

……いんふおるむ(第45回)……

質問の仕方の違いによる回答への影響

—— カールトンとシューマンのレビュー(2) ——

Graham Kalton and Howard Schuman (1982): "The Effect of the Question on Survey Responses: A Review", *Journal of the Royal Statistical Society, A*, 145. Part .PP42 - 73
より

翻案 鈴木 達三
帝京平成大学教授
統計数理研究所名誉教授

今回は質問の仕方による回答への影響について、「実態項目の質問」と「実態項目以外の質問」とにわけて考察してきた。これは、この二つのタイプの質問のワーディングや形式についての研究上の差を反映させるため便宜上仕分けした区分である。

したがって、一方でのべた問題点が他方では全くないということではない。これらのことを含めて、以下では「実態項目」「実態項目以外」の区分のない、全般的な質問の影響について考える。

4 全般的な質問の影響

(general question effects)

これまでの実態項目のところでのべた問題点も意見項目で生じる。たとえば、センシティブな事項は意見項目のときに、よりはっきり生じるし、記憶の問題も、意見の変化に関する調査の時には問題になる(もちろん、過去にさかのぼって質問するやり方で正確な情報を収集することは特に困難なことである)。

同様に、実態以外の質問項目の質問文と回答形式について、いろいろのべたことがらは、

そのままでは実態項目の質問にはあてはまらないけれども、実態項目の質問も質問文や回答形式の変更に影響される。

たとえば、Locander and Burton (1976) は、家族収入を聞く質問項目で、質問の仕方と回答のとり方の違う4種類のタイプを作成し、これを電話調査により実施し、調査結果の収入分布に大きな差異が生じることを示した(回答肢の順、回答のとり方(yes, no)のタイプの組合せのちがいで)。

この4つのタイプは

タイプ 家族収入は ドル以上か?

(順に収入が上の方になるようにきき“no”というまで続ける)

[= 5000ドル、7500ドル、1万ドル、15000ドル、2万ドル、25000ドル]

タイプ は逆に25000ドルから順に下にきき“yes”という回答が得られるところまで続ける。

結果はタイプ では37.5%が家族収入は15000

ドル以上と回答し、タイプ では、これに対応する収入層のパーセントは63.7%であった。

このように、質問と回答の形式が変更されると回答が影響されることを示した。

調査票にある他の質問の影響

最後に重要な質問の影響は調査票にある他の質問の存在である。また、それらの質問の場所、すなわち、これらの質問と今問題にしている質問との相対位置の影響である (context effects)。

質問の順および質問の文脈や前後関係の影響は実態項目および実態以外の項目の両方に生じる。しかし、影響の仕方は異なるように見える。

意見項目の質問の回答に対する質問の順の影響については数多くの実験がある。多くの場合、順序の影響 (order effect) は見出せなかった (主題に密接に関連した質問のときも)。

しかし、あるタイプの質問順では二つの場合に順序による回答への影響 (order effect) がみられる。これはもっと検討する価値がある。

一つは、一般的質問 (general) とそれを特定化 (specific) した質問の組合せのときである。

特定化した質問の回答結果は影響されないが、一般的質問の回答結果は、

特定化質問が前、一般的質問が後 のときと 一般的質問が前、特定化質問が後 の時で変化する (一般的質問が後のとき、回答分布 (%) は減少する)。

この例を二つあげる。

〔第一の例〕^(注)

Schuman et al. (1981) は中絶に関する二つの意見質問で、これについてのべている。

Schumanらは全く違う目的のため1979

年3月中絶に関する態度の項目を、NORCの一般社会調査 (GSS 調査) の中絶問題のシリーズ質問の中から一項目借用し、電話調査した。

調査結果の回答分布は1978年GSS調査のCodebookに報告されている%と大きく異なっていた (表1)。

はじめは、この差は電話調査とGSSの面接調査の差か、属性構成などによる差と考え、差がなくなるようにいろいろ統計的な修正を試みたが、これらの修正もうまくいかなかった。

この調査のとき、単独で借用した中絶の質問項目はGSS調査の中絶問題の質問項目のシリーズでは二番目に質問される項目であることに着目して、1979年6月にこの二項目の順序の影響 (order effect) に関する split - ballot (折半法) による実験調査をおこなった。

表1 NORC (GSS調査) とSRC (ミシガン大社会調査研究センター) 調査の間の中絶に関する一般的質問結果の差 (%)

	NORC _{a)} GSS - 78	SRC _{b)} 79 - 3月	学歴で _{c)} 補正
はい	40.3	58.4	54.9
いいえ	59.7	41.6	45.1
計	100.0	100.0	100.0
サンプル数	(1483)	(490)	

