

質問紙法 —その使い方—

越 河 六 郎*

Questionnaire: Its Usage

Rokuro KOSUGO*

This thesis shows a study on a questionnaire method that is to be used how to evaluate the level of fatigue and stress in an occasion of industrial workplace context. This evaluation method can perform in following two purposes; one is a personality inventory method where individual personality would be identified and another is holistic investigation on particular aspects of one's own circumstances, such as his/her workload etc. The subjective assessment of fatigue, for example, can work for the latter purpose. It should analyze not only the results from the checklist answers but also the relationship between those checklist answers and the working conditions in the workplaces. With this respect, it seems to be that the questionnaire method has been used with lack of this point of view.

ある職場から依頼されて、質問紙法による「調査」を実施しようとする、調査対象者に予定された職員の人たちから、また調査ですか、いろんな調査が続いて閉口しています、というクレームを聞くことがある。そうかと思う、すっかり調査ズレしてしまったのか、適当に答えていればいいんですといった態度の人もいる。

労務管理・健康管理の分野に限ってみただけでも、「職務満足度調査」、「ストレス調査」、「健康調査」など標準尺度の形式を整えた市販ものから、一般のアンケート調査の類まで含めると、「尺度」「調査表」の数は相当のものになる。それらが、現場に必要なあつての調査ならまだしも、「研究者」からの依頼、理論検証のための資料を得るという目的でなされる例も少なくない。調査される人たちは、そのような場合、単なる「対象者」でしかないのか。

本稿では、いわゆる「社会調査」批判を試みようとしているわけではない。あまりにも安易過ぎる使われ方となっている質問紙調査法について、40年以上にわたる「調査」の経験から、反省を含めてその問題となるところを、ほんのわずかだが取り上げてみようと思う。

質問紙法（評定尺度）の用い方は、対象となった人の「個人的特性」を検査するために使う場合と、個々人の「周囲の事情」を調べるため、の2つに大別することができる。筆者は、労働科学研究所をベースに、労働負担評価に関連する調査事項の一つである「疲労感調査」を手がけてきた。どちらかといえば、後者の目的による使い方となる。評定尺度の作成経過、フィールド調査でのエピソードなども絡めて、考えを述べることにする。

* 松蔭女子大学
Shoin University

1. 採用時などの性格検査

民間の病院で、わが国有数の総合病院の経営を見事に務められた方との酒席での話である。“採用時の適性検査の一つとして、ある性格検査を実施したが、かなり有名なテストとされているのに、あれは使いものにならないね。たとえば、仮に「協調性」などを考えて、とにかく成績の良いものを採用した。しかし、各部署に配置した後を見てみると、どうもあまり上手くいっていない。そうかと思うと、検査の成績では、これはどうかと危ぶんでいた人が反対によく勤務している。マニュアルの効能書きのとおりにはいきませんな。どうしたものかね”，というわけである。この質問に対して、その場では、それは使い方に難点がありますね、と答えてはおいしたが、他愛も無い酒飲み話とは異なることに気づいた。

使い方が悪いとした理由は、マニュアルに書いてあるとおりに「使った」ことを第一に挙げることができる。もちろん、この種の検査ではマニュアルどおりに使うことは当然であるが、出来合いのものとして、つまり、使いこなせるだけの力量が無いままに、書いてあることをなぞっているだけの使い方だということである。

そのマニュアルにも疑問がある。応答結果（検査成績）から、職務適性とか社会的適応性などが「判定」されている点である。検査法を標準化するとき、サンプル数の問題もあるが、妥当性、信頼性の検証を行ったとはいえ、せいぜい性別とか、年齢、業種、職種等といった「名目的な」区分を手がかりとした検証レベルで、どうして具体的な「現実の」職場に適応性があるとか無いとかを指示できるのか、曖昧な点がある。検査法を用いる立場の人は、逆説的かもしれないが、少なくともこれらの曖昧さを了解して使用すべきだと思う。

採用に当たっての「適性検査」となると、いかにも検査実施の目的がはっきりしているようだが、はたしてそうだろうか。適性検査の本来の意義は、職場配置や職務訓練のための基礎的情報を得るところにある。採用の可否を決めるための添え物的なものではないことを認識したい。

