

## 「心理学研究法総論」をどう教えるか(II)

著者	吉村 浩一
出版者	法政大学文学部
雑誌名	法政大学文学部紀要
巻	79
ページ	111-125
発行年	2019-09-30
URL	<a href="http://doi.org/10.15002/00022419">http://doi.org/10.15002/00022419</a>

# 「心理学研究法総論」をどう教えるか（Ⅱ）

吉村浩一

---

## 要約

---

実験法と調査法を扱った前号に引き続き、本号では、検査法、観察法、面接法に関する重要な問題のいくつかを指摘し、筆者自身の見解を示す。検査法の章では、調査法との違い、心理テストの標準化の難しさ、テストバッテリーを構成することの重要性を説明する。観察法の章では、音声や動画の記録システムが観察法のあり方をどのように変えたか、参与観察とエスノグラフィーの関係、KJ法（川喜田、1967、1986、1997）の特徴を解説する。面接法の章では、デプス・インタビューとグループ・インタビューの特徴と、犯罪捜査に用いられる認知インタビューの特徴を解説する。最後の章では、サンプルサイズの決め方と心理学では2種類のQDAという用語がむしろ反対の意味で用いられていることを指摘する。最後に、学生が卒業研究を行うときの姿勢についてコメントする。

キーワード：心理学研究法、検査法、面接法、観察法

---

## 5. 検査法

### 5.1 調査と検査の違い

調査と検査はともに質問紙を用いて行うのが一般的なので、実施上の共通点は多い。心理学で行う調査と検査は、本質においてどう違うのだろうか。実際には両者の区別は連続的なものと言うべきだが、典型例を考えると、調査は、独自に作成した質問を用いて、あることがらに関する意見や態度を捉えようとするのに対し、検査とは、評価の定まった（尺度化された）質問に回答を求め、その点数や回答パターンから被検査者を査定するものと言える。

「評価の定まった（尺度化された）」とは、どのような意味なのか。一言で言えば、一定の質問群で構成される質問票を用いて、対象（者）のある心的特性または状態を測るための物差しとして使えること言う。いわゆる、標準化されたテストである。たとえば、性格という心的特性を測るため

にさまざまな性格テストが開発されているが、そのいくつかは「標準化されたテスト」である。「標準化」されたものであることの最低要件は、「妥当性」と「信頼性」を有していることである。これら2つは、心理テストを学ぶ際、よく耳にする用語である。そのテストが測りたいものを測れていること（妥当性）と、測るたびに異なる結果になることなく安定・一貫して対象を測れていること（信頼性）は、標準化に向けての基本条件である。そうしたチェックを経た上で、その検査の結果（得点など）を評価する際の基準をもって、物差しとして使えるのが「標準化されたテスト」である。こうした条件を満たしていることが必要だと知れば、何らかの心的性質を捉えるために自分で適当に考えた質問群を気軽に「…を測定するための尺度」とは呼べないことが理解できよう。次節では、心理検査に用いられている心理テストを例に、心理テストの標準化過程を見ていく。

## 5.2 心理テストの標準化：MMPI 新日本語版の標準化作業を例に

上述のように、調査の際の質問票とは違い、検査に用いる質問票は標準化されたものであることが多い。パーソナリティを測定する心理検査の1つであるMMPI新日本語版の標準化作業を例に、標準化された検査を作る過程を見ていきたい。

MMPIとは「ミネソタ多面人格目録」のことで、その名の通り、ミネソタ（大学病院）で開発された英語のパーソナリティ・テストである。すなわち、アメリカ合衆国で用いられているパーソナリティ・テストである。その日本語版を作るには、英語を適切な日本語に翻訳するだけでよいと思う人がいるかもしれないが、実はそうではない。文化が違えば、同じ質問でも「是認率」が異なりうる。換言すれば、アメリカ人には当たり前の態度や行動が日本人には避けるのが常識的ということがある。人格を多面的に査定するため、MMPIは550項目もの多項目で構成されており、それら1つ1つに対し日本人の是認率を測定し、アメリカでの値と付き合わせていく。しかも、日本語訳の微妙な違いによって是認率が変わる可能性もあるため、何種類かの日本語訳を比較しながら進めなければならない。

標準化を行うためには、標本抽出にも配慮を要する。通常の心理学調査では、ある授業の受講者など身近な人に回答してもらうことが多いが、それでは真の意味での「日本語版」とならない。「MMPI新日本語版」では、居住する地方・年齢・職業・教育歴を性別ごとに国勢調査に基づく人口比で按分し標本抽出がなされた。

もちろん、このような大規模な標準化作業は一研究者では行えず、「MMPI新日本語版」の場合、「MMPI新日本語版研究会」が組織され、手分けしてサンプル抽出が行われた。代表者である田中(1997)を中心に、その標準化過程が残されている。既存テストの日本語版作成でさえこれだけの作業が必要なことから、「尺度を作成する」と安

易に言うことは控えなければならない。

## 5.3 テストバッテリーという考え方

標準化のため、多くのしかもいろいろな人の得点や是認率などを調べていくと、平均値付近の値をとる人が多く、その値から隔たるにつれて度数が少なくなる、いわゆる「正規分布」になることが広く認められる。このことは、心理機能についてもある程度当てはまるが、身体機能を例に用いると、より当てはまりがよい。たとえば、誕生後の赤ちゃんの月齢を横軸に、体重を縦軸にとると、体重の成長過程を追うことができる。こうした発達成長曲線には平均値だけでなく、しばしば「90パーセンタイル値」という指数が描き込まれている。この指標があれば、正常範囲をかなりの幅をもって捉えることができ、少しぐらい平均からずれていても過剰に反応せずにすむことになる。しかし、90パーセンタイル値から外れると、「異常」と見なすことにもなる。しかし、「要注意」との警戒を発することの意義は重要だが、「異常さ」を平均値からの逸脱のみで評価することには、別の意味で慎重であるべきである。日本人の虫歯の平均本数は0ではない。おそらく0本は90パーセンタイル値からも外れているだろう。しかし、虫歯0本の人は決して異常でない。分布上の位置だけではなく、適応性の基準も正常—異常の評価には必要なのである。