出所 Schuman and Presser (1981, Table 2.5)

a) 中絶に関する質問項目の二番目に出てくる、前の質問文は“あなたは、もし、妊婦がお腹の赤ちゃんに重大な障害の可能性が高いときには、合法的に中絶できるようにするべきだと思いますか？”

b) 中絶に関する先行質問はない。

c) SRCの結果をNORCの学歴分布で標準化した結果。

結果は、前に偶然得られた差を、全く同様に再現した。

表2の1979年6月の結果からみると、一般/特定化のときの一般の回答分布(%)より、特定化/一般のときの一般の回答分布(%)は減少することが分る。

それ以降の調査の場合も同様で、全体をまとめたデータからみると、11%も減少する(注)。

このcontext effectsの傾向は基本的には頑健な知見であるといえるが、何故そうなるのかの説明(解釈)は非常に困難である。

表2 中絶に関するcontext effectsの実験(9回のくり返し)結果(RDDによる電話全国調査)(注)

調査年	月	一般的質問で既婚婦人中絶に好意的な%		差
		既婚(前)/ 障害(後)	障害(前)/ 既婚(後)	
1979	6	60.2(309)	48.0(302)	12.2
1979	8	66.6(326)	49.2(305)	17.4
1981	2	54.6(163)	45.5(178)	9.1
1981	5	55.8(113)	51.8(112)	4.0
1981	7	52.5(120)	47.5(120)	5.0
1981	10	63.1(187)	54.1(185)	9.0
1983	7	62.0(158)	39.8(196)	22.2
1986	7/8	63.2(337)	54.7(327)	8.5
1989	3/4	56.3(245)	47.4(247)	8.9
まとめたデータ		60.2	49.2	11.0
サンプル数		(1,958)	(1,972)	

()内 各調査のサンプル数
出 所 Schuman, H. (1991) (表2.2)

(注) この後にも Bishop, Oldendick and Tuchfarber (1985) の実験やさらに大規模な実験により、今では“abortion context effects”として知られている。
Howard Schuman(1991): “Context Effects: State of the Past/State of the Art”, in Norbert Schwarz and Seymour Sudman(eds); “Context Effects in Social and Psychological Research”, Springer-Verlag, Chapter 2. pp5-20 より引用

〔第二の例〕

Kalton et al (1978) は自動車の運転

の仕方に関する二つの意見質問の例を示した。これは、「一般の人の運転の仕方」と「若い人の運転の仕方」についての質問で、普段の自分の運転の仕方よりよくないという割合は、「一般の人」をはじめに聞いたとき34%、であるが「若い人」の後で「一般の人」について質問したときは27%で7%減少している。しかもこの順序(order)の影響は回答者の年齢が45歳以上のときのみで大きく12%も開いた。

この理由はまだはっきりしていないが可能な説明は「割引効果」である。すなわち、「一般の人」を質問したときは「若い人」も入るが、「若い人」について質問してからの「一般の人」に対する回答には、「若い人」を除いているという考え方である。

〔実態調査の例〕

実態調査のときにも他の質問が、その質問の回答に影響することがある。これは同じような、項目の長いリストがあって次々に回答するとき、たとえば、新聞、雑誌の閲読率などの質問項目では、いくつもある新聞名のリスト、雑誌名のリストをつぎつぎみながら回答する。ともすればリストの後方に位置するものの閲読率は低くなる。

イギリスの全国雑誌閲読率調査で Belson (1962) は提示リストの順を変えて調査した。週刊誌の提示順は最も影響を受け、あるものは最初のときにくらべ最後のときでは3/4になった。

他の例として、全国犯罪調査(National Crime Survey)では中核の被害調査の前に態度をきく質問項目の補足調査を実施してみると

補足の態度調査+中核の被害調査のとき、被害体験は増加し、平均して20%くらい個人的犯罪被害の経験率が高く報告

された。また、所有物に対する犯罪被害経験率では13%増加した。この影響は、補足質問がよりよく報告を引き出したということで、これは、過去12カ月間に受けた被害の記憶をはっきりさせる効果があったということであるが、単純にそれだけとするわけにはいかないということも分った。

このように補足質問を加えたとき、中核になる質問に本質的な影響があったことは、異なる調査の間の比較可能性について重大な関心を引き起こした。

調査分析者は比較しようとする質問について、結果の異なる調査の間の差についての解釈には十分よく注意して、慎重であるが、上にのべたようなことから、調査票の残りの(他の質問についての)差異にも関心をもち、検討しなければならないと言える。