上にあげた「性格検査」には、もう一つの話がある。地方公務員の試験を受け、一次試験、二次試験に合格し、採用時の健康診断・面接の折に、「補足

的なものです」といわれて、この「性格検査」に回答を求められた青年がいる。その結果、「あなたは性格的に問題があるので採用できません」として不採用になったという。彼は、次の年も同じ公務員試験を受験し、同じく性格検査も受けた。今度は採用となり、その後20年以上とくに問題なく勤務を続けている。2回目の性格検査では、どのようにして応答したかを尋ねたら、各質問について、社会的に望ましいと思われる答えを選択したという。なんのことはない、検査法の不備を突いたわけである。そうしたら、1回目はどうしたのか。そのときは、質問（読み上げ）の意味がよくわからないものがあったので、答えもそうなったのかもしれませんが、ということであった。この例は主観評定法の難しさを示すものであろう。彼の1年間を補償するものはなにもなかった。

筆者は5年ほど前に自動車運転教習所で「運転適性検査」を受験した。その検査には一種の性格検査が含まれていたが、幸いなことにその検査を受けるのははじめてであった。結果は、各人のファイルの中に評価表の形で報告されることになっていたが、自分のファイルにはそれがなかった。「カード」が入っていて、特定の指導員（適性検査法の研修を終えている担当者）の所まで出頭しなさいとある。面談での指導が必要だったわけだ。「あなたは運転席に座ると人間が変わる性格の持ち主で乱暴な運転になるから、免許がとれてもなるべく自分で運転することは避けられたらよいでしょう。検査の結果がそのように出ているのです」と宣告された。こちらは、そうですかと引き下がるほかにはなかった。できるだけ素直に答えたつもりだったが、応答に片寄りがあったようだ。このテストは初めてでも、後で述べる「疲労感の評定尺度」などを作成した経験があるものだから、どうしても質問の“ワーディング”などが気になって、yes, no, ?, の選択のうち、?印のところにチェックが集まったためではないかと自己弁護している次第である。笑い話で済ませてしまえばそれまでだが、一つの検査結果だけで「決め付けてしまう」扱い方には抵抗を感じた。

この運転適性検査受験の経過を、概して同様の検査法を開発研究している人に、応用心理学会の大会の折、立ち話して訴えてみたことがある。そうしたら、その人は冗談と受け止めたらしく、越河さんに

ぴったりの結果ではないですか。その検査法使えますね、と冷やかされてしまった。マニュアルがあつて市販されている「スケール」を使用するときは、「冗談」では済まない部分のあることを理解してもらいたかったのだが。

上に紹介した公務員試験失敗の例などはまさにこのことである。蛇足になるが、小生の走行距離は2万キロ近くになる。今のところ至っておとなしく乱暴な運転はしていない、つもりだ。また、教習所での「検査結果」は、面談指導には用いられたが、走行技能訓練時に活かされていたのかどうか、教習ノートには記していない。

2. 質問紙作成にあたって

質問紙（評定尺度）の作成段階でも問題は多い。質問紙法は主観を問うわけであるから、まずもって、質問の仕方、「言葉づかい」には十分な配慮が必要である。答える側からすると質問の意味がわからなければ答えようがないからである。したがって、応答が得られたとしても、それらが何を表現しているのか、応答の数値化にあたっては、「もとのところ」をしっかりと吟味してかかるべきであろう。

東京都の授産所で仕事をしている60歳以上の人たちの作業負担を調べさせてもらったことがある。当時よく使われていた疲労感評定尺度の一つ『自覚的症状調査表』（1967年に、産業衛生学会・産業疲労研究会によって改定され、『自覚症状しらべ』となる）を適用した。被調査者は高齢ということもあって、調査表の質問が読めない。自分で読んで回答する方式の調査表なのだが、こちらで一つひとつ読み上げてやって、口頭で「症状の有無」を答えてもらった。やってみたら、実に読みにくいことに気づいた。調査は、勤務前とか勤務終了時など、「その時点での」症状を尋ねているのだが、なんと「時制」（Tense）が合っていないのである。特に精神的症状としてまとめられていた項目群には苦労した。早速、おばあちゃん（対象者）の一人からクレームがついた。そのいくつかをあげてみよう。

