を自然と自分や自分の周りのことに無理なく重ねて理解できるという形の転用が、「自然な一般化 (naturalistic generalization)」という用語とともに、質

的研究の質として言及されることがある (Denzin, 1989)。読み手は、記述されている対象が自分のよく知っている人と類似していると思うかもしれない。また、客観的には自分と全く境遇が違うにもかかわらず、ある部分で自分と共通の体験をしていると感じるかもしれない。そうした場合、研究対象が読み手自らの体験とつながるものとして捉え直されているということになり、横方向への転用が起こっていると言えるだろう。ここで得られる個人的な了解感、人が対象を共感的に理解したり、あるいは、記述されている問題を自分に関係のある世界のこととして一即ち、他人ごとではないこととして一引き受けたりすることの基盤になりうる。

自然な一般化が生じるようなかたちで対象記述を行うためには、レポート記述において適宜ローデータを提示しながら「厚い記述」を行っていく必要がある。Wolcott (1990) は、読み手があたかも自分の目で研究対象を見ているかのように感じられるようなレポートを書くことの重要性を指摘し、そのためには観察データや対象者の発話データをなるべく多く引用するようにと述べている。こうした工夫は、自然な一般化のための十分条件とは言えないにしても必要条件の1つであることは確かだろう。実際、吉村 (1989) は、臨床実践のためには普遍的な行動法則の提示よりも事例の詳細な記述の方がずっと役に立つという臨床家の実感を紹介しつつ、その理由として、行動のある側面だけに焦点をあてるのではなく、状況とこみにした個体の全体性を読み手に伝えられるという点を指摘している。その全体性の感覚を読者に伝えるためにも、厚い記述は重要な戦略の1つなのである。無論、引用が多ければいいというものではなく、それが筆者の解釈と整合して提示されないと意味がないわけだが、それは以下で述べる基準と関係してくる。

他の事例への「転用」 自然な一般化に加えて、質的研究のレポートにおける対象の記述が他の事例を理解する際の手がかりになると感じられるほどに、その研究は質が高いとされる。その場合の転用を可能にするのはアナロジーと呼ばれる認知機能である。Holyoak & Thagard (1995, 鈴木他訳, 1998) の示すように、人は知覚される類似性を根拠にして様々な推論や問題解決を行っており、新たなアイデアを練り上げたりする際にもアナロジーが使われる。質的研究においても、直接の研究対象についてデータから導き出された命題が読み手の経験する事象と似たところをもつならば、そ

の命題を利用して身の回りの事象に関する新たな気づきを得たり、理解を深めたりする可能性が生じる。その事象の予測やコントロールなど、現実への働きかけの手がかりになる場合もあるかもしれない。そうした転用が可能になるためには、ここでもやはり、研究レポートにおいて対象に関する厚い記述がなされているかどうかという点が重要である。厚い記述は読み手の経験との類比を容易にする手がかりを、より多く含んでいるからである。

その研究が他の事例を位置づけるものさしを提供する場合にも、研究結果の転用可能性は高い。見田(1978)によれば、質的研究では関心の対象の平均的な像を描き出すよりも、〈典型〉の特徴を明らかにしていくことがしばしば重要であるという。ここで言われている〈典型〉は、平均とは異なり、他とは区別されるその対象の独特な属性を明瞭に示す事例のことである。例えば、「フリースクール」の〈典型〉とは「一般の学校」との対照が際だっているような事例を言う。〈典型〉が詳細に記述されていれば、別の事例を見るための目安としてそれを役立てることが可能である。しかし、選択された対象がそのカテゴリーの〈典型〉であることを示すためには、「一般の学校」の属性がある程度わかっていることが必要であり、最終的なレポートにはその一般的な属性との差異が詳細に記載されていなければならないだろう。

