

より望ましい質的研究のための一提案 —交絡 (Confounding) の側面から厚い記述を考える—

三上 仁志

Center for Applied Linguistics, University of Warwick

概要

本稿は、質的研究におけるデータの取集と報告に関するものである。質的研究において (1) 研究内で報告されたデータが他の研究のそれと比較できること、つまり比較可能性 (Comparability) が高いこと、そして (2) データの分析結果や結論の適用可能範囲が検証できること、つまりトランスファラビリティ (Transferability) についての評価が可能であることは、その科学性や妥当性を保証する上で重要であると考えられている。これらの条件を満たすために厚い記述 (Thick Description) をおこなうことは、広く推奨されている。本稿では、研究の比較可能性を高め、トランスファラビリティの評価を可能とするための厚い記述について、交絡 (Confounding) との関係からこれを議論し、研究枠組みの提案をおこなう。

Keywords: 質的研究, 交絡, 厚い記述, 比較可能性, トランスファラビリティ

1. はじめに

質的な研究をおこなう上で望ましいとされるデータ報告の特徴に、厚い記述 (Thick Description) というものがある (Charmaz, 2003; Richards, 2005)。厚い記述とは、文字通り研究データの報告を手厚くおこなうことである。だが、「データの報告が手厚い」とは、具体的にどのような状態のことを指すのだろうか？ 調査対象者 (達) にまつわる全ての関連情報を網羅して、それを論文中に書き記すことなのだろうか？ 論文に文字数や枚数の制限が無いとすればそれも可能かもしれないが、それが出来るのは学位論文やごく一部の投稿論文でだけだろう (大概のジャーナルには、文字数や枚数の制限が存在する)。また、例えば調査参加者の人生の包括的な記録を論文に書かれても、(その人生がつまらないからではなく、話が長すぎて) とても読む気がしないだろう。

論文の (査) 読者が必要とするのは、(1) 可能な限り簡潔で (2) 読みやすくとめてあり、かつ (3) その研究テーマに関して私たちの知識を深めてくれる情報だろう。この内、情報が簡素で読みやすいことの重要性は、自分が読者になった時のことを想像すれば容易に察しがつく。仮に偉大な研究内容だとしても、文章のまとまりが悪く、読むのに苦労して、しかも 100 ページもある論文を読みたい人は多くないだろう。文章の簡素さ・読みやすさは、作文技術の問題でもあり、その技術を

高めるための方法論や枠組みは、数多く存在する。例えば、常に読者の視点を意識して推敲をおこなうことは、より読者に伝わりやすい文章を書くことに繋がるだろう(Hyland, 2009)。このような道標が(3)知識を深めてくれる情報の選定方法についても存在すれば、厚い記述を達成することも多少は容易になると思われる。本稿の目的は、この点についてConfounding(交絡)という概念を軸とした枠組みを紹介することである。本稿で紹介する枠組みは、(a)調査をデザインする段階での調査対象項目の選定、および(b)論文執筆時の報告情報の優先順位付けに用いることが出来る。

なお、厚い記述は、(a)と(b)の中間にあたる「データの収集時と分析時」の段階でも積極的に起こることが推奨されている(Charmaz, 2003; Richards, 2005, pp.56-68)。これらの文献では、データの記録者(多くの場合、分析者自身)は¹、データを収集している最中に、その場で起きている現象や、その場で覚えた感覚、それらについての解釈などを併せて記録しておくべきであるとされている。具体的には、例えばボイスレコーダーを使ってインタビューを録音している最中に、メモ用紙などに調査協力者の身振り手振りや表情、それらに対する記録者の解釈、アドリブで追加された質問項目(と質問が追加された理由)などを併せて記録しておくことが、ここでの厚い記述となる。これらの記述があることで、研究者は、(1)副次的な記録(この場合は、ボイスレコーダーで録音できない調査対象者の身振りや手振り、表情についての情報)や、(2)データ収集時の記録者の感覚および解釈、(3)研究者がインタビューを聞いている最中に下した解釈と(4)インタビュー内容を全て把握した後で下した解釈、の全てをデータ分析に用いることが可能となる。また、データ分析の過程(つまり、最終的な解釈に行き着くまでの分析の流れ)についての詳細な記録を残しておくことで、研究の妥当性を検証する際により体系的(Systematic)な分析記録(log)を参照することが可能となる。本稿は、データの収集時と分析時における厚い記述については扱わないが、この点に興味がある方は、前述の2つの文献(Charmaz, 2003; Richards, 2005)や、その他にも戈木クレイグヒル(2013)などを読まれると良いだろう。

2. 厚い記述, 比較可能性, トランスファラビリティ

前章で述べた様に、「厚い記述がされている状態」とは、ある研究テーマについて我々の知識を深めてくれるだろう情報が、論文の中で十分に(そして望むべくは簡素かつ分かり易く)報告されていることである。これを達成するための第一歩として、まずは調査対象や問題の所在を限定して、そして設定された問題の解明がどのような(一般的もしくは学術的な)意味を持つのかを読者に伝えることが必要である(これがないと研究の方向性が分からないので、提供される情報をどの様に読み取るべきか、読者が判断できない)。この前提を満たした上で、ある研究で報告されているデータが別の研究のそれと比較可能であることは、各研究の研究手法や議論、結論の適切さを評価する上で重要である。この条件を、比較可能性(Comparability)という(SLAA研究会, 2013; 三上, 2014)。比較可能性の重要性は、それが満たされていない場合を想像すると分かり易い。データの報告が不十分であれば、まずは追行調査をおこなうこと、データの収集・分析方法や分析結果の解釈の妥当性を評価することが(部分的もしくは全体的に)出来なくなる。また、類似したテーマを