「考えるのがいやになる」

…「いま、なにも考えていません」

「話をするのがいやになる」

…「いま、話してないでしょう」

「ねむくなる」

…「ねむい、ということ？」

「する事に間違いが多くなる」

…「いま、間違ふようなことしていません」

「どわすれする」

…「いま、度忘れするのですか」

と、いったぐあいである。仕方なく「翻訳して」応答を求めた次第であつた。時制がずれているだけでも、これだけ、答えにくいことを知らされたのである。このおばあさん、「肩がこる」という設問には、抵抗なく「ありません」と答えてくれたのだが、質問一回答が終わったら、すぐ近くにあった「肩揉み機」に座って肩を揉み始めたのである。やはり、肩の凝りがあるのですか、と聞いたら、「こうすると気持ちがいいのよ」という返事だった。

ストレススケールなど、外国で作成されたものを日本語に訳して「標準化」されたものも少なくないが、日本語として意味不明の設問は頂けない。翻訳のまずさだけではないと思うが、「人前に出ると赤面する」といった種類の質問項目も回答しにくい。読み上げ方式の調査では、回答に戸惑っていると、あまり深く考えないで、さっと答えなさいと言われてしまう。「人前」とはどんな人前なんだろうなどと考えていたら怒られる仕組みになっているようである。こんなふうにして「性格」などが特定されてしまうのかと思うと恐ろしい気もする。

検査法の標準化の手続きである妥当性や信頼性の検証は、統計的解析法の進歩で比較的容易になされるようになったことは確かである。「パッケージ」（統計解析ソフト）を当てはめれば、因子構造がわかるというわけである。以前は、一つの相関係数を算出するだけでも一兩日を要したことから考えると隔世の感がある。

そういった「妥当性」、「信頼性」に関することだが、もとは知能検査など「検査法」の具備すべき条件としての検証事項であつた経過がある。いまは、主観を問う調査表「質問紙」でも、妥当性・信頼性が「検証」されていないとスケールとして扱われない状況にあるといってよいであろう。ただし、「検査法（テスト）」と「質問紙法」は、その中味がはっきり異なることを忘れてはならない。さきに、性格検査といったように「検査」という表現をその

まま用いたが、違いをわきまえてかからなければならぬ面がある。どういうことかという、これも当然のことだが、検査法は「できるかできないか」、機能的な側面から個人差を扱う手法であり、質問紙法（評定尺度）は「何らかの事象についてどのように感じているか、思っているか」の主観を問うている。検査法は正解か不正解かははっきりとえられるが、質問紙法では応答を単純に正解、不正解とするわけにはいかない面がある。検査法は、「課題が解けるか」であり、主観応答の場合は各個人のそれぞれの判断基準で、まさに主観を「回答」すればよいのである。いわば、質問紙法は「課題性」が希薄だということになる。したがって、個々人の主観応答を数量化し、多数の人の中での「分布上の位置」を求めることができて、それは見かけ上の位置でしかないことになる。主観応答は個人の内的判断基準に依存するだけに、「応答数や訴え数」といった量だけでは処理できない側面を残している。

妥当性とは、当該テストが「何を測定しているのか」または、「測定しようとしている機能を完全に測定しているか」ということである。その検証法には、理論的検討、他の基準との相関を調べるなどがあるが、下位テスト間の相関をもとにした因子分析的手法はテストの性質を明らかにするという点でもっとも客観的な方法とされている。主観的な側面を問う質問紙でも、妥当性の検証は可能である。この場合は、各質問項目間における「応答の整合性」といった面からの、質問項目の「特性分類」ということになろう。