転用される「他の事例」の範囲は、研究結果としての命題をどのような抽象レベルで表現するかによって違ってくる。すなわち、研究対象の特徴をより抽象度の高い視点から概念化することで、直接の研究対象に関する問題をいっそう多くの事象に関わる問題として提示することが可能である。例えば Angrosino (1994) は、知的障害をもつ1青年の語りと行動を検討し、「バス」が、街を自由に行き来できる自分という彼の理想の象徴となっていることを示したが、もちろん「知的障害者」と呼ばれる人がすべて「バス」をそのように見ているわけではない。Angrosinoはこの結果を、「自己像の重要な側面はその人の言動のなかである象徴を通じて表現される」というより抽象的な仮説にまとめている。この仮説は、知的障害者のみならず人びと一般に適用できる抽象水準をもち、その抽象水準で記述することによって、結果を他の事例に適用する際の範囲が大幅に広がっていると言える。

質的研究のレポートを評価する

以上、伝統的な評価基準を参考にしつつ質的研究の質を判断する視点と質を高める方法について述べてきたが、そのうちいくつかは研究のプロセスで行われるべき工夫であり、質的研究の結果を記したレポートには直接は表れないことが多い。とりわけ字数が制限されている雑誌論文の場合には、分析手続きの詳しい記述や長い質的データの提示には限界があり、例えば、データ収集時の記述の厚さを細かくチェックしたり、データからの推論過程を辿ったりすることはむずかしい。外部の評価者は、質的研究の質をレポートに盛り込まれた情報だけから判断しなければならないのが通例である。この節では、そうした状況で使われる評価基準を、上とは少し別の観点から簡単に解説しておきたい。Lofland & Lofland (1995)などを参考にしてまとめると、その基準は便宜的に、「現実性」、「新奇性」、「関与性」の3つに分けることができる。

1) 「現実性」という基準

「現実性 (reality)」というのは、記述されている結果が「現実」と呼ばれるものにつきあたっていると感じられる程度のことである。それが、研究者だけの単なる思いつきと思われるときや、イデオロギーとか理念に導かれているように見えるときには、この「現実性」は低いということになる。では、どのような基準でその「現実性」が判断されているのだろうか。西(2001)によれば、人は日々膨大な情報を受け取りながら、それがほんとうらしいかどうかといった現実性の判断をしているが、その際に、受け取る情報を「現実そのもの」と逐一对応づけて判断するわけではない。例えば、目の前にある石ころが夢ではなく現実にそこに存在していると判断する根拠は、その存在がありありとした実感を伴って感覚器官に受け止められること、何度見てもそこにあると感じられること、視覚だけではなく触覚を通してでも感覚が得られること、さらに、他人もまたその石の存在を疑わないことなどである。言い換えれば、それは与えられる情報の具体性や直接性、異なる時点で与えられる情報の連続性、異なるチャンネルからの情報の一貫性、他者からの承認などが、私たちの日常的な現実性の根拠となっているのである。

同じことは、質的研究の「現実性」を判断する際の基準についても言えるだろう。すなわち、重要なのはまず、研究対象について導き出された命題が具体的な

データの引用によって例証されており、その内容が生き生きと実感されることである。それによって、先に述べた「自然な一般化」が促進されるとも考えられる。次に、レポートの内部において矛盾がなく、そこで与えられている情報に論理的な一貫性があることも必要である。それは、データとそこから導かれる解釈や結論との間に見られるそれなど、比較的狭い範囲の一貫性かもしれない。あるいは、論文の問題提起と結論など、論文全体にわたる一貫性であるかもしれない。また、異なるチャンネルや他者からの情報の一貫性としては、先ほど「信用性」の基準において述べた、トライアングレーションやメンバーチェックによる見解の一致の確認を挙げることができる。これらの手がかりを適宜使いながら、そのレポートの「現実性」が評価されるわけである。