90パーセンタイル値から外れていることから短絡的に「異常」と判断せず、「要注意」との警戒のもと、さまざまな角度から検査を行い、最終評価を示すべきである。心理機能に対する「さまざまな角度からの検査」とは、どのように実行されるのだろうか。前節のMMPIは、その名の通り、病的状態を含め、性格のさまざまな面を多面的に査定できることをうたっているが、550項目の実施にはかなりの時間を要する上、おそらく注目点を見極めるのにあまり必要でない項目群も含まれているだろう。しかも、MMPIなどの質問紙法の検査だけでは見逃がされてしまうものもある。焦点を当てた事柄について多角的に検討する

には、たとえばMMPIは関係する一部だけを使い、性質の異なる別の検査結果と照らし合わせて査定することが有効である。たとえば、無意識レベルでの特徴を捉えるためTATなどの投影法テストと組み合わせるなどである。1つの心理検査で捉えられる心的特徴は限定的で、それぞれの検査には長所と短所がある。複数の心理検査等を組み合わせることで焦点となる心理機能を多角的に捉えるテストを構成することを「テストバッテリーを組む」と言う。テストバッテリーを組む作業は固定的に行えるルーチンワークではなく、指示を出す医師を含め、検査者の能力と研究心が求められる創造的作業である。

病院臨床にあっては、心理的問題・認知機能・神経学的所見を付き合わせ総合的に捉える力が問われる。テストバッテリーのさらなる拡張と言えるその例を示して、本節を終えたい。

あるとき、70代の女性が、最近ご主人を亡くし気分がふさぎ込み体調がすぐれないと、家族につれられて来院した。身体検査所見には特に異常はなく、精神的なことの関与が疑われたため、心理スタッフに依頼がきた。そして、臨床心理士が、その患者さんに文章完成テストを行った。その報告書には、「ご主人を亡くしたということについて同じことが何回も書いてあり、“裳の作業”ができていない」と記載されていた。転院してきたその患者さんの面接では、臨床心理士の所見通り、同じことを何度も繰り返していた。医師からの依頼で行ったウェクスラー成人知能検査では、同じことを応答するエラーが所々に認められた。このエラーパターンからは前頭葉機能障害も疑われるため、ウィスコンシン・カード分類テストやストループ・テストなどの前頭葉機能検査を施行した。その結果、保続や抑制障害という前頭葉機能障害の特徴を認めた。画像所見からも、前頭葉の血流低下が認められ、医師はその他の臨床症状から前頭葉型痴呆と診断した。文章完成テストでみられた「ご主人を亡くしたということが何回も書

いてある」という所見を、臨床心理士は、患者さんの生活背景から、「裳の作業ができていない」と考察した。確かに、その通りのことが起こってはいるのだが、病院臨床における心理スタッフには、臨床心理学的アプローチだけでなく、臨床神経心理学的アプローチも必要である。(吉村・高岩, 2006, p.23-24)

## 6. 観察法

### 6.1 音声・映像記録の利用

観察時に観察対象を映像・音声記録することが一般的となったのは、ホームビデオやカセット・テープレコーダが普及した1970年代以降である。それ以前は、観察現場で目や耳をフル稼働させ、メモをとったり、あらかじめ作成しておいたチェックリストにチェックを入れる作業を行っていた。音声や映像をテープに丸ごと記録できるようになったことは、観察現場での慌ただしい作業から研究者を解放し、生の様子をじっくり観察する余裕を与えてくれた。大衆化したビデオや音声レコーダが観察法を格段に発展させたのである。

しかし、観察法を武器に研究している研究者(大学院生などの学生も含め)の中には、安価なビデオテープやカセットテープ(今日ではデジタル・メモリ)を湯水のように投入し、記録されたテープからデータを生成しないまま何百、何千本もの記録媒体を研究室にため込んでいる人が少なからずいる。以前は、データ生成を観察現場でリアルタイムに行っていたが、現在では録音・録画されたテープを観察後にくり返し再生しながらデータを生成しなければならない。この時間消耗的な研究スタイルが、データ化以前のテープの山に埋もれる事態を生みやすくさせている。

動画や音声を記録できる前の時代は、現場での観察ですべてをデータ化できないことを当然とし、観察対象時間内に起こることがらのうち、何をいつ記録するかを限定することにより、「事象見本法」「時間見本法」などを用いるのが一般的であった。そのおかげで、逆に観察を終え現場を

離れるときには、データが生成されていた。それに対し、手軽に録音・録画できるようになると、現場から持ち帰ったテープを見直さなければデータは生成できない。それは、停止や巻き戻しを繰り返す度に何回も行う時間消耗的作業であるため、上述したように未データ化のままのテープの山に埋もれる事態を生まれやすくするのである。

こうした現実を考えれば、くり返し見直せるのだから「あらゆることをデータ化しよう」と欲張るのではなく、テープ記録以前の時代と同様に、「事象見本法」「時間見本法」でサンプリングして録画されたものを見ながらデータ生成するのが現実的である。テープ記録以前の時代とは違い、何をサンプリングするかを事前に決めておく必要はなく、テープを流し見しながらサンプリング対象を選定できる利点も手に入った。また、行った計数が正しいかどうかのチェックも、録画・録音テープを見直すことで可能である。

## 6.2 それでもサンプリングしないデータ収集を

分析対象となる事象を一部に絞ることはやむを得ないにしても、注目する事象が他の（記録から外された）事象と関連することから気づくことは大いにある。そこで、取りあげるべき事象を追加できるようにしておくことが重要である。そのためには、もとの録画・録音記録に戻れる体制を確保して置く必要がある。特に、限られたタイミングだけを記録対象にするのではなく、その事象の開始時点・持続期間・終了時点をすべてデータ化できていれば、他の事象との関係を的確に評価することが可能になる。

近年は、記録する事象を事前に決定しておくのではなく、必要に応じてあとから追加でき、かつ時間情報も連続的に記録できるコンピュータ・ソフトウェアが整いつつある。まだ高価なものが多いようだが、ビデオ編集ソフトや音楽編集ソフトを使ったことがある人なら、こうしたソフトの価値をイメージできるだろう。1つのチャンネルに記録映像全体を貼り付け、その中で注目する事象を別チャンネルにコード化して記入していく。こ

うしたコードは1つだけでなく10でも20でも独立したチャンネルを増やすことができる。最終的にできあがるのは、さながら、さまざまな楽器の楽譜を重ねたオーケストラのスコアのような時間軸に沿ったコード群である。