信頼性となるとどうか。妥当性と同じ意味で扱う場合もあるが、同一テストを同一被検者に一定の期間を置いて、2回以上実施したときの成績の「変動量」をあらわす、これが一般的である。質問紙法ではどうなるか。主観を問う。何らかの事象についての主観である。そうすると、「何らかの事象」が問題となる。その事象、たとえば労働条件の大きな変化があったとする。そのときの主観応答の結果に「変動量」として何らの相違も見出されなければ、仮に、労働条件の変化を評価するといっても、その目的は達成されないことになろう。性格検査においても、個人の性格を信頼性のある「固定的なもの」として扱うのはどうかということになる。もし、性格検査は、行動場面の相違に応じて変化する部分を

捉えているのではないとするならば、職務適応性とか、職場適応性を云々することはできなくなるはずである。なぜなら、適応性という場合は、抽象的な職場名、職種名ではなく、現実の職場があって、労働条件も違えば、設備等の環境条件も異なる。いっしょに仕事をする同僚、上司からの管理の違いもある。仕事の性質、難しさなどといったら、簡単にとらえられるものではない。これらの「条件・事象」との関係を理解する努力が不足しているからである。適応性というときには、行動場面の相違・変化に対応する「行動様式の変容」が期待されているのではないだろうか。

『作業後症候しらべ』という調査表がある（太田垣，1958）。疲労感調査法の一つであって、さきに触れた『自覚的症候調査表』、および『蓄積的疲労徴候インデックス・CFSI』（越河・藤井，1987）の前身といってもよい。『作業後症候しらべ』は60項目の設問があり、「身体的症状」「感覚的症狀」「神経的症狀」「観念的症狀」「意思的症狀」の5項目群に分類されている。これらの分類は作成者の「選択」によるものであるが、職場、作業場面観察のための有効なツールとして用いられてきている。

昨今では、繰り返しになるが、統計的解析を経ない「スケール」などは認められそうにないけれども、要は、その“洗練された”スケールの使い方というか、調査の目的との関連が、もうひとつ問われていることに気づかなければならないと思う。何らかの調査用紙を持ちこんで回答を求めると、“被検者”から「この調査をやるって何がわかるの」「結果を何に使うの」と聞かれることがよくある。このような質問に対しては、すぐ答えることはなかなか難しい。マニュアルの効能書きを伝えるぐらいではどうにもならない。これが現場調査である。

“疲労調査など、現場調査は、まず、その現場の「問題」の有無を見わけること、問題があれば、それらの解明、そして改善対策を考えることになる。だから、その現場、すなわち、労働の場面・作業の場面を直接的に観察することだ。とにかく現場に「触れること」、そこで仕事している人の話を聞いたり、何回も足を運ぶことが大事だ。そうしていると、だんだんモノが見えてくる。何らかの質問紙を配布し、それに回答を求めれば、それだけですべてわかるわけではない”（太田垣さんの談話から）。

質問紙作成に当たっては、質問の「言葉づかい」のほか、質問項目数についても考慮が必要である。質問が多ければ多いほど正確で詳しい情報が得られると思ったら大間違いである。回答者は、まともに答えなくなることすらある。調査となると、調査者の一方的な意図で実施される場合が多いが、被調査者の方は、その「調査」に都合よく準備された状況にあるとは限らず、膨大な質問を前にして躊躇するのは当然である。

産業精神衛生学の領域を研究している人が、あるメンタルヘルス関連の尺度について、その作成者に、設問が数百項目もあるのは、なんとしても多すぎはしないかと尋ねたら、これくらいの数の質問に答えられないような被検者は、もともとまともに仕事などできる人ではないですよ、と反論されたという。答えた人が本当にその尺度の作成者なのかどうか疑わしくなる話である。

産業衛生学会・疲労委員会で、先ほどの「自覚的
症状調査表」の改定を協議する集まりのとき、梶原さん（梶原三郎氏）は、「疲労感」^{註1）}をあらわす言葉が一つで済むのならいいのだが、そうもいかないしね、と話されたという。ある程度の項目数はいたしかたないということである。

3. 調査の実施に関連して

「疲労感調査」などでは、評定法に拠ることが多い。一応、標準スケールの形を整え、マニュアルも作成しているわけだが、それらのスケールを実際
に使用する場合は、スケールの主旨を充分理解することが肝要である。マニュアルどおり、そのままなぞってやれば使えるというわけにはいかない。