扱った研究との比較から各研究のデータや結論について議論し、新たな知識体系を構築してゆくことも困難となる(Richards, 2005)。この意味で、比較可能性が高いことは、研究の透明性が高いことと同義であり、根拠に基づいた議論を展開するための十分条件となる。

比較可能性が高い水準で満たされている場合、研究のトランスファラビリティ(Transferability)を評価することが可能となる。トランスファラビリティとは、特定の状況や調査協力者を対象として得られたデータや研究結果が、異なる状況や調査協力者に対しても適用できる程度/度合のことである(Brown, 2001, 2006)。英語の教室学習にポートフォリオを導入する例を使って、この概念を解説してみよう。ポートフォリオとは、学生に自らの言語的な成長や学習の進捗状況などを定期的に内省してもらうことで作成された、通時的な学習記録のことである(Council of Europe, 2011)。学生が継続的に自らの学習活動を内省することは、自分が鍛えるべき能力を具体的に意識させ、また(学習目標が可視化されるので)学習意欲(Motivation)の向上にも有効だとされる(British Council, 2007)。例えば、ある教育機関において最新のポートフォリオが開発されたとしよう。この機関では、この枠組みを用いた教育プログラムの施行によって生徒の学習意欲が大幅に向上したとする。データは、ある教師が担当した10人の生徒から、インタビューを用いて収集されたものであるとしよう。調査内容は論文の形で発表され、そこで提供されたデータは、比較可能性を高い水準で満たしていたとする。この様なデータがあれば、別の教育機関でも、この新規的なポートフォリオの有効性を検証したり、教育現場の状況に合わせて改善すべきポイントを洗い出したりすることが可能となる。仮に枠組みそのものは素晴らしいポートフォリオでも、使用する状況が異なれば、達成目標となる言語能力、生徒の年齢、生徒の性質や学習スタイル、教師と生徒の関わり方など多くの要素が変化するため、枠組みに何らかの変更が必要されることは多いだろう。トランスファラビリティについて評価し、議論することは、この様な違いを生み出す要因について考え、より効果的かつ生態的妥当性(Echological Validity)の高い教育法を考える上で重要となる。なお生態的妥当性とは、「研究の内容や方法が、学習者を取り巻く現実世界の実情を反映している程度(三上, 2014, p. 6)」のことである。

質的な研究は、研究知見の一般化を(必ずしも)目的としない。先ほどのポートフォリオの例でも触れた通り、調査環境が変われば、おのずと調査の結果も変わるという前提が質的研究の出発点となることも多い(Ushioda, 2009)。よって注意しなければならないことは、質的研究においては、トランスファラビリティの高さ(=データや研究結果の普遍性が高いこと)と、優れた研究であることとの間に線形的な関係が成り立たない(場合が多い)ことである。トランスファラビリティを評価することの意義は、研究間の比較を通して(1)共通の理論やモデルが適用できる範囲と(2)状況によって可変的な要素について、より合理的な判断を下すことが可能となることにある。先ほどのポートフォリオの事例では、例えば教育工学や心理学、または応用言語学の研究領域で広く受け入れられている理論や知見を土台として研究をデザインし、それらの事前情報と特定の状況や調査協力者を対象として観察された事例を比較することで、各理論やモデルの特定の研究テーマにおけるトランスファラビリティを評価することが可能となる。また、研究テーマによっては、可変的な要因(例えば

学習環境)が、学習の過程や成果に最も影響する要素となることもあるだろう(Freed, Segalowitz, & Dewey, 2004)。トランスファラビリティの高さと研究の質が必ずしもイコールにはならないのは、このような理由によるものである。

この様に、比較可能性を高め、トランスファラビリティの評価を可能とする厚い記述は、質的研究の妥当性や科学性を保証する上で重要な要素となる。では、我々は、実際の調査においてどのような基準を用いて研究と関連しそうな情報を選定し、調査対象とすれば良いのだろうか？非常に正統的かつ効率的な方法は、自分が研究上の関心を持つ分野の先行研究を読み込み、先行研究が使っている枠組みや情報を参考とし、自分の調査用に変更を加えることだろう(Charmaz, 2003)。先行研究の多くが、査読や多くの議論、もしくは十分な改訂や編集作業を経て発表されたものであるとすれば、それらの研究が用いた理論や方法論には、一定の妥当性が期待できる。よって、それらを土台に調査対象項目の選定をおこなうことは、(反論もあるだろうが)ある程度適切な方法であるといえる。これとは逆に、文献の読み込みが足りていない状況で研究を始めることは、先行研究での議論を考慮していないために重要なデータが調査対象から外れる可能性と、調査の効率が悪化するリスクが高まることになるだろう(三上, 2014)。

3. 交絡 (Confounding)

比較可能性とトランスファラビリティの観点から厚い記述を考えた際に有効と考えられる情報選定の枠組みの一つは、自らの研究に含まれるだろう交絡変数 (Confounding Variable もしくは Confounder) の存在を考慮し、交絡 (Confounding) の観点から見て特に重要と思われる情報を調査対象に加えることである。本章では、まずは交絡 (変数) とは何であるのかを理解するために、その解説をおこなう。なお、交絡変数は共変量や剰余変数とも呼ばれる (星野・繁樹, 2004) が、本稿ではまとめて交絡変数と呼ぶこととする。