この結論は調査結果のくり返し(replication)にとって極めて重大な影響をもたらす。というのは、通常かなり安易に単独の質問、あるいはいくつかの質問の組を比較のためくり返すことがよくある。このとき、質問票の全体をくり返すのは極めて困難である。

くり返し調査の重要性と変化の測定は調査データの分析にとって重要であるという見地から、この分野の研究が切に望まれる。

5 結語 (Concluding Remarks)

まとめと注意

このレビューの全般的な結論は、調査で質問することは物事をはっきり決める手段ではないということである。

調査関係の文献には、実態質問項目について重大な回答誤差の可能性、実態に対する影響の大きさ等を示唆するものが多い。また、意見項目に関する多くの実験では、質問された

ことのほんのささいな目につかないような変更でもときには大きく影響することもあることを示している。

今実施されている調査の質問の仕方の限界についてのこの結論は調査法研究者にとって申し分のないものであろう。

しかし、調査を実施したり、調査データを多方面の目的に利用したりする多様な人びとには十分に認識されていない。

このような事情(状況)を考えに入れて、経験のある調査の実験家は回答結果をそのままのレベルで取り扱うとき大変注意深く考え、どちらかといえば、集団内のサブグループ(各グループ)間の比較や、他の調査結果との比較に重点を置くようにしている。真の差異を推定しようとして、これらの対照結果を解釈することについては、ある一定の偏りが各グループにわたって、あるいは各調査にわたって存在するかも知れないが、これは、グループ(調査)間を比較、対照することによって消去されると考えていることになる。これは正に、前回のべたform resistant correlationの仮説である。

この仮説は、しばしば合理的な近似として役立つが、それでも無批判にこれに頼るべきではない。

意見項目について、それがうまくなかった例をいくつかとり上げて報告した。また実態項目についても異なる偏りが生じる可能性もあることを示した。

異なる調査の結果を対比させる場合の特別の問題としては、前節でのべた質問順序(質問の位置)と質問の前後関係(他の質問との関連)の影響がある。

実態項目および意見項目の両方とも調査票における他の質問項目の有無により、回答結果が敏感に動くことが示されたことは、調査間の比較・分析をよく行う調査研究者にとっ

て大きな関心事にちがいない。

この問題にとってよい説明（イラスト）は Turner and Krauss (1978)^(注)によって示された。彼らは全国的組織の信頼性のトレンドに関する NORC と Harris のデータの間の違いを説明することが可能な context の相違点を引き出した。もちろんこの説明は疑わしいということもできるので調査間の比較が困難なことを示唆している。

(注) Turner, C., & Krauss, E. (1978) Fallible indicators of the subjective state of the Nation. *American Psychologist* 33, 456 - 470

方法についての研究から得られた証拠により、調査手順のうち、質問方法に関する側面で改善の余地が多くあることが示された。

この分野では、これまでもかなりの量の研究がなされているけれども、我々には質問文のワーディングや形式の影響の性質について、また、どのようなときにその影響が出て、どんなときには出ないのか、さらに、どのようにふるまうのか等について分らないことが多く残されている。

これまで、実態項目に関する研究の多くは、単に回答誤差の大きさを評価するだけであり、意見項目についての研究は、質問文を二つ（あるいは多く）作成し、その間の差異を検証しているだけであった。

現在の研究は質問文の影響の研究を組織的に始めたところである。これは、まず、影響のタイプを分類することを試み、次に、その影響タイプに潜在する心理的過程を理解しようとする。研究のこの段階はまだ幼稚なものであるが労力は人をひるませる程のもので、進み方は遅く、主要な突破口も期待できないようなものである。

これまでのように、どちらかといえば一つ

一つが孤立した実験よりも、今必要なことは、質問文の影響に関する理論的構造を作り上げ、そのテストを目的とする開発的計画にそった一連の研究である。

質問過程に含まれる要因およびその相互関連のあり方をもっとよくはつきりと理解するまで、よい質問文を作成する基盤を欠くことになる。

可能性の一つは現在ある心理学の理論がこれまでどのべたタイプの影響について幅広く説明することが出来るということである。

たとえば、Bishop et al (1981) は最近の認知理論はそのようなことを取り扱うフレームワークを用意しており、彼らが、偶然に手に入れた単純な内容に関する影響を解釈するとき利用して、このことを示している。

しかし、理論が予測できるようになった形で利用できるのであれば、あるいは少なくとも、すでにあるいろいろな実例に応用できるのでなければ、一回限りの、偶然に出会った調査結果を事後にうまく説明できたという以上の理解を得るのは困難である。