CFSI（蓄積的疲労徴候インデックス）のマニュアルとして、「労働と健康の調和」（越河・藤井、2002）を刊行したところ、「実施手引き」のような簡単なマニュアルにしてもらえないかという要望があった。この人は企業の健康管理担当者で、それさえあれば、難しい説明はいらない。当方で実施できるからというのである。出来合いのスケールの使い方に馴染んだ人と思われた。

市販されているモノであっても、よく選んで、理

解して、使用法にも習熟して利用するときは、「出来合いの」使い方ではなくなる。狩野さん（狩野広之氏）は、「精神的能力の発達に関する逐年的研究」（1950）に着手されるとき、直接、鈴木治太郎博士のところへ通って「鈴木・ビネー知能検査法」を確かめられたと聞いている。

調査の実施に当たっては、被調査者に調査の目的を説明（インストラクション）する必要がある。これも配慮を要する。例えば、労働負担の様子を調べようとするとき、「これは、労働負担の調査です」というのはいささか無神経であり、誘導していることにもなりかねない。CFSI（蓄積的疲労徴候インデックス）は、ものものしい表現となっているが、実際の調査票では、「勤務と生活と健康に関する質問」という表題であり、インストラクションも、「この質問票は、起床・就床の時刻など、生活の様子と勤務時間にかかわる事項、および普段の健康状態等についてたずねております。それぞれの質問に答えてください」としている。

調査結果の取り扱いに関連して、「本調査は、個人的に迷惑のかかることは一切ありません」と明記することがあるが、これは大体信用されない。ある「健康調査」では、個人の結果は、個人に返し、例えば、健康上異常ありと解釈された場合は、専門医の受診を勧めるという「保健指導」を行い、会社側には全体的な傾向を報告するという段取りで実施されている。しかし、実際のところそのとおりかどうか。調査活動においては、調査者と被調査者の間の信頼関係が大事である。

記名方式か、無記名方式かの違いは、調査の目的にもよるが、「応答」への構えも変わるので留意しなければならない。性別とか年齢など、属性の記入も同様である。

回答の求め方。一定の場所に集まってもらって、一斉に記入してもらう方式とか、持ちかえり方式などさまざまである。これも調査の目的に関連する事柄で、それぞれの方式に合う設問の仕方の工夫もあるから、すぐ、どれがよいか、どれが悪いかなど、きめられないことがある。

「読み上げ方式」で回答を求めるやり方はどうか。このことについては、さきにもちょっと述べたが、あまり強くこだわる必要はないと思う。記入者は、何秒とか、調査者の決めたペースに合っている人だ

註1）「疲労感」：本稿では「何らかの活動を続ける過程で生ずる心身の違和感・自覚される症状」を総称するものとして用いている。

けではないからである。ゆっくり回答していると、あまり深く考えるなどか、速く答えなさいとかうるさい。タイムリミットのテストならともかく、主観調査では、この辺のところの配慮が望まれる。

なお、機能検査などを含む「疲労検査」の場所で、調査者の目の前で、「自覚的症状調査表」に回答してもらった場合と、同じ職場で労働組合を通して配布・回収（持ちかえり方式）された場合とでは、平均訴え率で数ポイントの相違がみられたことがある（1964）。

労働組合経由の訴え率が高く出ている。この結果についての解釈は別として、いずれにしても、調査方式の違いによる応答結果への影響は否定できない。

4. 調査結果の扱い

質問紙への回答は、「エントリー」作業に送られ、集計分析の段階に入る。マークシート方式では、すぐさま機械的に処理される形になる。回答原票を1枚1枚丹念に目を通すことは少ないようだ。○×方式の回答でも、そのつけ方に表情がある。丁寧に回答した人とか、事務的に答えた人、誤りのあった人などいろいろある。疲労感の調査などでは、ぴったりの項目があったりすると、◎印で答える人、アンダーラインを付す人もいる。反対に投げやりな回答もあり、データとして使用できない場合もある。いずれにしても、○や×以外の情報を読み取ることができる。多変量解析に持っていく前の段階での処理が大切である。