3.1 交絡とは？

交絡は、(その存在が質的研究で述べられていることもあるが) 特に量的な研究で頻繁に扱われる概念なので、用語として耳なれない読者もいるかもしれない。そのため、ここでは先ず交絡という概念について、量的な研究の話を交えながら基本的な紹介をおこなう。本稿では、あまり本雑な説明や議論は避けるが、逆に交絡について更に詳しく知りたい読者の方は、各所に詳しい解説をおこなった文献を示しておいたので、必要に応じてそれらを参照されたい。

量的な研究においては、研究目的を達成するために推測統計が頻繁に用いられる (Lowie & Seton, 2013; 竹内・水本, 2012)。推測統計とは、大きくいえば、ある集団の平均的な傾向を明らかとしたいのだが、興味のある集団が非常に大きい (もしくは、何らかの意味で希少である) 時に、「集団全体 (母集団という) の中から、それを代表できる集団 (標本という) を調査対象としてデータを取り、そこで得られたデータの分析結果から、

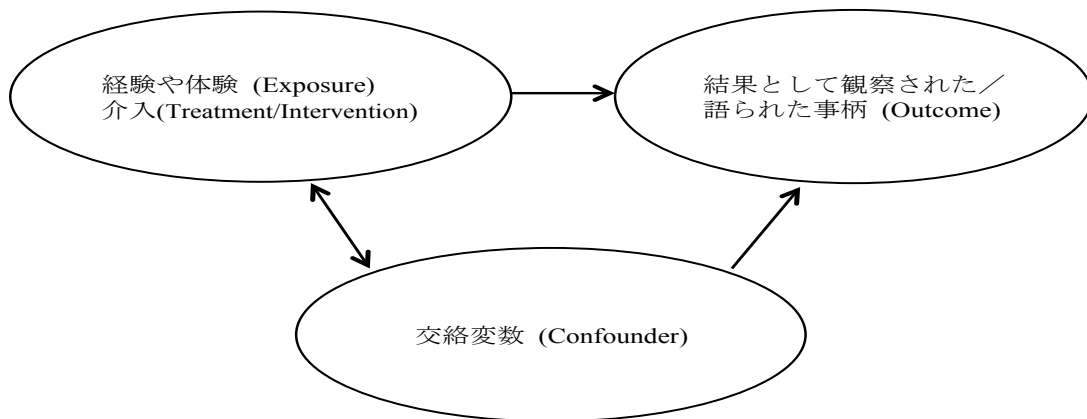


図 1. 交絡の概念図

集団全体での傾向を予測する（三上, 2014, p. 3）」方法のことである。また、推測統計を使う時に、研究者は、しばしば事象間の（因果）関係の解明に興味を持っている。図 1 は、因果関係を図式化したものである。この図の上部には、2 つの楕円がある。この内、左側にある楕円が原因（経験や体験・介入）であり、右にあるものが、結果（結果として観察された／語られた事柄）である。なお、医療などの領域では、Exposure を（薬剤などへの）暴露と訳すことがある（星野・繁樹, 2004）が、筆者は教育関係の人間なので、本稿においては「経験や体験」を訳とした。また Outcome は、単純に結果と訳しても良かったのだが、質的研究では、収集されるデータが、主にインタビューや（記述式の）質問紙から得られた文字データとなるため（Dörnyei, 2007; Smith, 2003）、図 1 では「結果として観察された／語られた事柄」とした（この後の説明では、必要に応じて Outcome とも呼ぶことにする）。さて、本題に戻ろう。量的な研究においては、データの分析に数値データが用いられ（三上, 2014; 竹内・水本, 2012）、図 1 の左の楕円（経験や体験・介入）が、右の楕円（Outcome）に与えている影響の程度を明らかとしようとする（左から右に流れている矢印は、因果関係を意味している）。

先程のポートフォリオの例を使って、この因果関係を説明してみよう。仮に、ポートフォリオを学校教育の現場で活用することで、生徒の学習意欲をどの程度高めることが出来るかに興味があるとしよう。図 1 でいえば、左上の楕円が「ポートフォリオの活用」であり、右上の楕円は「学習意欲（の程度）」である。これを質的に研究する場合、例えば、ポートフォリオを活用した英語学習を生徒に「体験」してもらい、プログラムが完了した時にインタビューなどを用いて、学習意欲がどう変化したかを生徒に「語ってもらう」ことができるだろう。

量的な研究では、図 1 で示されたような因果関係を調べるために、調査協力者を一定数募り²、例えば彼／彼女らを実験群（ポートフォリオを活用するグループ）と統制群（ポートフォリオを活用しないグループ）に振り分ける。この上で、2 つのグループの調