2) 「新奇性」という基準

「新奇性 (novelty)」とは、そのレポートのなかで語られていることがこれまでの常識や先行研究などになかった知見を含んでいると感じられる程度を示している。単純なところでは、そのテーマや対象がこれまでの時代には見られなかった新しい現象を扱っている場合には、新奇性は高くなるだろう。例えば、“フリーター”と呼ばれる「生き方」は比較的最近現れたものだが、そうした生き方を選ぶ人々の行為や体験の特徴をいち早く記述し、その意味を明らかにしたとしたり、— やや社会学的ではあるかもしれないが — 新奇性のある質的研究ということになる。また、これまであまり注目されてこなかった対象を選択した場合にも、新奇性は高まるだろう。例えば、障害をもつ人々や特殊な状況にある人などの声をわかりやすく整理して呈示していくことは、質的研究でしばしば目標とされるところである (Ferguson, Ferguson, & Taylor, 1992)。

しかし、新奇性はテーマの目新しさだけに関わるものではなく、それまでの理論や常識的なものの見方がその研究によって変更されるときによりいっそう高まることになるだろう。そうした変更は、大ざっぱに言って次のような4つのパターンに分けて考えることができるという (Lofland & Lofland, 1995 進藤・宝月訳, 1997)。

- ① 構造化されていないと見えるなかに構造がある、あるいはその逆。
- ② 異質と見える現象が実際には1つの要素からなる、あるいはその逆。

③ 安定と見えるなかに変動がある、あるいはその逆。

④ 類似と見えるなかに差異や反対の属性がある、あるいはその逆。

確かに、科学史上の著名な発見や理論的進歩の多くは、この4つのいずれかと関わっていると思われる (小山, 2003)。例えば、メンデレーエフによる元素の周期律表の発見は①に関係している。それまで元素はばらばらに発見されていただけで、その間に関係があるとは考えられていなかったからである。②に対応する発見としては、フランクリンが嵐をあげて雷が摩擦によって起こる電気と同じものであると示した研究をあげることができる。それまでは、雷と電気は異質のものともみなされていた。③については、例えば、安定していると考えられていた大陸が長い時間枠で見れば動いていることを主張した、ヴェーゲナーの大陸移動説と対応させられるだろう。④は、現在のところまだ研究中ではあるが、老年期の物忘れなどとは違う病理症状としての痴呆症状の定義などに関係している。「正常な」老化と痴呆 (認知症) の違いがはっきりと区別できる指標が見つければ、それは早期の治療に道を拓くことになるかもしれない。

質的研究では新たな仮説の生成が重視されるが、「新しい」という評価がなされるためには、こうした考え方の転換がどこかに含まれていることが必要であろう。その転換が、レポートのなかで明確な形で記述されている場合に、受け取り手はその研究の価値を強く感じるようになる。

3) 「関与性」という基準

最後に「関与性 (relevance)」だが、これは受け取り手がその研究を自分や自分の生きている世界に関係し、重要だとみなす程度を示している。人は様々なレベルで周囲に関心や欲求をさし向けながら生活しており、その関心や欲求との相関の高さに応じて、その対象に意味があるとか意味がないとかいった判断を下す (西, 2001)。例えば、マッチの箱のデザインに関心がある人にとって、珍しいマッチの空き箱は非常に重要な意味をもつが、関心のない人にとってはゴミにすぎない。研究にしても、評価者の関心に響きあう度合いに応じてより意義のあるものと判断されることになる。ただし、研究を評価する場合には日常的な評価とは異なり、「私個人」にだけ関与していればそれで済むわけにはいかない。評価者は、意識的にせよ無意識的にせよ、「われ