異なる事象を時間軸上に精密に位置づけていくと、肉眼観察では見えないさまざまなことが見えてくる。古典的な例ではあるが、異なる事象同士の間で微妙な時系列関係を捉え、ジェスチャーのもつ重要な特徴を捉えた McNeill (1987) の研究を紹介しよう。当時はまだ行動を解析するソフトウェアはなく、音声と同時に記録したビデオ映像（毎秒60フレーム）を用いての発見であった。人はジェスチャーを交えながら相手と話すことが多いが、仕草の開始時刻がその内容を口頭で発話するタイミングよりわずかに先行することを、マクニールは画像と音声の時間軸上での併置から見いだしたのである（吉村、1998より引用）。

近年の行動分析ソフトウェアをもってすれば、時間軸をズームインしてゆけば、時間的にも空間的にもミクロな関連性が容易に視覚化できる。そうすると、逆に全体が見えにくくなるが、その場合には、逆にズームアウトし、時間軸を粗く捉えることでコード間のマクロな関係が可視化できる。

## 6.3 隣接科学における観察法の利用：参与観察とエスノグラフィー

人間行動の観察は、心理学だけでなく文化人類学や社会学などでも盛んに行われてきた。それらの分野では、「参与観察（参加観察法）」や「エスノグラフィー」と呼ばれる観察法が用いられている。

個人や集団を観察する場合、科学的見地に立てば、観察者である自分の存在が観察対象の意識や行動に影響を与えないことが望ましい。しかし、観察したいことが対象者の自発的行動では捉えにくいなら、行動の発現を促すことも必要である。より積極的に、状況設定をしかけたりして反応を引き出すことも有効となる。文化人類学などにお

ける「参与観察」は、対象との関わりや働きかけを通して、より豊かなデータ収集を目指す観察法である。三隅・阿部(1974)は、心理学における参加観察法の定義の中で、参加はするが観察対象に対してできる限り影響を与えないことをたてまえとし、逆に対象に対して積極的に働きかけその変化の過程を観察しようとする方式を「アクション・リサーチ」として、文化人類学における参与観察を参加観察と区別することを提案した。用語上の問題は残るものの、観察対象者に対して働きかけを行う参与観察法がもつ長所と問題点を理解した上での提案と言えよう。

「エスノグラフィー」も、同じく文化人類学や社会学で用いられる用語である。この言葉にはとらえにくい曖昧さもあるが、心理学において質的研究法が広がりつつある現在、質的データ生成の中心を担うのが観察法であり、中でもエスノグラフという形をとることが多い。この点を考えると、ここでエスノグラフィーについて検討しておくことは有益である。

社会学者の佐藤(2006a)は、「エスノグラフィー」というのは実に曖昧で不思議な言葉だと評し、すくなくとも3つの意味があるという。第1に、これを「民族誌」あるいは「調査モノグラフ」と訳せば、フィールドワークの成果をまとめた報告書となり、第2に「民族誌的アプローチ」と訳せば、フィールドワークという調査の方法あるいはその調査のプロセスそのものを指すことになる。そして第3は、「民族誌学」という意味ともなる。佐藤によれば、現在は1か2の意味で用いることがほとんどで、これら2つの意味は重なっているという。

このような理解を受けて箕浦(1999)は、この用語を第1の意味で使い、次のように主張する。フィールドワークが終了した時点で、少なくとも数冊のフィールドノート、数十本のテープ、写真、フィールドで集めた資料などがうずたかく積み上げられているのが普通である。エスノグラフィーとは、それらにできるだけ忠実に書くものと思いつている人が多いが、そのような記録を

読まされる読者は災難である。エスノグラフィーは、自分が理解していることを読者に伝える作業で、自分のためのフィールドワークの記録でないことを肝に銘じるべきで、フィールドノートとエスノグラフとは異なるものだとする。

以上のことを理解した上で、参与観察とエスノグラフィーに関する興味深い心理学研究(村本, 1996, 2006)を紹介し、本章を終えたい。

(村本は)東京都内の公園での早朝ラジオ体操に集まる人々を対象としてフィールドワークを行った。毎朝6時30分からの10分間、この公園では20~30人の人びとが集まって、一緒に体操をする。その活動は、悪天候の日を除けば一日の休みもなく、また、いかなる公式なアナウンスや勧誘があったわけでもない。それは、いつの間にか、自然に作られた集合体であった。参与観察とインタビューを通じて、人びとの行動・相互作用・共有された規範などの性質を1つひとつ抽出した結果、この公園の集合体は、従来の社会心理学において定義されてきた「集団」の概念に合致する特質と、それとはまったく相反する特質とを併せ持っていることが明らかとなった。すなわち、この集合体は、明確な組織をもつ「集団」とも言い切れず、かといって、大勢の人がたまたま居合わせただけというような単なる「集合状態」とも違う、中間的な存在であった。このような曖昧な集合体は、これまでの社会心理学研究ではほとんど現れていなかった。(村本, 2006, p.226-227.)

#### 6.4 文化人類学からのもう1つの贈り物: KJ法

KJ法とは、文化人類学者の川喜田二郎が考案したもので、フィールドワークで集めた情報などをカードを使ってまとめていく方法である。その方法は、チームでブレインストーミングしながらの研究や意志決定を進める上でも有効だとし、のちに川喜田(1967)は『発想法』と題する書籍にまとめた。心理学でも観察データを始め、調査で

の自由記述で集めた多くの記述データを整理し系統立てていく目的で、KJ法は活用されている。

KJ法は、1枚のカードに単位となる記述（とその見出し）を書き、類似性に基づいてカードをグループ化していくという、ポストイット機能に似た素朴な方法のように思える。したがって、「我流」で行っている人も多い。しかし、正式なKJ法では、進め方等にルールがあり、商標登録もなされている。たとえば、「A型図解化」と「B型叙述化」に分けるなどである。まずは「KJ法」を「KJ法もどき」から区別しなければならない。

