

質問の順序効果の2類型

——位置の効果, 位置と質問の交互作用効果——

山 口 洋

〔抄 録〕

調査票調査で複数の対象をある観点から評価させるとき、質問の順序効果が様々な形で生じることが知られている。本稿の目的は、効果の検定方法や対象の評価方法に関わる統計学的な観点から、過去の研究で示された種々の効果を類型化し、各類型の特徴を明らかにすることである。2個の質問に2個の回答カテゴリがついた最も単純なケースにしほって検討した結果、種々の効果は「位置の効果」と「位置と質問の交互作用効果」に大別できること、位置の効果は **Mainland-Gart** の方法で、位置と質問の交互作用効果は **Hills-Armitage** の方法で検定できることが明らかにされた。最後に、スプリット・バロット法による調査で、以上の各類型の効果が観察されたとき、どんな集計方法で各対象の評価を行うべきかが検討された。

キーワード 調査方法論, 質問の順序効果, 位置の効果, 位置と質問の交互作用効果

1. はじめに

調査票調査において、同一質問項目の回答傾向が質問の順序によって変わることを質問の順序効果 (**question order effect**) と呼び、古くからやっかいな問題のひとつとされてきた。

調査論の分野で集中的に研究されてきたのは、様々な対象をある観点から続けて評価させる場合の順序効果である⁽¹⁾。例えば「家庭生活」「職業生活」といった生活領域別に満足度を答えさせたり、政党名を列記して、それぞれの政党に政権担当能力があるか否かを答えさせたりするケースである。こうしたとき、「家庭生活」を先にきいたときと「職業生活」を先にきいたときとで、それぞれの満足度が違ってきたり、ある政党を先にきいたときと別の政党を先にきいたときとで、それぞれの政党の評価が違ってきたりする。

スプリット・バロット法⁽²⁾を用いた実証研究によって、この種の順序効果には様々なタイプがあることが明らかにされてきた。**Schuman & Presser (1981)** が示した整合効果 (**consis-**

tency effect) と対比効果 (contrast effect), Moore (2002) の加算効果 (additive effect) と減算効果 (subtract effect), Schuman&Ludwig (1983) が例示した公平性の規範 (norm of even-handedness) といったものがそうである。これらのタイプ分けは順序効果を生み出す様々な社会心理学的プロセス^③を明らかにする上で有益だったが、それぞれのタイプの統計学的な意味合いの違いや、これらの効果が生じた場合に、どのような分析手順で結論 (すなわち個々の対象の評価) を導けばよいのかについては、十分明らかにされてこなかった。

本稿の目的は、①上記の様々なタイプの順序効果を包括する基本的な類型として、統計学的な観点から「位置の効果」と「位置と質問の交互作用効果」を定義し、②各類型の効果の検定方法を示すとともに、③各類型の順序効果がみられたときに、調査本来の問題関心である対象の評価についてどのように結論を導くべきかを示すことである。検定法や結論の導き方と連動した体系的な知識が得られる点が、本稿の類型化の特色である。

本稿の守備範囲を明確にするため、3点ほど付け加えておこう。第1に、順序効果には、個々の項目の回答の平均値や比率に対する効果と、項目間の相関に対する効果があるが、本稿が提示する2類型はもっぱら前者に関する類型である。ちなみに上記の整合・対比、加算・減算といった類型も主に前者に関するものである。第2に、本稿の類型化はもっぱら統計学的あるいは操作的な観点から行われる。つまり本稿の類型は、スプリット・バロット法を使った調査の集計結果のパターンを分類するものであって、順序効果を発生させる社会心理学的なプロセスを分類するものではない。位置の効果も位置と質問の交互作用効果も、それぞれ様々な社会心理学的プロセスを経て発生するものと考えられ、本稿でも折に触れてそれを論じることになるが、それが本稿の主な目的ではない。第3に紙幅の関係上、本稿では、一定の観点で複数の対象を評価させる質問群のうち、最も単純なケースのみを扱う。すなわち、ふたつの質問のそれぞれに2個の回答カテゴリー (集計時にふたつのカテゴリーに合併した場合を含む) がついているケースのみを扱う。また、それぞれの類型に対応した検定法も、この最も単純なケースに対応したものだけを示す。ただし本稿の類型化の基本的な発想は、こうした最も単純なケース以外にも適用可能である。