われ」にとっての関与性を判断しながら、その研究を評価することになる。

人が「われわれ」と感じる共同性は1つではなく、したがって、関与性のレベルも複数存在する。例えば、最近マスコミで問題として取り上げられているような社会問題に研究が直接関係していたり、あるいは研究レポートの筆者によってそれと関連づけられていたりする場合には、現代社会の成員としての「われわれ」にその研究が関与していると評価されるかもしれない。また、心理学の学的伝統の中で問題となっていることとの結びつきがレポートのなかで示されていれば、心理学の世界に生きる「われわれ」にとっての意義が直感されることになる。むろん心理学のなかでも様々な学派があり研究史が存在しており、より小さなそうした「われわれ」において重要性が認識されることもある。例えば、臨床心理学の分野で重要とされているテーマと認知心理学の分野でのホットなトピックは必ずしも結びつかないであろう。いずれにしても、そうした「われわれ」との関与を読み手に明確に伝えているかということもまた、研究の評価の一側面である。

終わりに

以上、近年発表されることが多くなった質的研究の実践とその結果の質をどのように評価したらいいのかという点を、筆者なりの観点から簡単にまとめた。

まとめながらときどき思い出していたのは、次の小林秀雄の言葉である。

美しい花がある、「花」の美しさというようなものはない。(「当麻」)

有名な言葉なので、おそらくご存じの方も多いただろう。この言葉をもじるとしたら、“よい研究がある、研究のよさというものはない”ということになるかもしれない。私も、半分くらいはその通りではないかと思う。ただ、美やよさを味わうだけならそれで済ませてもかまわないのだが、美やよさを創造しようとするときには、何か手がかりがほしいと思うのが人情というものである。その手がかりは、伝統芸能などの分野では、古来「型」という形で伝えられてきた。それは絶対的な着地点ではなく、とりあえずの中間点としてめざされるものであり、さらにそこから出発して型を発展させることもできる。本稿で提示したのは、そういう型のよいものと考えていただければちょうどよいだろう。

それなりに整理して提示したつもりではあるが、そ

れでもまだややこしいとか曖昧であるとかいった不満が出てくるかもしれない。それは半ば筆者の力不足によるところが大きいだが、同時にその曖昧さは、言葉を主に用いて研究を進めていく質的研究の特徴に見合ったものとも考えられる。やはり言葉(概念)によって現実の一面を捉えようとする心理検査の質の判断基準についても、似たようなつかみどころのなさがある。心理検査は「知能」、「性格」など目には見えない構成概念を測定するものだが、その心理検査の質を見るためには、最終的に「構成概念妥当性」が使われる。その妥当性の基準は複数あって、どれか1つだけで決められるものではないし、また、そのどれもが一定の意義をもっている(Walsh & Betz, 1985)。質的研究の質の評価においても、同様の基準の多様性が認められ、そうした多様な基準を組み合わせて使いながら、個々の研究が評価されていくことになるだろう。

引用文献

- Angrosino, M. V. (1994) On the bus with Vonnie Lee: Explorations in life history and metaphor. *Journal of Contemporary Ethnography*, 23, 14 - 28.
- Creswell, J. G. (1998) *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks: Sage.
- Denzin, N. K. (1978) *Research act* (2nd ed.) New York: McGraw-Hill.
- Denzin, N. K. (1989) *Interpretive interactionism*. Thousand Oaks: Sage.
- Ferguson, P. M., Ferguson, D. L. & Taylor, S. J. (eds.) (1992) *Interpreting disability: A qualitative reader*. New York: Teachers College Press.
- ギアーツ, C. 吉田禎吾・柳川啓一・中牧弘允・板橋作美 (訳) (1987) *文化の解釈学* 岩波書店 (Geertz, C. (1973) *The interpretation of cultures*. New York: Basic Books.)
- 南風原朝和・市川伸一・下山晴彦 (編) (2001) *心理学研究法入門 — 調査・実験から実践まで*. 東京大学出版会
- 平井洋子 (2003) 信頼性と妥当性 南風原朝和・市川伸一・下山晴彦 (編) *心理学研究法* 放送大学教育振興会 pp. 93 - 105.
- ホリオーク, K. J. & サガード, P. 鈴木宏昭・河原哲雄 (監訳) (1998) *アナロジーの力ー 認知科学の新しい探求* 新曜社 (Holyoak, K. J. & Thagard, P. (1995) *Mental leaps: Analogy in creative thought*. London: MIT Press.)
- 伊藤哲司・能智正博・田中共子 (編) (2005) *動きながら識る、関わりながら考える — 心理学における質的研究の実践* ナカニシヤ出版
- 小山慶太 (2003) *科学史年表* 中公新書
- ラングネス, L. L. & フランク, G. 米山俊直・小林多寿子 (訳) (1993) *ライフヒストリー研究入門 — 伝記への人類学的アプローチ* ミネルヴァ書房 (Langness, L. L., & Frank,