査開始時と完了時の学習意欲の程度も、例えば 6 件法の質問紙などを使って記録しておき、グループ間での最終的な学習意欲の伸び率を比較することで、ポートフォリオの使用が、学習意欲に与える影響を検証することができる（もちろん、他の調査方法や統計的手法を使うことも考えられるし、データの収集を質的におこなっても差し支えない）。さて、ここでの統制群とは、実験群の生徒達が特定の経験や体験をしなかった（＝ポートフォリオを活用しなかった）状況を人為的に作り出したものである（Yamamoto, 2012）。このような方法を用いて「ある経験や体験の有無」が Outcome に与える影響を適切に検証するために、研究者は、調査対象となる特定の「経験や体験・介入」の有無以外の点では性質が（比較的）似通った学生達に調査に協力してもらい、その上で統制群と実験群を作る必要がある（Gum, Thamilasan, Watanabe, Blackstone, & Lauer, 2001; Rosenbaum & Rubin, 1983）。このような調査対象者間の均質化を実現するために、研究者は、調査対象者の選び方（サンプリングという）やデータ分析の方法を工夫し、興味のある事象について、可能な限り純粋な関係を描こうとする（Brown, 2006; 狩野 & 三浦, 2002; 村山, 2012）。

可能な限り純粋な関係を描こうとする過程で、研究者は、研究対象となる経験や体験・介入以外の要素で、かつ生徒の学習意欲に影響を与える要因について考え、それらの影響を調整（Adjusting）する。この、調整されるべき要素のことを交絡変数とよぶ（図 1 の中央下の楕円）。より正確には、交絡変数は、Outcome には影響関係をもっていて（交絡変数から影響関係を示す矢印が伸びている）、かつ経験や体験・介入とは相関関係（どちらかの要素が片方に一方的に影響してはならず、相互に影響する関係）にある（Rosenbaum & Rubin, 1983; 浦島, 2004）³。

再び、ポートフォリオの例を使って交絡変数の説明を試みよう。今回は、これまでの経験や体験・介入と Outcome の関係に加えて、自己調整学習能力という交絡変数の存在を仮定してみる。自己調整学習能力というのは、文字通り、より効果的な学習を学習者自身で組み立てることが出来る力のことである（Zimmerman, 2006）。自己調整学習をおこなうための条件のひとつとして、学習者は、自分の学習の進捗や効率性を効果的に自己観察（内省）できる必要がある。ここで、先行研究の結果から、自己調整学習能力が高い人は、学習意欲も高まりやすいことが分かっているとすると（つまり、正の関係があるとする）。また、自己調整学習能力が高い学習者は、ポートフォリオの様に形式が定まったものではなくても、学習の振り返り（内省）を既に自己流でおこなっていると仮定しよう。この場合、調査の開始時点で既に高い自己調整学習能力を有する学習者は、そうでない学習者と比べて、ポートフォリオという内省方法により向いている（＝つまり、正の関係がある）可能性がある。この筋書きに従えば、自己調整学習能力は、ポートフォリオを使った学習と学習意欲の関係を描く上での交絡変数となる（ここでは話を簡単にするために、自己調整学習能力以外の全ての交絡変数は、調整済みだと仮定しよう）。さて、ここで更に、実は自己調整学習能力は、今回のポートフォリオを用いた学習と比べて学習意欲の向

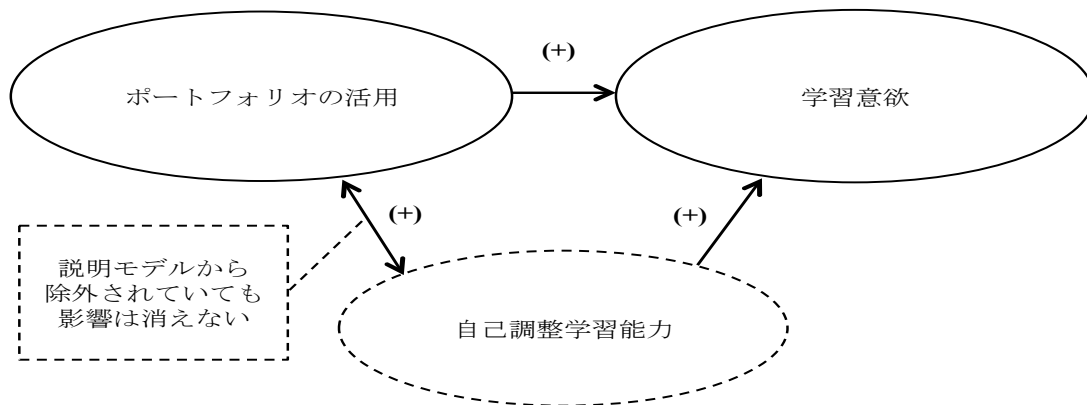


図2. 交絡による過大評価の例

上に2倍も大きな影響力を持っていたと仮定しよう。この場合、自己調整学習能力の存在を考慮しないデータ（Crude Data）を使って、ポートフォリオを使った学習と学習意欲の変化の関係を調べると、データ分析の結果は、以下の一連の理由から、ポートフォリオ使用の有効性を過大評価することになる（図2）：（1）ポートフォリオを使った学習への適合性と自己調整学習能力は正相関し、また、そのどちらも学習意欲と正の関係にある；（2）だが、自己調整学習能力の方が、学習意欲の変化への影響力は2倍強い；（3）自己調整学習能力は、今回の説明モデル（＝ポートフォリオを活用した学習が学習意欲に与える影響を説明するための枠組み）には組み込まれていない；（4）自己調整学習能力が（ポートフォリオの活用を介して）学習意欲に間接的に与えていた影響によって、今回の説明モデルで確認されたポートフォリオの使用と学習意欲増加の関係は、本来あるべき姿よりも強く描かれた。この様に、説明モデルで見逃されている要因が分析結果を曇らせることを交絡という（Rosenbaum & Rubin, 1983）。因果関係を曇らせる要因は、総じてバイアス（Bias）と呼ばれ（Delgado-Rodríguez & Llorca, 2004）、そのため交絡は、交絡バイアス（Confounding Bias）と呼ばれることもある。本稿の最終的なまとめでも触れるが、交絡は、大きく分けて3種類あるバイアスの一種である。また、交絡変数、経験や体験・介入、Outcomeそれぞれの関係が正か負であるかによって、交絡が起きた際に因果関係を過大評価するか過小評価するかは変化する。ただし、交絡変数を見逃すことで分析結果が曇ることが把握できれば、本稿のこの後の議論は十分理解できる。そのため、過小／過大評価についてより詳しく知りたい場合、浦島（2004）が（分野は医療系だが）特に分かり易い説明をしているので、参考とされたい。