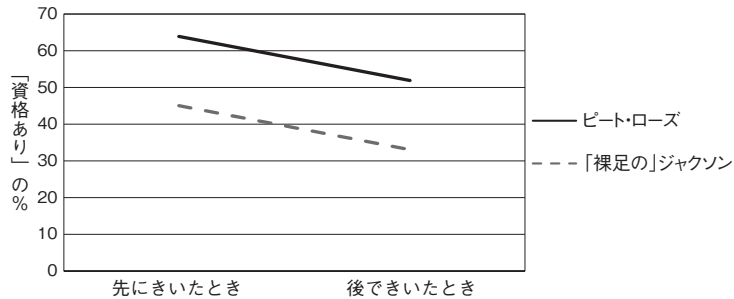
2. 類 型 化

2.1. 位置の効果

第1の類型は「位置の効果」である。これは質問項目の位置によって、各項目の回答に、ある共通した変化が現れる場合をさす。例えば、2個の対象を評価させる場合、先にきいたときより後できいたときの評価が高いという傾向が、どちらの対象にも共通して見られるとすれば、位置の効果が生じていることになる。

図1は位置の効果の実例である。これは米国のギャラップによって1999年7月に行われた

図1 野球殿堂入りする資格があるか？ (N=1061)



Moore (2002) より筆者作成

世論調査 (N = 1061) の結果である。ここでは、過去に偉大な記録を残しながらも、野球賭博にからむスキャンダルのせいで野球殿堂入りしていない2人の選手について、それぞれ野球殿堂入りする資格があるか否かを (スキャンダルの概略を説明した上で) 質問している。サンプルはランダムに2分割され、半数はローズ→ジャクソンの順に、半数はジャクソン→ローズの順に質問された。図1のように、いずれの選手についても「資格あり」との回答は、先にきいたときの方が多くなっている。逆に言えば、いずれの選手も後できくと不利な評価を受けやすい。Moore (2002) によれば、掟破りの選手の殿堂入りという「例外」が認められるのは、せいぜい一人目 (先にきいた選手) までで、二人目 (後できいた選手) の例外は認めないという回答者がいたため、こうした結果が生じたのだという。

図1のパターンについて Moore (2002) は、先にきいたときよりも後できいたときの比率が低くなることから「減算効果」と呼ぶ。同じく Moore (2002) は、先にきいたときよりも後できいたときの比率が高くなる事例も挙げ、これを「加算効果」と呼んだ。しかし、この区別はあまり本質的ではない。そもそも図1を「資格なし」という回答の比率にしたがって作れば、後できいたときの比率の方が高くなるから、「減算効果」ではなく「加算効果」が生じているということもできるからだ。

むしろ図1の特徴として重要なことは、グラフの2本の線が同じ側に傾いており、ほぼ平行だということである。グラフの線が傾いているということは、同じ質問であっても、置かれる位置によって回答傾向が変化することを意味している。また、2本の線が平行だということは、どちらの項目についても、回答の変化の方向性が同じであることを意味している。Moore (2002) が挙げる加算効果の実例も、グラフ化すれば (紙幅の都合上、省略するが) これと全く同じ特徴を示す。違いは2本の線が右肩上がりになることだけである。このように、Moore (2002) のいう減算効果と加算効果は「位置の効果」として一括して考えることができる。

2.2. 位置と質問の交互作用効果

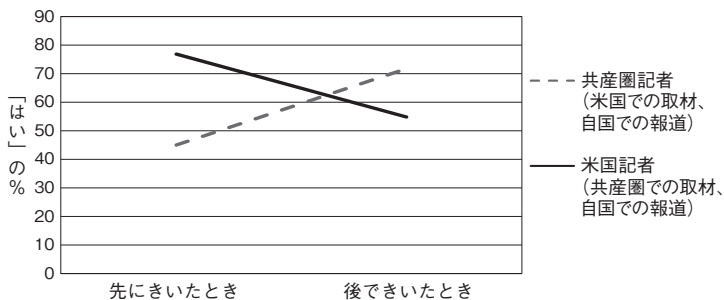
もうひとつの類型は「位置と質問の交互作用効果」である。これは質問によって、位置の効果の方向性が異なる場合をさす。例えば2個の対象を評価させる場合、一方は先にきいたときの評価が高くなり、もう一方は後できいたときの評価が高くなったとすれば、位置と質問の交互作用効果が存在することになる。