- G. (1988) *Lives: An anthropological approach to biography* (4th ed.) . Novata, CA: Chandler & Sharp.)
- Lincoln, Y. S., & Guba, E. G. (1985) *Naturalistic inquiry*. Beverly Hills: Sage.
- ロフランド, J. & ロフランド, L. 進藤雄三・宝月誠(訳) (1997) 社会状況の分析 — 質的観察と分析の方法 恒星社厚生閣 (Lofland, J., & Lofland, L. (1995) *Analyzing social setting* (3rd ed.) Belmont: Wadsworth.)
- 箕浦康子 (編著) (1999) フィールドワークの技法と実際 — マイクロエスノグラフィ入門. ミネルヴァ書房.
- 見田宗介 (1979) 現代社会の社会意識 弘文堂
- 村上陽一郎 (1979) 新しい科学論 — 「事実」は理論をたおせるか 講談社ブルーバックス
- 無藤 隆・南 博文・サトウタツヤ・やまだようこ・麻生 武 (編) (2004) ワードマップ 質的心理学 新曜社
- 西 研 (2001) 哲学的思考 — フッサール現象学の核心 筑摩書房
- 能智正博 (2005) 質的研究のめざすもの 伊藤哲司・能智正博・田中共子 (編) 動きながら識る、関わりながら考える — 心理学における質的研究の実践 ナカニシヤ出版 pp. 21 - 36.
- 能智正博 (2004) 質的データの分析 — データの読みという視点から 児童心理学の進歩 43, 272-293.
- 能智正博 (2001) 「障害」をもつ人々たちへのアプローチ 伊藤哲司・尾見康博 (編) 心理学におけるフィールド研究の現場 北大路書房 pp.110 - 119.
- Robinson (1951) Logical structure of analytic induction. *American Sociological Review*, 16, 812-818.
- 佐藤郁哉 (1992) フィールドワーク — 書をもって街に出よう 新曜社
- 佐藤達哉 (2002) モードⅡ・現場心理学・質的研究：心理学にとっての起爆力 下山晴彦・子安増生(編) 心理学の新しいかたち — 方法への意識 誠信書房 pp.173-212.
- 下山晴彦・子安増生 (編) (2002) 心理学の新しいかたち — 方法への意識 誠信書房
- Schofield, J. W. (2002) Increasing the generalizability of qualitative research. In A. M. Huberman, & M. B. Miles (Eds.), *The qualitative researchers' companion*. Thousand Oaks : Sage. pp.171-203.
- Seale, C. 1999 *The quality of qualitative research*. London: Sage.
- Walsh, W. B. & Betz, N. E. (1985) *Tests and assessment*. Englewood Cliffs: Prentice Hall.
- ウィリッグ C. 上淵寿・大家まゆみ・小松孝至 (訳) (2003) 心理学のための質的研究法入門 — 創造的な探求に向けて培風館 (Willig, C. (2001) *Introducing qualitative research in psychology*. Buckingham: Open University Press.)
- Wolcott, H. F. (1990) On seeking – and rejecting – validity in qualitative research. In E. W. Eisner & A. Peshkin (Eds.), *Qualitative inquiry in education: The continuing debate*. New York: Teachers College Press. pp. 121-152.
- 吉村浩一 (1989) 心理学における事例研究法の役割 心理学評論, 32, 177 - 196.

(2005. 3. 受稿)