3.2 交絡の問題点と質的研究への応用

交絡が、量的な研究にとって問題であることは明らかだろう。煎じ詰めれば、量的な研究において交絡の問題を無視すれば、興味のある現象を正確に捉えることができないの

である (Gum et al., 2001; Rosenbaum & Rubin, 1983)。この問題に対処するために、量的研究では、サンプリングやデータ分析に関して、方法論の精緻化や対処法の提案がされてきており、現在も、その議論は続いている (Brown, 2006; Oswald & Plonsky, 2010; Plonsky, 2013; 星野 & 繁樹, 2004)。

交絡の存在を無視することで分析の結果が曇ることは、実は質的な研究においても同様である (村山, 2012)。その理由は単純で、例えば図1や2のような関係を描くことが目的の研究であれば、使用するデータの種類を (数値データから言語データに) 変更しただけで交絡の影響が統制されると考える理由がないためである。そのため、まずは非常に現実的な理由から、事象間の (因果) 関係を取り扱う限り、交絡は常に意識されるべき問題となる。

しかし、質的研究において交絡の存在を意識する理由には、より積極的なものもある。交絡について考えることは、厚いデータ報告を達成するために有効である。まずは、研究をデザインする段階において、交絡の存在を意識しながら先行研究の読み込みをおこなうことは、先行研究が使っている枠組みや、そこで報告されているデータが、研究の妥当性をどの様に保証しているのかについて、より実用的 (Practical) な視点から考えることを可能とする。また、交絡の存在を念頭において先行研究を読み込むことで、過去の研究では言及されていない (が重要と考えられる) 交絡変数の存在について思い当たることもあるだろう。これらの変数を自らの研究の調査対象とすることは、研究の比較可能性を上げることに繋がる。先ほどから触れている通り、質的研究では、これらの変数について必ずしも数値データを収集する必要はない (三上, 2014)。例えば、交絡変数になりそうな要素について、インタビューや自由記述式の質問紙から言語データを収集することが可能である。また、例えば量的な分析と質的な分析を統合した Mixed Methodology を採用することに意味がある場合、言語データに加えて、質問紙や言語テストを用いて数値データを収集することもあるだろう (Cohen, Manion, & Morrison, 2007)。また、交絡の概念を意識することで、量的な研究において報告された知見を、より積極的に質的な調査に取り込むことが容易となる。

以上の理由から、自分の研究テーマに関連する潜在的な交絡変数に関する情報を調査の対象として、それらの情報を選定理由と共に論文の中で報告することは、厚い記述をおこなう上で有効な方法となる。これらの情報に加え、データの分析に用いた理論的枠組みや (Allen, 2010; Smith, 2003)、分析の道筋や他の研究結果との類似点・相違点を議論することで、論文全体で厚い記述を達成することが可能となる (Charmaz, 2003; Richards, 2005; 戈木クレイグヒル, 2013)。更に、交絡についての情報を提供することで、今度は量的な研究に慣れた者が、質的な研究から学ぶことも容易とするだろう。先ほども述べた様に、質的研究の目的は、研究結果の一般化 (Generalization) にはないことも多い。しかし、交絡という量的質的の別を問わず研究の結果に影響を与える要因についてのデータを収集し報告することで研究の比較可能性は高まり、トランスフ

アラビリティの評価も可能となる。住(2011)で紹介されている様に、質的・量的という点で異なる枠組みを用いる研究者が、互いの研究の妥当性や科学性について適正な評価を下すことは時として困難であるが、交絡について考察することは、その架け橋となりえる。

以上の通り、交絡について考えることは、質的研究において厚い記述をおこなう上で有用である。次の第4章では、本稿のまとめとして、調査をデザインする段階で交絡変数を選定する具体的な方法を紹介する。

4. 調査のデザイン段階で交絡変数を選定する方法

本稿のまとめとして、交絡変数を選定する手順の一例を紹介する。交絡変数の選定は、(専門的な知識が不可欠なので) 量的・質的研究の別を問わず文献の読み込みから始まると考えてよい。この過程を経て、研究者はまず、交絡変数の候補をリストアップする。ただし、交絡変数候補の数は、文献を読むにつれて増加し続けることが予想される。そのため、交絡変数と研究対象の関係性や交絡変数の影響力について、ある程度体系的にまとめておく・図式化しておく、後の研究作業がより効率的になる。