図2は位置と質問の交互作用効果の実例である（Schuman, Kalton, & Ludwig, 1983）。これは米国のサーベイリサーチセンターが1981年9月に行った電話による全国調査の結果（N=215）である。米ソ冷戦の時代を背景に、米国は共産圏の記者の米国での取材と本国での報道を許可すべきか、また共産圏の国々は米国記者の共産圏での取材と米国での報道を許可すべきかを、サンプルの半数にはこの順番で、別の半数には逆の順番できいている。「はい（許可すべき）」の割合は「米国記者」では先にきいたときに、「共産圏記者」では後できいたときに高い。

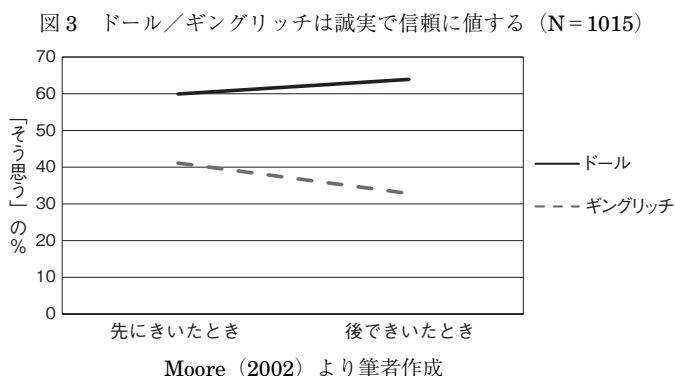
Schuman & Ludwig (1983) によれば人々の公平性の規範（norm of even-handedness）がこのような順序効果を生み出す。回答者は米国人だから、基本的には「米国記者の活動は許可すべきで、共産圏記者の活動は許可すべきでない」と感じている人が多い。しかし、米国記者→共産圏記者の順にきくと、「米国記者の活動は許可すべき」と答えた7~8割の回答者の大部分は、公平性を保つため「共産圏記者の活動も許可すべき」と答えることになる。逆に、共産圏記者→米国記者の順にきくと、最初に「共産圏記者の活動は許可すべきでない」と答えた約半数の回答者の多くは、やはり公平性を保つため「米国記者の活動も許可すべきでない」と答えざるをえない。また、Schuman & Presser (1981) は、後の質問の回答が前の質問の回答に近づくことから、図2と同様のケースをまとめて「整合効果」とも呼んでいる。

図3も位置と質問の交互作用効果の実例である（Moore, 2002）。これは米国のギャラップが1995年3月に行った世論調査の結果（N=1015）に基づいている。当時の共和党を代表する2人の政治家が誠実で信頼に値するかどうかをたずねている。ドールは先にきいたときよりも後できいたときの方が「そう思う（誠実で信頼に値する）」の割合が高いが、ギングリッ

図2 取材および報道を許可すべきか？（N=215）



Schuman, Kalton, & Ludwig (1983) より筆者作成



チはその逆になっている。ドール→ギングリッチの順にきくと、評判のよいドールと比較されて、ギングリッチの評価はますます低くなり、ギングリッチ→ドールの順にきくと、評判の悪いギングリッチと比較されて、ドールの評価はますます高まるという傾向が読み取れる。Moore (2002) は、この種の順序効果を、Schuman & Presser (1981) の用語を使って対比効果と呼ぶ。

図2も図3も、質問によって位置の効果の方向性が異なるという点では共通した特徴を持っている。したがって「公平性の規範」「整合効果」「対比効果」などと呼ばれてきた事例は、一括して「位置と質問の交互作用効果」と呼ぶことができる。