ところが、「KJ法」自体にもいくつかのバージョンがある。1967年版、1986年版、1997年版（川喜田、1967、1986、1997）の3種類がある。田中（2010）は、三者の違いを一覧表示し、どのバージョンのKJ法かを明示しないと混乱が生じる可能性があるとして指摘する。田中はまた、「KJ法」のもつ「4つのNOT」と題して、「KJ法もどき」がとりがちな方法に対し注意を促している。たとえば、本来の「KJ法」はカテゴリー分けの分類は行わない。「KJ法」では集まったデータだけを用いて似ているもの同士をグループ分けするのであって、あらかじめ設定した分類カテゴリーに各カードを割り振っていくものではない。グループ化の基準を何に置くかは重要である。「KJ法」では創造的発想が重視される。同時代の同じく文化人類学者の梅棹忠夫は、B8判サイズのカードを用いてひとつつながりになった構想を他のカードの一部が見えるようにずらしてホチキス止めする（その形が籠のこざねに似ていることから自ら「こざね」と名づけた）という知の作業を考案したが、その際には「分類するのではなく、論理的につながりがありそうだ、とおもわれる紙きれをまとめていく」（梅棹、1969）。何に基づいてグループ化しているかを自覚して臨むことが重要となる。

## 7. 面接法

心理学で「面接」といえば、心理療法の1つで

あるカウンセリングのことをまず思い浮かべるが、心理学研究法の1つとしての「面接法」は、面接者が対象者と対面し、質疑応答を進めることの中から質的データを収集する方法全体を指す。面接法にはいくつかのスタイルがあり、進め方のみならず、目的も異なる。一般的な分類法は、「構造化インタビュー」「半構造化インタビュー」「非構造化インタビュー」の3分類であるが、このうち「構造化インタビュー」は、あらかじめ設定した項目を手順通りに尋ねていき得られた回答を記録していくもので、設定項目からの逸脱を最小限にとどめつつ量的集計を目指すものである。そこで、あえて面接法とは言わず、「他計式調査」（回答用紙への記入を回答者本人ではなく質問者など他者が行う方式の調査）として調査法に含めてもよいかもしれない。ここでは、あとの半構造化と非構造化インタビューを中心に、質的データの収集を目的とする面接法に限定して、いくつかの類型を捉えていきたい。

### 7.1 デプス・インタビューとグループ・インタビュー

調査法の章で言及した「意識調査」では、ある事柄に対しその人がどのような意見や態度をもっているかを回答してもらうものであった。そうした意見や態度は、一般的なインタビュー（面接法）でも聞き出すことは可能だが、「デプス・インタビュー（深層インタビュー）」とは、意見や態度そのものよりも、そうした意見や態度をもつ理由やいきさつ、さらにはそれらに関連する事情などを捉えることに主眼を置く。意見や態度が複雑で込み入っている場合や、対象者自身が必ずしも意識していない理由がある場合なども、それらを引き出すのに有効な方法である。したがって、この種のインタビューは誰にでもすぐできるものではなく、訓練やセンスを必要とする。心理カウンセラーの行う面接はデプス・インタビューの典型と見なせるが、マーケティング調査など、心理面接とは全く異なる状況でも利用しうるものである。

一方、グループ・インタビューは、マーケティング調査での利用が主で、わが国においても多くの民間調査会社が、さまざまな企業の商品開発や顧客調査などを請け負い、さかんに行っている。その盛況ぶりは、法政大学文学部にまで及んでいる。最近、民間調査会社に委託し、授業改善を目的とした学生からの意見を聞き取るため、わが文学部でもグループ・インタビューが2度にわたって実施された。

グループ・インタビューについては、パブリック・アートの鑑賞をテーマに筆者らが行った研究で検討したので(吉村・伏見・関口, 2008), ここでの見解を要約して説明していきたい。グループ・インタビューは、レヴィン(Kurt Lewin)が第二次世界大戦終了直後に創案したグループ・ダイナミックスの考え方(Lewin, 1945)に依拠して考案されたもので、その根底にグループは個人の集合とは異なるダイナミズムをもつという考え方がある。要するに、数人を一緒にインタビューすれば効率がよいという理由からではなく、1人の発言が他のメンバーを刺激し、単独インタビューでは得られない発言が促進されると考えるのである。レヴィンはゲシュタルト心理学者であり、上述の考えは、「全体(グループ)は部分(メンバー1人1人)の総和とは異なる」というゲシュタルト心理学の根本原理に沿うものである。確かに、インタビューとの一対一のインタビューでは求めにくいさまざまな利点がグループ・インタビューにはある。たとえばHess(1968)は利点として、「相乗効果性」「雪だるま性」「刺激性」「安心感(グループが安らぎをもたらし、率直な反応を促進する)」「自発性(参加者はすべての質問に答えるよう要求されているわけでないので、反応はより自発的で純粹である)」をあげている。

しかし他方、人数が複数であるから量的評価も行いやすいと考えてしまいがちだが、そう考えてはならない。他のメンバーの発言に刺激されて表明する意見は、独立した意見としてはカウントできない。そもそも、グループ・インタビューは量

的評価を目的にするものではなく、あくまで発言の広がりや深まりを目指すものである。

また、上で紹介したHessのグループ・インタビューに対する評価は、あくまで欧米文化の中での評価であり、日本人の場合、初対面の数名が「安心感」をもって発言できるようになるにはかなりの時間を要し、通常行われる2時間程度のグループ・インタビューでは、ファシリテータ(司会者)によほどの能力がないと、各質問に(半ば強制的に)順番に答えていってもらおうという「自発性」のないセッションとなりかねない。「刺激性」を生かすには、セッション終了後のお酒付きの懇親会が必要と言われるゆえんである。マーケティング・リサーチの分野でグループ・インタビューを含めさまざまなデータ収集を実践してきた中山(2005)は、わが国でのグループ・インタビューを手厳しく批判する。批判の理由はこうである。

グルインって約2時間くらいかけてやる。ここに6人とか8人くらいを集めるわけですよ。ということは、1人平均のタイムシェア、どう計算しても、せいぜい20分、自己紹介のための時間もあります。司会者のしゃべる時間も必要です。いいとこ1人15分しかとれません。

しかも普段行ったこともないお堅い「会議室」で、時には「別室でモニタ(監視)」するための隠しカメラまでついています。もう、バレバレ!