交絡について(1) 具体的なモデルが先行研究で提案されている場合は、そのデータを参考として交絡変数、経験や体験・介入、Outcome の関係を整理すれば良く(星野・繁柘, 2004)、(2) 具体的なモデルが見つからない場合は、想定できる範囲内で関係の強弱を予測しておく。予想に特定の筋道がある訳ではないが、ここでは Connors et al. (1996) で使用された方法に多少の変更を加えたものを紹介する(図1を参照しながら読まれと良い)。例えば、各交絡変数候補について(1) Outcome に関係が全くないとはいえない、(2) 経験や体験・介入よりは関係が弱そうだが、Outcome とある程度関係するだろう、(3) 経験や体験・介入と同程度か、またはそれ以上に Outcome と関係する可能性がある、の三段階程度で評価を付けておくと、調査の対象とする変数を限定する必要が生じた際に、その絞り込みが楽になる。また、図1に示されている様に、交絡変数と経験や体験・介入の間にはある程度の相関が期待されるべきだが、仮に両者の間に弱い相関しか期待できない場合でも、交絡変数と Outcome 間では強い関係性が予測される場合、それらの変数は説明モデルに組み込んだ方が良くとされる(Westreich, Cole, Funk, Brookhart, & Stürmer, 2011)。よって、この三段階の予測をする作業では、経験や体験・介入と交絡変数の関係も考えつつ、特に交絡変数が Outcome に与えている影響の強さを念頭において予測を立てると良いだろう。なお、相関の強弱の解釈は、研究分野によって変わるため、どれくらいの数値であれば意味があるのかについて、親指ルール(Rule of Thumb)はない(例えば、水本, 2012)。その解釈は、先行研究の文脈からなされるべきであり、その意味でも、文献の読み込みは重要である。

また、特定の研究分野の専門家に交絡変数のリストを作成して貰い、自らが作成したリストと見比べることで、変数候補に見落としがないかを確認することも有効である

(Connors et al., 1996)。Appendix の質問紙は、この様なリスト作りを依頼する際に、どの様なリストを作成して欲しいかを伝えるための例である。本題にどの様な設問を使用するかは、研究対象次第といえるが、基本的には、例題の形式を応用して作成が可能だろう。例題では、散歩をすることが肥満の解消につながる例を使用し、散歩量と肥満の程度の間関係を描く上での交絡変数例を挙げている。先程の例と同様、交絡変数には、経験や体験・介入および Outcome との関係から三段階の予測評価が与えられることになる。筆者の個人的な経験則では、(協力者の手間や回答の実用性などから考えて) 三段階程度の予測が使いやすそうであるが、予測の段階は、より多くても少なくても構わないだろう。なお、ここで挙がってくる交絡変数候補についても、経験や体験・介入とは弱い関係しか仮定できないものがあるだろう。その様な候補についても、最終的には交絡変数が Outcome に与えている影響の強さを考え、その取扱いを決めると良い。

リストの作成については、調査対象に近い性質を持つ人々に協力してもらうことも考えられる。ポートフォリオの例でいえば、言語能力や学習意欲の点で調査対象者に近い学生達に協力を仰ぐことが考えられる。ただし、この様な人々にリストを作成して貰う際には、学習者と専門家が、(状況によってはかなり)異なる視点からリスト作成に取り組む可能性があることを考慮すべきである。医療分野での例を用いれば、例えば、ある病気にかかった患者は、その病気を体験した者でなくては分からない情報を提供してくれるだろう(例えば、病気になった際の心理状態など)(Charmaz, 2003)。一方で、専門家(この場合は医師)は、同じ病気について、例えば白血球数や血圧の変化などのより客観的/学術的な視点からリストを作成する可能性がある(平石順, 2015年2月9日のコメント)。よって、どちらの立場の人々であっても、リスト作成に協力してもらうことは可能であるが、それぞれが異なる立場から同じ事象を観察する可能性は考慮しておくべきである。最終的にどの様な対象者にリストの作成を依頼するかは、研究の目的によることになる。もちろん、諸般の事情から、研究者自身が作ったリストのみを使って交絡変数を選定することもあるだろう。

また、調査をおこなっている最中に(例えばインタビューの最中に)、新たな交絡変数の存在に気が付くこともあるだろう。この様な場合、幸い質的な研究には、データ収集と分析が並行して行われても良いという特徴があるため(Richards, 2005などを参照のこと)、特定の項目について、その場で新たな質問をすることや、既にデータを採り終った調査協力者についても、その項目について、再度質問をすることも可能である。

5. おわりに

本稿では、質的研究において厚い記述をおこない、研究の比較可能性を向上させトランスファラビリティの評価を可能とするための一方法として、交絡の側面から調査対象情報を選定する方法を紹介した。3.1でも触れたが、交絡は、とても大きく分けて3種類あ

るバイアスの一種である（ここでのバイアスとは、研究結果を曇らせる要因全般のこと）。本稿では、最後に残りの代表的なバイアスである選択バイアス（Selection Bias）や情報バイアス（Information Bias）（Delgado-Rodríguez & Llorca, 2004）についても簡単に触れ、バイアスと厚い記述の関係について考えたい。

選択バイアスとは、調査協力者達の言語能力や性質、ふるまいが、例えばサンプリングの偏りのせいで、母集団を代表しないことを示す（Brown, 2006）。質的な研究では、調査協力者が必ずしも母集団を代表する必要はない（Richards, 2005）。しかし、研究上の興味と関係のある範囲内で、彼／彼女（ら）の情報を厚く記述しておくことは、やはり研究の比較可能性を高める上で重要である。