なお、ここまで挙げた事例は位置の効果のみ、位置と質問の交互作用効果のみが比較的純粋な形で現れた例であり、実際のデータでは両者の性質が同時に現れる場合もある。Moore (2002) も、整合効果および対比効果（本稿の括りでは位置と質問の交互作用効果）は、加算効果および減算効果（本稿の括りでは位置の効果）と同時に生じる可能性があるとして論じている。

3. 位置の効果および位置と質問の交互作用効果の検定法

3.1. スプリット・バロット法とクロスオーバー法

順序効果を観察するためのスプリット・バロット法の統計学的な形式は、医学実験におけるクロスオーバー法（cross-over study：交差研究）の形式とほぼ等しい。薬剤 A と薬剤 B の効果を比較する際、被験者集団をランダムに半分に分け、集団①に対しては A→B の順に薬剤を投与し、集団②に対しては B→A の順に投与するような実験デザインのことを、クロスオーバー法と呼ぶ⁽⁴⁾。この例の薬剤 A と薬剤 B を、質問 A と質問 B に置き換えれば、クロスオーバー法の統計学的形式は、スプリット・バロット法とほぼ同じであることがわかる。

したがって、クロスオーバー法のために開発された順序効果⁽⁵⁾の検定法はスプリット・バロット法による順序効果の検定にも用いることができる。具体的には、位置の効果の検定とし

て Mainland-Gart の検定 (Mainland, 1963 ; Gart, 1969)⁽⁶⁾が、位置と質問の交互作用効果の検定として Hills-Armitage の検定 (Hills & Armitage, 1979) が利用可能である。次節以下、質問の順序効果の例を使って具体的に説明しよう。

3.2. 検定法

無作為標本のデータから、整合効果・対比効果、加算効果・減算効果の実例を示した過去の研究 (Schuman & Presser, 1981 ; Moore, 2002) でも、効果の有無に関する統計的検定が行われている。しかし、これらの文献の検定法は、そこで示された各類型に直接対応した方法ではなく、一つの事例に現れる様々な比率の差を個々に検定 (カイ二乗検定) するものである。しかも、そうした複数の検定が相反する結果を示したとき、そこからどのように総合的判断を下すのが説明されていない。一方、本稿が示すふたつの検定法の特長は、それぞれが位置の効果および位置と質問の交互作用効果という類型に直接対応していること、また、一つの事例につき各一回の検定を行えば、その事例が各類型に当てはまるか否かを判断できることである。

位置の効果および位置と質問の交互作用効果の検定を行うには、まず、表1のような4行2列のクロス集計表を用意する必要がある。これは4×2のクロス表だが、実質的には、質問Aの回答×質問Bの回答×調査票種別という3重クロス表である。

まず、位置の効果(のみ)が存在するとしたら、表1のクロス表のセル度数がどのような分布になるかを考える。表2は、質問A, Bともに先にきいたときよりも後できいたときの方が「はい」の割合が高くなる場合について考えたものである。セル中の+と-は、回答パターンと質問順序とが独立だと仮定したときの期待度数よりも、実際の度数が多く(+), または少なく(-)なることを意味している。「はい(A) いいえ(B)」という回答は「A→B」よりも「B→A」のときに多くなるはずである。逆に「いいえ(A) はい(B)」という回答は「B→A」よりも「A→B」のときに多くなる。その他の回答パターンの度数は順序に影響を受けないはずである⁽⁷⁾。同じく表3は、質問A, Bともに先にきいたときの方が後できいたとき

表1 順序効果の検定のためのクロス表

回答パターン		順序	
		A→B	B→A
質問A	質問B		
はい	はい		
はい	いいえ		
いいえ	はい		
いいえ	いいえ		

表2 位置の効果が存在する場合(1)

回答パターン		順序	
		A→B	B→A
質問A	質問B		
はい	はい		
はい	いいえ	-	+
いいえ	はい	+	-
いいえ	いいえ		

+ : 期待度数より多い - : 期待度数より少ない

表3 位置の効果が存在する場合 (2)