たった15分で初対面の司会者の前で、初めて会う人たちに囲まれて、カメラを気にしながら、ホンネが語れるものでしょうか? 私には絶対にできません。

でもそれを期待してやっているのがグループインタビューってやつです。

そんな状態ですから、声の大きな人、正論を吐く人、ちょっと専門っぽい意見をひけらかす人……の意見に影響されやすい。(p.61-62)

ここでは、グループ・インタビューの問題点を

中心に説明したが、吉村ら(2008)では、わが国で活用するにあたっての留意点を指摘し、積極的利用に向けての提案を行っているので、関心のある人は参照してもらいたい。

両タイプのインタビューとも、マーケティング業界での使用が盛んである。それは、意見分布の量的裏づけを得るためだけでなく、消費者に対する理解を深め、自らの商品や広告の特徴を捉え、よりよいものづくりのための発想を得る目的で行われている。

## 7.2 認知インタビュー

「インタビュー」という行為は、極めて広い状況において行われている。心理面接やオーラル・ヒストリーの作成、調査票を用いるだけでは得られないデプス・インタビューなどは、心理学において行われている「インタビュー」である。加えて、グループ・インタビューなど、マーケティング業界中心に利用されているものもある。さらには、入社試験での面接もインタビュー、マスコミ記者による取材もインタビューである。このような広がりの中、「認知インタビュー」は認知心理学が開発し、犯罪心理の分野で利用されている領域限定的な方法である。そういえば、警察による尋問や聞き込みも英語ではインタビューである。

「認知インタビュー」については、吉村(1998)で解説したので、それを要約しながら説明していきたい。尋問や聞き取りの証言の質の向上を目指してこの方法を創案したのはフィッシャーとガイゼルマンで、彼らは1987年からこの手法の開発に取り組み、Fischer and Geiselman(1992)で系統的解説を行った(この書籍は2012年に邦訳されている)。認知心理学における記憶研究の成果を生かし、インタビューを受ける人の記憶向上を目指す工夫が盛り込まれ、インタビューする人と受ける人の意思疎通の向上を図る社会的コミュニケーション技法も取り入れて行われるものである。

記憶検索を向上させる工夫の説明に続き、社会的コミュニケーション技法についても吉村(1998)

の解説から引用する。まずは、記憶検索を向上させるための4つの方略からである。

- (1) 事件が起こった場面の環境のイメージや印象を形成するため、目撃した事実の文脈を心の中で再構成するよう教示する。またそのとき、どのような感情や考え方を抱いているかを思い出すように求める。すなわち、“状況依存効果”を利用する方略である。
- (2) たとえ重要でないと思っても、あるいは断片的にしか思い出せないことであっても、編集をいっさい施さず思い出すまますべてを報告するように求める。
- (3) 一連の事象をさまざまな時間順序で、また出発点を変えて報告するように求める。
- (4) 異なった場所から、あたかも別の人の視点に立って事象を思い出す作業を求める。

(中略)

記憶理論に基づくこれらの方略に加えて、社会的コミュニケーション技法も重視されている。具体的にはインタビューのペースの指定や再生作業の組み立てを証言者の側に委ねたり、オープン・クエッションを用いたり、あるいは証人の発言を遮ったりしない、などの配慮である。(吉村, 1998より引用, p.122)

認知インタビュー手法を用いたKoehnken, Schimmossek, Aschermann, and Hoefler(1995)の研究では、通常のインタビュー法に比べ、目撃者の報告数が2割以上増加した。ただし、「正しい報告」のみならず、「誤った報告」の数も同じように増加した。彼らはこの結果を、「正答率を犠牲にすることなく目撃者から引き出す情報を本質的に増加させる手段」として、積極的に評価した。誤った情報に対しては、のちの捜査段階での情報照合によって、誤りであることを検出するチャンスがあると見なせるからである。

## 7.3 断片化されることへの抵抗

インタビューを書き起こしたものをデータ化す

際には、ひとまとまりの記述ごとに区切り、記述内容をコード化してデータとすることが多い。それは、発話者を「没個性化」し、語られたことの内容だけをデータとして採用する行為となる。加えて、前後の文脈の中での発言であることを無視する「脱文脈化」の行為ともなる。科学とは分析することを旨とするもので、複雑で多面的な内容の会話記録を単純化し、データとして扱いやすくすることは、科学としての常套的手段に適っている。

しかしながら、没個性化や脱文脈化は、インタビュー内容の重要な部分を文脈からはぎ取る危険を伴う。秋田・藤江(2019)は、質的研究法を解説したFlick(2007)の指摘を引いて、次のように述べている。「会話分析のように、特定の形のトークの組織化のような、形式的な構造の解明が主たる課題である場合や、発話や語りの内部構造や全体性をもったまとまりに関心がある場合は、コード化を用いることは難しい」(秋田・藤江, 2019, p.27)。コード化と文脈の保持とはトレードオフの関係にあり、どちらかを犠牲にせざるをえないものなのだろうか。

質的研究の進め方について、近年提案がなされているさまざまな方法の中には、コード化によりデータを断片化されたままで終わらせるのではなく、再文脈化することまで視野に入れた進め方も提案されている。たとえば、大谷(2019)の「SCAT」(Steps for Coding and Theorization)がある。この方法は、語られたり記録されたりした文章をテキスト・データとし、それを用いた理論化までのプロセスを「可視化」することを目指している(最初の提案は大谷, 2008において行われた)。一般に、質的研究への批判は、テキスト・データからの理論化プロセスがブラックボックスの分析者の主観に基づいて行われる点にある。それに対し、SCATでは、テキストを〈1〉テキスト中の注目すべき語句の抽出、〈2〉テキスト中の語句の言い換え、〈3〉それを説明するような概念の検出、〈4〉(前後や全体の文脈を考慮して)テーマ・構成概念の考案を経て、「ストーリーラ

イン」を作りあげるという各プロセスの出力を書き残すことによって可視化を図る。プロセスの可視化は、科学としての要件を満たすもので、ストーリーライン作成が「再文脈化」プロセスを担う。大谷(2019, p.317)は、自らのSCAT以外にも質的研究で再文脈化の手続きを用いているものがあるとし、KJ法の「叙述化」をあげている。これまでは量的研究の研究者と質的研究の研究者はお互いに自分の長所を賞賛し、相手の欠点を批判し合う関係に終始してきたが、分析プロセスの明示化は科学としての要件を満たす質的研究の側からの努力と言える。