情報バイアスは、集められたデータに歪みがあるために、正確な分析をおこなうことが出来ない状況である。例えば、人は、同意することが反社会的と見なされるような質問をされた時に本心と異なる回答をすることがある（Dörnyei & Taguchi, 2010）⁴。また、正直な回答をおこなった場合でも、時間が経ってしまった出来事や昔の感情を正確かつ詳細に思い出すことは大変困難である（Dunlosky & Metcalfe, 2009; Gass & Mackey, 2000）。研究者自身の先入観も、データの解釈に影響するため、研究目的によっては情報バイアスの温床となりえるだろう（三上, 2014）。（言語）能力の測定についても、測定機に固有な・測定者に固有な・外部的な・偶発的な理由から、測定誤差が生じる（Brown, 1999, 2001）。これらの情報バイアスを可能な限り取り除いた研究をおこなうためには、先行研究や幅広い領域の専門書を読み込み、研究手法について正確な知識を身に付けることが必要である。適切に選択された調査方法について順を追った正確な記述をおこなうことは、勿論、より厚い記述をおこなう上で重要である。

これらに加えて交絡（バイアス）について考えることは、質的な研究において厚い記述を実現し、研究結果の比較可能性を高め、トランスファラビリティの評価を可能とする上で有効である。本稿では、質的な研究の調査対象情報を交絡の側面から考え選定する方法について提案をおこなった。また、交絡変数が Outcome に与える影響の程度に考えを巡らすことで、報告すべき情報の優先順位を決定することも可能となるだろう。

最後になるが、本稿は、いわば「質的研究における交絡 101」の様な位置づけとなるものであり、今後より具体的な分析手法を提案して行きたいと考えている。本稿の情報が、他の研究者の調査方法構築の一助になったとすれば、これに勝る喜びはない。

謝辞

本稿の着想は、中部大学の塩澤正教授、名古屋学院大学の中山麻美先生、名古屋大学の平石順久先生と開いている勉強会から得たものである。また、本稿の執筆に当たって、先生方から多くのアドバイスや助力を頂いた。ここに記して感謝したい。

注

1. 多くの場合、データの記録者は分析者自身であろうが、そうでない場合も想定できるので、ここでは、分析者とデータの記録者を別個の存在としている。
2. 量的な研究においてどれだけの調査対象者数（標本サイズという）があれば十分かということは、調査結果の一般化可能性（Generalizability）と関連する話である。この点に興味がある方は、以下の文献などを参考とされたい（Brown, 2006; 水本・竹内, 2011）。
3. なお、この関係が意味していること（例えば調査対象者の割り当て問題など）をより詳しく知りたい読者は、以下に示す参考文献を読まれると良い（Rosenbaum & Rubin, 1983; 星野・繁樹, 2004）。
4. 例えば「イライラしたら人を殴ったり小動物などを虐待したりしても構わないと思う」という問いに（本心でどう思っているとしても）強く同意することは、通常は考えにくい。

参考文献

- Allen, H. W. (2010). Language-Learning Motivation During Short-Term Study Abroad: An Activity Theory Perspective. *Foreign Language Annals*, 43, 27—49.
- British Council. (2007). Portfolios in ELT. Retrieved from <http://www.teachingenglish.org.uk/article/portfolios-elt>
- Brown, J. D. (1999). Standard error vs. Standard error of measurement. *Shiken: JALT Testing & Evaluation SIG Newsletter*, 3, 20—25.
- Brown, J. D. (2001). *Using Surveys in Language Programs*. Cambridge: Cambridge University.
- Brown, J. D. (2006). Statistics Corner. Questions and Answers about Language Testing Statistics: Generalizability from Second Language Research Samples. *Shiken: JALT Testing & Evaluation SIG Newsletter*, 10, 24—27.
- Charmaz, K. (2003). Grounded Theory. In S. J. A (Ed.), *Qualitative psychology: a practical guide to research methods* (p. 81—110). London: Sage.
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research Methods in Education* (6th ed.). London: Routledge.
- Connors, A. F. J., Speroff, T., Dawson, N. ., Thomas, C., Harrell, F. E. J., Wagner, D., ... Knaus, W. . (1996). The Effectiveness of Right Heart Catheterization in the Initial Care of Critically Ill Patients. *The Journal of the American Medical Association*, 276, 889—897.
- Council of Europe. (2011). European Language Portfolio. Retrieved from <http://www.coe.int/t/dg4/education/elp/>