回答パターン		順序	
質問 A	質問 B	A→B	B→A
はい	はい		
はい	いいえ	+	-
いいえ	はい	-	+
いいえ	いいえ		

+ : 期待度数より多い - : 期待度数より少ない

表4 位置の効果を検定するためのクロス表

回答パターン		順序	
質問 A	質問 B	A→B	B→A
はい	いいえ		
いいえ	はい		

よりも「はい」の割合が高くなる場合の度数分布を示したものである。

表2および表3から、位置の効果を検定するには、表4のような2×2のクロス表を作成し、これについて独立性の検定 (test of independence) を行えばよいとわかる。既に述べたように、この方法はクロスオーバー法における Mainland-Gart の検定と全く同じである。独立性の検定は、大標本について手計算するときにはカイ二乗検定で行い、小標本の場合、もしくは統計ソフトが利用できるときには正確確率検定 (exact test) で行うのがよい⁽⁸⁾。次節の計算例では正確確率検定を行っている。

次に位置と質問の交互作用効果を検定する方法を説明しよう。ここでも考え方の基本は、表1のような質問Aの回答×質問Bの回答×質問順序のクロス表である。やはり、位置と質問の交互作用効果 (のみ) が存在するとしたら、この3重クロス表のセル度数がどんな分布になるかを考える。

質問Aは前よりも後で大きく方が「はい」の割合が高く、質問Bは後よりも前で大きく方が「はい」の割合が高い場合、度数分布は表5のようになるはずである。「はい・はい」という回答は、順序「A→B」よりも順序「B→A」で多くなる。逆に「いいえ・いいえ」という回答は「B→A」よりも「A→B」で多くなる。その他の回答は順序の影響を受けないはずである⁽⁹⁾。同じく表6は、質問Aは後よりも前で大きく方が、質問Bは前よりも後で大きく方が「は

表5 位置と質問の交互作用効果 (1)

回答パターン		順序	
質問 A	質問 B	A→B	B→A
はい	はい	-	+
はい	いいえ		
いいえ	はい		
いいえ	いいえ	+	-

+ : 期待度数より多い - : 期待度数より少ない

表6 位置と質問の交互作用効果 (2)

回答パターン		順序	
質問 A	質問 B	A→B	B→A
はい	はい	+	-
はい	いいえ		
いいえ	はい		
いいえ	いいえ	-	+

+ : 期待度数より多い - : 期待度数より少ない

質問の順序効果の2類型 (山口 洋)

い」の割合が高い場合の度数分布を示している。

表5および表6から、位置と質問の交互作用効果を検定するには、表7のような2×2のクロス表を作成し、これについて独立性の検定を行えばよいとわかる。これはクロスオーバー法における Hills-Armitage の検定に等しい。Mainland-Gart の検定の場合と同様、独立性の検定は、大標本について手計算するときにはカイ二乗検定で、小標本の場合、もしくは統計ソフトが利用できるときには正確確率検定で行う。ちなみに次節の数値例では正確確率検定を行った。

表7 位置と質問の交互作用効果を検定するためのクロス表

回答パターン		順序	
質問 A	質問 B	A→B	B→A
はい	はい		
いいえ	いいえ		

3.3. 検定の数値例

図2で示した Schuman, Kalton, & Ludwig (1983) のデータについて、位置の効果と位置と質問の交互作用効果を検定してみる。いずれも有意水準は両側5%とする。まず表8のようなクロス表を用意する。これは原典で示された3重クロス表を本稿の形式に直したものである。なお表8では、共産圏記者に関する質問を質問A、米国記者に関する質問を質問Bとした。

表8 Schuman, Kalton, & Ludwig (1983) のデータ (図2の原データ：セルの数値は度数)

回答パターン		順序	
質問 A	質問 B	A→B	B→A
はい	はい	52	67
はい	いいえ	0	5
いいえ	はい	11	10
いいえ	いいえ	52	18

位置の効果を検定するための2×2のクロス表は表9のようになる。表の下に示した数値は、独立性の検定の結果を示すp値である。0.053 (正確有意確率：両側) という値を示すから、5%水準で有意ではない。したがって母集団において、位置の効果があるとは言いがたい。