## 8. すべての研究法を振り返って

### 8.1 サンプルサイズはどう決めるか

実験法・調査法・検査法・観察法・面接法が出そろったところで、いずれの研究法においても重要な検討事項となるサンプルサイズについて考えたい。Dunn(1964)は、「必要な標本の大きさを見積もるということは、きわめて大切なことである。というのも、すべての問題に適用できる標本の大きさとして、これが〈最良〉というものは存在しないからである」と重要さだけでなく難しさも指摘している。これは、サンプルサイズについて考察した村井(2006)からの引用だが、村井自身も、「それぞれの状況に応じて個別に判断しなければならないので、即答できない」と慎重である。それでも、他の研究者の見解として、具体的な数値を紹介している。Cohen(1990)は、学生時代に「群比較するとき適切なサンプルサイズは各群30である」と習ったそうで、また田中(1996)には「人間の心理現象についていえば、一群20~30人くらいで安定するように思われます(個人的な経験則にすぎません)」と書かれているようである。実際に行われている研究よりも多めかもしれないが、3.3で指摘した「大数の法則」に少しでも近づくという意味からは、多めの方がよいかもしれない。

しかしながら、サンプルサイズが大きいこと

は、よいことばかりと言えない。実験法の場合、フィッシャーの実験計画法のところでも説明したように、そもそもその方法は、少人数の参加者から有意な結果を導き出すために考案されたものであった。大きすぎるサンプルサイズがもたらす問題点は、最近クローズアップされている「効果量 (effect size)」の観点から明白である。効果量の必要性が指摘され始めた頃の検定手順は、以下のようであった。まず、検定力 (通常は .80) と有意水準 (通常は .05) を決め、各条件に常識的数値の参加者数 (実験なら 20 名程度) を割り当ててデータ収集を行う。そして、有意差検定へと進み、有意差が認められたとする。ところが、最後に行う効果量評価において満足できる値が得られなければ、せっかくの有意差がしぼんでしまう。実は、ここに登場する 4 つのパラメータ、「検定力」「有意水準  $\alpha$ 」「サンプルサイズ」「効果量」には閉じた関係があり、3 つの値が定まれば残る 1 つの値も決まる (Borenstein, Rothstein, Cohen, and SPSS Inc, 2001)。そこで、上のようにデータ収集を行ってしまってから効果量の評価を行うのではなく、あらかじめ満足できる効果量を定めておき、そこから必要なサンプルサイズを導き出してデータ収集を行うのが効率的である。村井 (2006) には、効果量から必要なサンプルサイズを算出してくれるソフトウェアとして Sample Power などが紹介されている。また、「t 検定で、効果の大きさが A、検定力が B、有意水準  $\alpha$  が C のときの、サンプルサイズは？」などパラメータ間の関係を読み取り表の形で示してくれるソフトウェアとして Cohen (1988) のものなどが紹介されている。

実験法では、サンプルサイズが各条件 50 名や 100 名などということはまれで、ほとんどの場合 20 名程度以下なので、有意差が得られた時点で効果量もある程度高いと期待できる。それに対し、調査法の場合は 100 人単位でのデータ収集が普通なので、効果量問題は深刻である。とは言え、調査法でサンプルサイズが大きいことには、別に理由がある。調査の目的が、有意差検定では

なく、母集団の性質をより狭い範囲で推定することであることの方が一般的だからである。精度の高い区間推定を行うには、サンプルサイズは大きいほどよい。また、因子分析により安定した因子を抽出するためにも、項目数 (列) の数倍の人数 (行) の確保が必要で、それにより変数間の相関関係の評価の信頼性が高まるため、人数は多いほどよい。

しかし、調査法で得られたデータを用いて有意差検定しようとする、上述したように事情が違ってくる。効果量の問題が、直接的に関わってくるからである。50 人、100 人ものサンプルサイズがあれば、わずかな値の違いでも平均値や相関に有意差が生じる。このような場合には、効果量を必ず算出し、結果評価の必須指標として表示すべきである。

以上の通り、実験や調査を行ったり、検査法を作成したりする場合、すなわちデータを量的に評価する際には、サンプルサイズが重要な意味をもつ。それでは、面接法や観察法で質的データを集めるときには、サンプルサイズ問題は生じないのであろうか。結論から言えば、全く別の意味で、サンプルサイズをどの程度にするべきかが問題となる。たとえば、あるテーマについて、それに関係する人たちから意見や考え方を聞き出そうとするインタビューの場合、1 人から聞いただけでは、意見・情報の偏りや、重要な情報の収集漏れがあるに違いない。それならば、2 人ではどうか。この問いへの答えを得ることは理論的にはできない。しかし、メドがまったくないわけではない。製品の使い勝手を評価するユーザビリティ分野でのことではあるが、ユーザビリティ工学の開拓者ヤコブ・ニールセンは、次のような公式を提案している (Nielsen and Landauer, 1993)。

$$\text{Usability Problems Found}(i) = N(1 - (1 - \lambda)^i)$$

ここで  $i$  はテストユーザーの数、 $N$  はユーザビリティ問題点の総数、 $\lambda$  は 1 人のテストユーザーが問題点を発見する可能性である。Nielsen and

Landauer (1993) の研究では、実測値として、N の平均値は 41 で、1 人のユーザーが問題点を発見する可能性  $\lambda$  の平均値は 31% であった (Nielsen, 1994 p. 135 より引用)。これを受けて、樽本 (2012) は次のように解説する。テストに費やすコストとそれによって得られる成果のバランスを考えると、問題の総数 N の大半 (約 85%) を発見するには、5 人をテストすればよい (Nielsen and Landauer (1993) に従って、 $\lambda=0.31$  とする)。すなわち、上の式にこれらを代入すると、Usability Problems Found(5)=0.8436N となる。5 人をテストすることで、問題点の総数の約 85% が発見できるのである。この公式に用いられた定数は、あくまでユーザビリティ分野での問題発見についての提案で、問題総数が有限個で 1 人による発見率が 31% と想定するユーザビリティ分野に固有の式である。そのため、5 人という数値はインタビューによる新しいことがらの指摘全般には当てはまらない。しかし、インタビュー領域においてもニールセンの考え方をもとに、必要なインタビュー対象者数を経験的に提案できる日が来るかもしれない。