どちらかといえば変わった疑いもなくゆっくりとした、より一層骨の折れる接近法の一つは、我々の最近の研究から示すことができる。

これは、前述の abortion (中絶) の二つの項目の文脈（前後の関係）による影響の分析から出発する。

最初の影響はすでに述べた通り偶然に見つかった。次には実験的に NORC の GSS 調査にある二つの隣接した項目を含めくり返し調査した。最初の段階は、中絶に関する質問をオリジナル質問項目と概念的に同じと考えられる他の項目ととりかえることであった。これは影響が NORC の元の質問項目の質問の正確な文章（ワーディング）に限られているかど

うかを定めるためである。

これを一つの項目にその時実行した。そして今では文脈（前後の関係）の影響は中絶を扱うオリジナル項目だけでなく、他の平行して聞いた質問をもこえて、より一般的に及ぶことが分っている。

次の段階は、この影響はこれまでの実験がすべてそうだったように項目が隣接している必要があるかどうか、あるいは二つの質問項目が関連のないいくつかの項目で分離されているときにも生じるのかどうか確かめることである。

はじめの結果は前者の必要性を支持するものだった。すなわち、先行質問の影響は隣接の条件をなくすと消失するようにみえた。しかし、これは一回だけであり、くり返しが大切であり現在それを実施している。

最後に中絶に関する context effects が質問文のワーディングや質問項目の位置にかかわらず、一定度の一般性を確立することができたとすれば、この影響の原因となるものについて、仮説を立てそのテストを試みることができる。

これまでに、回答者に対し、追及の質問で自由回答形式の項目を利用してみたがうまくいかなかった。回答が一つの文脈から他の文脈へかえても変わらないからである。

いろいろな仮説を厳密にテストする方法を考えているが、いまでもまだ、オリジナル項目の影響の原因について、満足できる結論に達するにはほど遠いところにある。

一般化と原因探究のこの全過程には一連の実験が必要である。実験は、その一つ一つが、はじめの偶然的知見に対する理解をより広く、より深くしようとするものである。

その一つのケースについて、終結にもちこむことができれば、他のケースもこれと同じ

ようにしてもちこむことができるかどうか、あるいは、そうならなければ、別に他の知見を集め、他の説明を考えるとところからはじめるか、きめなければならない。

この接近法は、明らかに時間をくい、いやになるほどゆっくりしたものである。しかし、このような一連の組織的探究がまず context effects の性質について、つぎにより一般の response effects の性質について理解するために必要であると考えられる。

他の分野における科学の進歩も通常スモールステップをくり返して徐々に進む形である。

われわれの関心のある調査における回答の影響についての理解を進めるところが同様の努力をしないですむということはありません。

もし、われわれが偶然出会った effects とその事後的説明だけですませることから先に進むとすれば上にのべたような中くらいの規模の探究的研究は本質的に重要である。

ところで、調査の実際家はこの論文で示したような人工的産物（すなわち、多様な質問の利用、多様な内容、多様な研究調査の方式の利用等）に対する強力な防具をもっている。

response effects により調査結果が損なわれることは希ではないが、それがすべての調査項目に同じような仕方でも広く生じることもない。調査にとって重要な概念を少なくともいくつかの項目に結びつけ、それらを形式やワーディング、context 等でそれぞれ異なるようにしておけば研究者は本質的な知見に対する人工的な回答ミスによるタコツポにはまりこむ危険は避けることができる。

以上、Kalton ,G .and H .Schuman(1982): “ The Effect of the Question on Survey Responses : A Review ”, Journal of the Royal Statistical Society , A 145 . Part

1 . pp42 - 73 .

の概要をのべた。一部はSchumanらのその後の研究のうち、“ abortion context effects ” に関することをSchuman ,H.(1991): “ Context Effects : State of the Past / State of the Art ” ,in Norbert Schwarz and Seymour Sudman(Eds) , “ Context Effects in Social and Psychological Research ” Springer - Verlag , Chapter 2 , pp5 - 20 . から一部引用している。

これからも明らかかなように、このレビュー以降に、response effectsに関する研究が進み、多くの研究者が参加してシンポジウムが開かれ、その論文集が出版されている。たとえば

Hans . J . Hippler , Norbert Schwarz , and Seymour Sudman ,(Eds)(1987): “ Social Information Processing and Survey Methodology ” Springer - Verlag . New York , Tokyo

Schwarz ,N .and Sudman ,S ,(Eds)(1991): “ Context Effects in Social and Psychological Research ” Springer - Verlag ,New York , Tokyo 等である。Schumanが “ まとめ ” でのべているようにcognitive psychologyとの関連なども研究されている。

ここでとり上げた“ abortion context effects ” を含めた context effectsのことはまたの機会に紹介したいと考えている。