質問A：共産圏の記者による米国の取材・報道を、米国は許すべきか
 質問B：米国の記者による共産圏の取材・報道を、共産圏の国々は許すべきか

位置と質問の交互作用効果を検定するための2×2のクロス表は表10のようになる。

表9 位置の効果の検定 (Mainland-Gart の検定：セルの数値は度数)

回答パターン		順序	
質問 A	質問 B	A→B	B→A
はい	いいえ	0	5
いいえ	はい	11	10

p = 0.053 (正確有意確率：両側)

表10 位置と質問の交互作用効果の検定 (Hills-Armitage の検定：セルの数値は度数)

回答パターン		順序	
質問 A	質問 B	A→B	B→A
はい	はい	52	67
いいえ	いいえ	52	18

p = 0.000 (正確有意確率：両側)

独立性の検定結果は $p=0.000$ (正確有意確率: 両側) となり、有意である。したがって、母集団において位置と質問の交互作用効果があるということが出来る。すなわち、質問によって位置の効果の方向性が異なるという図2の傾向は、母集団においても存在すると推定してよい。

4. スプリット・バロット方式の調査票から結論をどう導くか？

4.1. 順序効果がある場合の3種類の評価方法

方法論研究や予備調査ではもちろんのこと、実質的な目的を持った本調査でも、ここまでみてきたようなスプリット・バロット法は有効である。現実の調査では、順序効果の発生が懸念されつつも、関連する質問を続けて行うことが不可避な場合が多いと思われるからだ。

では、スプリット・バロット法による調査の結果、位置の効果や位置と質問の交互作用効果がみられた場合、調査者本来の問題関心についての「結論」をどう導いたらよいだろうか？すなわち順序効果が存在する場合に、ふたつの対象の評価の高さに関して、結論をどのように導いたらよいだろうか？

順序効果が存在する以上、ある質問順序の調査票だけから結論を導くのは論外とすると、大きく分けて3種類の集計・評価の仕方がありうる。すなわち①2種類の調査票の結果を合算して評価する、②先にきいたときの集計結果で評価する、③後できいたときの集計結果で評価する、という3種類である。ただし合算する方法は、2種類の調査票がほぼ同数のサンプルに割り当てられているときにしか使えない。これらのどの集計法が妥当であるかは、順序効果の種類による。特に位置の効果のみが存在するケースと、位置と質問の交互作用効果が存在するケースに分けて考えることが必要である。

4.2. 位置の効果のみが存在する場合

位置の効果のみが存在する場合、すなわち図1のようなケースでは、先にきいたときを比較しても、後にきいたときを比較しても、ふたつの対象の相対的な評価は変わらない。図1の場合であれば、いずれの比較においても、ローズの評価(資格ありとする人の割合)はジャクソンの評価を上回っている。したがって、両者の相対的な評価のみが問題関心であるなら、2種類の調査票を合算して比較するのがよいだろう⁽¹⁰⁾。合算することによってケース数が倍になるから、より信頼性の高い結果が得られることになる。

判断に迷うのは、評価の絶対的な高さ、つまり図1のケースで言えば、それぞれの選手について「資格あり」とする人の割合(%)に関心がある場合であろう。先にきいたときの%と、後できいたときの%に違いがみられるからである。

極力、順序効果を排除した結論が得たいのであれば、先にきいたときの%だけを信頼すべき

だろう。郵送調査や留置調査では、後方の質問から先に回答する人もいるかもしれず(Schwarz & Hippler, 1995)、この方法で順序効果が完全に排除できるとは言い切れないが、電話調査や面接調査ならば、先にきいたときの%は単独できいたときの%と同義なので、ほぼ完全に順序効果を排除することができる。

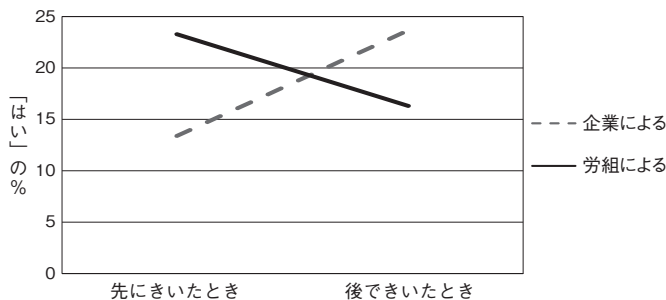
これに対して、2種類の調査票の結果を合算する方法は一見、公平な判断ができるように思われるが、後できいたときの集計結果が半分混じっているため、順序効果の影響が完全に排除されるわけではない。また、後できいたときの結果を利用することは、それなりの積極的な理由がある場合（後述）を除いて避けるべきだろう。