## 8.2 2つの QDA

インタビューデータを分析する際、筆者は MAXQDA という商品名のソフトウェアを使っている。これは QDA ソフトと総称される質的データを分析するためのソフトウェアの 1 つである。ここで言う QDA とは、Qualitative Data Analysis のことで、このソフトの日本での利用を積極的に推奨している社会学者の佐藤 (2006b) も、この意味で QDA を使用している。

ところが、同じ QDA という略語が、官能検査の世界ではむしろ逆の意味で使われている。神宮・土田 (2008) には「定量的記述的試験 (QDA: Quantitative Descriptive Analysis) 法」とある。しかも、この用語使用は、日本工業規格の JIS Z9080 で規定されている。Q だけを見ると、Qualitative (定性的) と Quantitative (定量的) で MAXQDA の場合と全く意味が逆である。さ

らに、D の使われ方も違っており、官能検査での D は Descriptive (記述的) で、その内容はむしろ定性的なものである。話が込み入ってきたので、官能検査での Descriptive にだけ注目し、それがどのような手続きを意味しているかを神宮・土田 (2008) に基づき説明しよう。

人がもの (サンプル) と接したとき、そのものもっている多様な属性を意識し評価がなされる。どのような属性を意識しているかを明らかにするための方法が、「記述的試験法」である。一般には、5 人以上の専門家 (そのサンプルをよく知っている人たち) に、そのサンプルを表現するにはどのような言葉が適切かを表現してもらう。得られた言葉の中で似たもの同士を集めてカテゴリー化し、使用頻度を考慮して整理していき、評価語を決定する。(p. 41)

ここまでが Descriptive に関わる作業で、ここから先は Quantitative (定量的) 作業となる。それは、決定された評価語を用いて評定尺度による定量的データ処理を進める作業である。

注意を促したいことは、QDA が心理学に関わる隣接領域において、まったく異なる意味で用いられている点である。

心理学研究において、量的研究と質的研究はことあるごとに対立してきた。科学の名のもと、量的研究のみを科学とする風潮もお強いが、本論で扱った観察法や面接法などは質的研究に支えられている部分が多い。筆者がライフワークとしてきた「逆さめがねの世界への知覚順応過程」を捉える研究などは一人の着用者に対する分析、すなわち事例研究の典型で、質的研究としての性質が強い。そこで、事例研究法、引いては質的研究のもつ特徴を正負両面から捉える必要性に直面してきた。かなり以前の整理ではあるが、吉村 (1989, 1996) に掲げたリストを、以下に示しておきたい。

事例研究のネガティブな面

- (1) 小標本 (サンプルの代表性が保証されな

い)

- (2) 定性的データ中心（特に主観性の強いデータ）
- (3) 事後研究（独立・従属変数の未決定性）
- (4) 被験者主導型研究（条件統制，操作がなされていない）

#### 事例研究のポジティブな面

- (1) 個性的なものは法則的なものの残りが残らない
- (2) 個には普遍的なものが現れている
- (3) 心理機能の理解は生活環境から切り離さずに行う方がよい
- (4) 希少な異常例の検討は正常状態の理解を助ける

最近でも、「量的研究 vs. 質的研究という二項対立を越えて」という枠組みで、村本（2006）が近年の論者の主張を紹介している。もちろん、質的研究の側に重心が移ってしまうと、科学としての健全さを失ってしまう。しかし、質的研究が方法論的役割の一翼を担い続けることも、心の科学としての健全さの維持にとっては必要なことである。

### 8.3 終わりにかえて：最近の卒業論文

最近の心理学科の学生が取り組む卒業論文テーマには、取り上げた現象そのものの心理的側面を捉えることを目的にしているものが多い。たとえば、「SNSを利用する大学生の自己承認欲求」などである。おそらく彼らは、その現象そのものに関心や疑問をもっていて、その現象が解明できればそれ以上のことは望まないであろう。ゆえに、抽象的な状況設定はしない。そうした具体的な状況において得られた見解に、彼らは達成感を感じるのであろう。

本来、研究で取り上げる現象は、その現象の背景にある一般的な心的事象からのサンプルであり、その現象を通して明らかにできたことから一般性ある見解を導くことが研究の価値であると、筆者は考えてきた（たとえば、4.3の村本の研究参照）。「A理論を検証するために、具体的にA'

という実験状況を設定してデータ収集を行う」というのが心理学の通常のパラダイムで、たとえ卒業論文といえども、そのパラダイムに則った研究が望ましいと考えてきた。しかし、そうした枠組みのもとで設定されるサンプル状況は、多くの場合、現実生活では体験しない不自然で、たいていはおもしろくない状況である。そのような状況で生じる現象そのものに、学生たちは興味・関心を示さない。SNSのような日常生活で体験する関心事とは無縁だからである。

4.3のサンプリングの節で示したように、母集団から無作為抽出されたサンプルを用いることは、事実上、不可能に近い。このことは、得られた知見を一般化する際にも当てはまる。設定した人工の実験条件には、背景となる母集団が現実世界の中にあるとは限らない。少なくとも、卒業論文に取り組む学生に、それを想定するよう強いることはできない。教員の側からの歩み寄りが必要である。

しかし、得られた知見の一般化には、無作為抽出とは異なる価値がある。適用できる範囲やことからは、ある程度まで論理的に想定できる。そうした見通しをもって卒業論文に取り組む知性を、学生の側に求めてもよいはずである。できあがった卒業論文や修論を読んでいると、「考察」の内容が「結果」で記述したことの繰り返しに終始しているものに会うことがある。結果から飛躍しない慎重な書きぶりや褒めるべきかもしれないが、得られた知見のもつ適用範囲に対する見通しのなさの反映であることも否めない。

エビデンスを重視する現代社会では、卒業した学生たちはさまざまな局面で主張の根拠となるデータを示す必要に迫られることになるだろう。資料重視の学問を学んだ人たちなら、証拠となる文献資料をあの手この手で収集しレポートすることだろう。しかし、データを生み出すトレーニングを受け続けた心理学科の卒業生は、自分たちで証拠となるデータを生み出す手段を培ってきたはずである。そのようなとき、あの手この手の方法があることを、本論を通して習得し、実行に向け