- Delgado-Rodríguez, M., & Llorca, J. (2004). Bias. *Journal of Epidemiology and Community Health*, 58(8), 635–641.
- Dörnyei, Z. (2007). *Research Methods in Applied Linguistics*. Oxford: Oxford University Press.
- Dörnyei, Z., & Taguchi, T. (2010). *Questionnaires in Second Language Research: Construction, Administration, and Processing* (2nd ed.). New York: Routledge.
- Dunlosky, J., & Metcalfe, J. (2009). *Metacognition*. London: Sage.
- Freed, B. F., Segalowitz, N., & Dewey, D. P. (2004). Context of Learning and Second Language Fluency in French: Comparing Regular Classroom, Study Abroad, and Intensive Domestic Immersion Programs. *Studies in Second Language Acquisition*, 26, 275–301.
- Gass, S. M., & Mackey, A. (2000). *Stimulated Recall Methodology in Second Language Research*. New Jersey: Lawrence Erlbaum.
- Gum, P. A., Thamilarasan, M., Watanabe, J., Blackstone, E. H., & Lauer, M. S. (2001). Aspirin Use and All-Cause Mortality Among Patients Being Evaluated for Known or Suspected Coronary Artery Disease. *The Journal of the American Medical Association*, 286, 1187–1194.
- Hyland, K. (2009). *Teaching and Researching Writing* (2nd ed.). London/New York: Routledge.
- 狩野裕, & 三浦麻子. (2002). 『グラフィカル多変量解析：目で見る共分散構造分析 (増補版)』現代数学社.
- 三上仁志. (2014). 積極的な理由から質的研究を行うために—その背景と科学性の保証について—. 外国語教育メディア学会中部支部外国語教育基礎研究部会 2013 年度報告論集, 1–20.
- 水本篤. (2012). 質問紙調査における相関係数の解釈について. 外国語教育メディア学会 (LET) 関西支部メソドロジー研究部会 2011 年度報告論集, 63–73.
- 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—. 外国語教育メディア学会 (LET) 関西支部メソドロジー研究部会 2010 年度報告論集, 47–73.
- 星野崇宏, & 繁榊算男. (2004). 傾向スコア解析法による因果効果の推定と調査データの調整について. *行動計量学*, 60, 43–61.
- 村山航. (2012). 妥当性 概念の歴史的変遷と心理測定学的観点からの考察. *教育心理学年報*, 51, 118–130.
- Lowie, W., & Seton, B. J. (2013). *Essential Statistics for Applied Linguistics*. Basingstoke: Palgrave MacMillan.
- Oswald, F. L., & Plonsky, L. (2010). Meta-analysis in Second Language Research: Choices and Challenges. *Annual Review of Applied Linguistics*, 30, 85–110.

- Plonsky, L. (2013). Study Quality in SLA: An Assessment of Designs, Analyses, and Reporting Practices in Quantitative L2 Research. *Studies in Second Language Acquisition*, 35, 655—687.
- Richards, L. (2005). *Handling qualitative data: A practical guide*. London: Sage.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70, 41—55.
- SLAA 研究会. (2013). 質的研究—基礎 : 定義・特徴量的研究との比較他.
- 戈木クレイグヒル滋子. (2013) 『質的研究法ゼミナール: グラウンデッド・セオリー・アプローチを学ぶ (第二版.)』 医学書院.
- Smith, J. A. (Ed.). (2003). *Qualitative psychology: a practical guide to research methods*. London: Sage.
- 住政二郎. (2011). 質的研究の科学性に関する一考察. 外国語教育メディア学会(LET) 関西支部メソドロジー研究部会 2010 年度報告論集, 30—44.
- 竹内理, & 水本篤 (Eds.). (2012) 『外国語教育研究ハンドブック—研究手法のより良い理解のために』 松柏社.
- 浦島充佳. (2004). 交絡因子. Retrieved from <http://dr-urashima.jp/pdf/eki-11.pdf>
- Ushioda, E. (2009). A Person-in-Context Relational View of Emergent Motivation, Self and Identity. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, Language Identity and the L2 Self* (p. 215—228). Bristol: Multilingual Matters.
- Westreich, D., Cole, S. R., Funk, M. J., Brookhart, M. A., & Stürmer, T. (2011). The role of the c-statistic in variable selection for propensity score models. *Pharmacoepidemiology and Drug Safety*, 20, 317—320.
- Yamamoto, T. (2012). Understanding the Past: Statistical Analysis of Causal Attribution: Business Source. *American Journal of Political Science*, 56, 237—256.
- Zimmerman, B. J. (2006). Development and Adaptation of Expertise: The Role of Selfregulatory Process and Belief. In A. Ericsson, N. Charness, P. Feltovich, & R. Hoffman (Eds.), *Handbook of Expertise and Expert Performance* (p. 707—724). New York: Cambridge University Press.

Appendix : 交絡変数の選定に使用可能な質問紙の例

例題

毎日散歩をすることが、肥満をどれだけ抑制するかに興味があるとします。
そして、量的な手法を使い、散歩と肥満（抑制）の関係を明らかにしようとしたとします。
この場合、

【散歩の量】→【肥満の抑制度】

という関係を明らかにしようとしています。

しかし、肥満に関係しそうな要素は、散歩の他にも色々と考えられます。

例えば、食べ過ぎは、肥満の原因となるでしょう。

逆に、定期的に運動していたら、散歩はしなくても肥満にはならなそうです。

そこで、質問です。

質問例)

①散歩の他に、肥満と関係しそうな要素を、思いつく限り挙げてください。

似た内容の事柄が重複しても構いません。

②挙げた候補のそれぞれについて、どれくらい強く**肥満の抑制**と関係しそうかを、

以下の3段階から評価をして下さい。

- 1 : 関係は弱いかも知れないが、関係がないとはいえないと思う
- 2 : 散歩程ではないかも知れないが、関係していると思う
- 3 : 散歩と同じ位か、それ以上に関係していると思う

解答例)

評価	肥満の解消に役立つような事柄
3	その人が散歩以外に日常的にしている運動の量
3	その人の一日の摂取カロリー
2	遺伝的な要素（＝親がどれ位太っているか／いたか）
2	その人の年齢
1	その人の収入
1	営業職に就いているかどうか（飲み会が多そうなので）