代理店指導に当たっていた営業マン（40歳代）が、担当の、ある代理店に少し間を隔てて訪問したところ、その代理店のおやじさんが、1週間ほど前に死亡されていたことを聞かされた。驚いて取って返し、喪服に着替え香典を包んで、あらためて御悔やみを済ませたが、その帰り、運転する自分の車の中で、くも膜下出血のため急死。労災が認定された事例である。この人は、大都市の支店に勤務（事務系）していたが、地方都市の支店に人事異動で、苦しい営業に配属されてから、それほど経っていない時期の事故であった。担当エリアは山間部が多く、自動車運転時間が1日平均200キロを超える状況で、それが事故前2～3カ月間続いていたという記録がある。事故の1カ月ほど前に、たまたま「健康

調査」が実施されていて、筆者の手元に調査原票があった。無記名方式の調査ではあったが、職務と年齢段階などから、本人であることが判明。労災認定に当たった意見書を求められていたので、早速調べてみたところ、疲労感スケールの欄外に、「すっかり疲れました」というメモが残されていた。疲労感の訴えは平均的なレベルであったが、このような調査票への走り書きなどは貴重な情報を寄せてきてくれているのである。報告書にもその旨は記しておいた（1978）。

疲労感調査などでは、設定されている心身症状項目のうち、該当する項目の数が多いか少ないかで、疲労感の訴えが高いとか低い、というように判定されるのが一般的である。しかし、その値と、「労働負担」の大小が必ずしも平行関係にあるということではない。「訴え」が多いから、疲れが大きいとか、反対に訴えが少ないから疲労は少ない、というわけにはいかないということである。なぜ、そのような応答が生じたか、その背景要因を探るのが目的のはずであって、どのような職務の状況にある人が、どんな「応答」を示したかの関係を求めることが肝心である。統計処理で、仮に職種間に「有意な」差が認められたから「差がある」として片付けしないで、なぜそのような結果が出たのかを「探す」のが調査であろう。

金融関係の企業の大きな支店で、作業分析を行ったことがある（1975）。調査の世話役をしてもらった支店次長さんに、前述のCFSIへの回答を求めた。CFSIは、81の心身症状項目について、最近、それらの体験があったかどうかを報告してもらい、評定尺度の一つであるが、次長さんの応答は、1項目だけ該当するというものであった（男子の平均訴え数は15項目）。しかし勤務状況は、残業が続いており、睡眠時間は1日平均3時間ほどしかない。一見、元氣そうで、“睡眠時間は訓練すると3時間で充分です”などと話してくれた。そんなはずはないですよ。これだけ忙しいのですから、と言ってみたものの聞き入れてくれなかった。しかし、調査が終わり10日間ほど経った頃、その支店を訪れたら、次長さんは欠勤しておられない。神経性の疾患で入院されたのだという。自覚症状の訴えが少ないから、問題がないとするわけにはいかない例としてあげてみたが、やはり、職務の状況とか、勤務と勤務外の過

し方とのバランスなどが点検されないと調べは不充分ということである。

5. 要 約

本論文では、産業場面でよく用いられている、疲労感評価や、ストレスチェックのための質問紙などに関して、その問題となるところを、調査の体験を通して考察した。質問紙法（評定尺度）は、大別すると、性格検査のように個人の特性を調べるものと、労働負担など、その人を取り巻く周囲の事情を探ろうとするものの2つに分けられる。疲労検査などは、後者の目的でなされるわけで、質問紙への結果だけでなく、応答と職場の労働条件などとの関係が

検証されるべきである。この点からすると、質問紙法の使い方にはかなり安易なところが目につく。

引用文献

- 狩野広之 1962 精神的能力の発達に関する逐年的研究
労働科学研究所。
越河六郎・藤井 亀 1987 蓄積的疲労徴候インデックス（CFSI） 労働科学 第63巻5号。
越河六郎・藤井 亀 2002 労働と健康の調和 労働科学研究所。
太田垣瑞一郎 1958 主観的疲労感の評価 産業心理ハンドブック。同文館，pp.792-795。