て努力してもらえればと願いたい。

### 引用文献

- Borenstein, M., Rothstein, H., Cohen, J., and SPSS Inc. (2001) *Sample Power 2.0*. Chicago, IL: SPSS Inc.
- Cohen, J. (1988) *Statistical power analysis for the behavioral sciences 2nd ed.* Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cohen, J. (1990) Things I have learned (so far). *American Psychologist*, 45, 1304-1312.
- Dunn, O. J. (1964) *Basic statistics: A primer for the biomedical sciences*. New York: John Wiley. (ダン, O. J. 田中真理子・日野寛三 (訳) 1986 医歯系・生物系の統計学入門 大竹出版)
- Fisher, R. P. and Geiselman, R. E. (1992) *Memory-enhancing techniques for investigative interviewing: The cognitive interview*. Springfield, IL: Thomas. (宮田洋 (監訳) 2012 認知面接: 目撃者の記憶想起を促す心理学的テクニック 関西学院大学出版会)
- Flick, U. (2007) *Designing qualitative research*. London: Sage. (フリック, U. 鈴木聡志 (訳) (2016) 質的研究のデザイン 新曜社)
- Hess, J. M. (1968) Group interviewing. In R. L. King (ed.) *New science of planning*. Chicago: American Marketing Association. pp. 51-84.
- 神宮英夫・土田昌司 (2008) わかる・使える多変量解析 ナカニシヤ出版
- 川喜田二郎 (1967) 発想法——創造性開発のために 中央公論社
- 川喜田二郎 (1986) KJ法——混沌をして語らしめる 中央公論社
- 川喜田二郎 (1997) KJ法入門コーステキスト4.0 KJ法本部・川喜田二郎研究所
- Koehnken, G., Schimmossek, E., Aschermann, E., and Hoefler, E. (1995) The cognitive interview and the assessment of the credibility of adults' statements. *Journal of Applied Psychology*, 80, 671-684.
- Lewin, K. (1945) The research center for group dynamics at Massachusetts Institute of Technology. *Sociometry*, 8, 126-136.
- McNeill, D. (1987) *Psycholinguistics: A new approach*. New York: Harper & Row. (鹿取廣人他 (訳) (1990) 言語心理学——「ことばと心」への新しいアプローチ サイエンス社)
- 箕浦康子 (1999) エスノグラフィーの作成 箕浦康子 (編著) フィールドワークの技法と実際 ミネルヴァ書房 pp. 71-86.
- 三隅二不二・阿部年晴 (1974) 参加観察法の定義と歴史 続有恒・苧阪良二 (編) 心理学研究法第10巻 観察法 東京大学出版会 pp. 139-147.
- 村井潤一郎 (2006) サンプルサイズに関する一考察 吉田寿夫 (編) 心理学研究法の新しい形 誠信書房 pp. 114-141.
- 村本由紀子 (1996) 集団と集合状態の曖昧な境界——早朝の公園で見出される多様なアイデンティティ 社会心理学研究, 12, 113-124.
- 村本由紀子 (2006) 心と社会を研究する方法 吉田寿夫 (編) 心理学研究法の新しい形 誠信書房 pp. 221-243.
- 中山マコト (2005) たった3日で売れ出すキキダス・マーケティング 日本能率協会マネジメントセンター
- Nielsen, J. (1994) *Usability engineering*. New York: Academic Press. (篠原稔和 (監訳) (2002) ユーザビリティエンジニアリング原論 東京電機大学出版局)
- Nielsen, J. and Landauer, T. K. (1993) A mathematical model of the finding usability problems. Proceedings of the INTERCHI '93 Conference on Human Factors in computing systems (Amsterdam, The Netherlands, 24-29 April 9, 206-213.
- 大谷 尚 (2008) 4ステップコーディングによる質的データ分析手法 SCAT の提案——着手しやすく小規模データにも適用可能な理論化の手続き——名古屋大学大学院教育発達科学研究科 (教育科学), 54(2), 27-44.
- 大谷 尚 (2019) 質的研究の考え方——研究方法論から SCAT による分析まで——名古屋大学出版会
- 佐藤郁哉 (2006a) フィールドワーク増訂版 書を持って街へ出よう 新曜社
- 佐藤郁哉 (2006b) 定性データ分析入門 QDA ソフトウェア・マニュアル 新曜社
- 田中富士夫 (1997) MMPI 新日本語版の構想と標準化 MMPI 新日本語版研究会 (編) MMPI 新日本語版の標準化研究 三京房 pp. 1-21.
- 田中博晃 (2010) より良い外国語教育研究のための方法 KJ法入門: 心的データ分析法として KJ法を使う前に 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, 17-29.
- 田中 敏 (1996) 実践心理データ解析——問題の発想・データ処理・論文の作成 新曜社
- 樽本徹也 (2012) アジャイル・ユーザビリティ オーム社
- 梅棹忠夫 (1969) 知的生産の技術 岩波書店

吉村浩一 (1989) 心理学における事例研究 心理学評論, 32, 177-196.

吉村浩一 (編) (1996) 特殊事例がひらく心の世界 ナカニシヤ出版

吉村浩一 (1998) 心のことば——心理学の言語・会話データ—— 培風館

吉村浩一・伏見清香・関口洋美 (2008) グループ・イ

ンタビューの可能性：パブリック・アート鑑賞システムの評価をめぐって 法政大学文学部紀要, 56, 49-60.

吉村浩一・高岩亜輝子 (2006) もう1つの病院臨床——神経心理学と基礎心理学の有機的連関を目指して—— 法政大学文学部紀要, 53, 17-29.

How to Teach the Class Subject *Research Methods in Psychology*:  
Part Two

Hirokazu YOSHIMURA

**Abstract**

Continued from our last issue which dealt with experiments and surveys, the author points out some important problems of the three methods, tests, observations, and interviews, and shows his own view on each of them in this issue. In the test method chapter, he explains how tests are different from surveys, the difficulty to standardize a test, and the importance to construct test battery. In the observation chapter, problems occurred when the voice and video recording system were introduced, the relationship of participation observation and ethnography, and the characteristics of KJ Method (Kawakita, 1967, 1986, 1997) are explained. In the interview chapter, he explains the characteristics of the depth interview and the group interview, and explains the procedure of cognitive interviews which are used for crime investigations. In the final chapter, he suggests how to decide the size of a sample and points out that the term QDA has two different meanings (Qualitative Data Analysis and Quantitative Descriptive Analysis) in psychology. Lastly, the author comments on students' attitudes when they work on their graduation thesis.

**Keywords** : method of psychological study, test method, interview method, observation method