4.3. 位置と質問の交互作用効果が存在する場合

次に、質問と位置の交互作用効果がみられる場合について考えてみよう。例えば図4の場合、先にきいたときと後できいたときと合算したときとで、それぞれ結論が異なる。これは1947年に行われたギャラップによる世論調査(N=2634)の結果であり、やはり公平性の規範が作用した例として、Schuman & Ludwig (1983)によって紹介されたものである。企業による政治献金、労組による政治献金のそれぞれが「許されるか？」とたずねている。先にきいたときを比較すると、「はい（許される）」の割合は、企業（13%）<労組（23%）となり、後できいたときを比較すると企業（24%）>労組（16%）と逆転し、合算すると企業（19%）≒労組（20%）となってほぼ等しくなる。

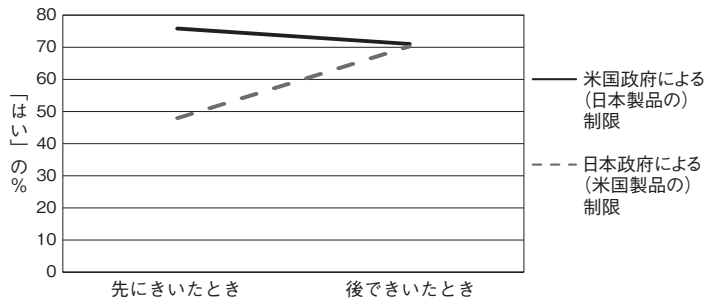
図4の場合、合算した結果から「企業献金が許されるとする人の割合は約20%で、労組のそれとほぼ等しい」と結論づけることは（一見、公平な判断のようだが）大きな問題をはらんでいる。第1に、この場合、先にきいたときと後できいたときで、正反対の結果が出ているのだから、単に「ケース数を倍にして信頼性を高める」という理由で、両者の結果を合算することは適当ではない。第2に、合算した結果には、後できいたときの結果が半分含まれており、順序効果を排除したことにはならない。こうした意味で、位置と質問の交互作用効果がみられる場合は、合算した結果で比較することは推奨できない。

図4 政治献金が許されるか？ (N=2634)



Schuman & Ludwig (1983) より筆者作成

図5 政府による輸入制限が許されるか? (N=419)



Schuman & Ludwig (1983) より筆者作成

順序効果を極力排除したいのなら、先にきいたときのみを比較して、労組による献金は23%の人が、企業による献金は13%の人が「許される」と答えた、と結論すべきである。双方の質問がそれぞれ単独で問われたなら（特に電話調査や面接調査の場合）、そのような結果になると考えられるからだ。

位置と質問の交互作用効果がある場合、順序の異なる調査票の回答を合算することは推奨できないが、先にきいたときの結果と、後できいたときの結果を並列して示すことが必要な場合もある。後できいたときの結果が意味を持つのは、回答者が複数の対象を比較評価すること自体に、何らかの積極的な意義を見いだせる場合であろう。例えば図5のようなケースでは、後できいたときの集計結果にも積極的な意義を見いだすことができそうである。

これも Schuman & Ludwig (1983) が公平性の規範の実例として紹介したもので、ミシガン大学のサーベイリサーチセンターが米国民を対象として1982年2月に行った電話調査の結果である (N=419)。日米貿易摩擦の最中に行われたこの調査では、米国政府による（日本製品の）輸入制限と日本政府による（米国製品の）輸入制限の是非が問われた。他との比較を前提としない場合、すなわち「先にきいたとき」には、当然ながら、回答者の自国政府である米国政府の制限を支持する意見が、日本政府の制限を支持する意見をはるかに上回る。しかし比較を前提とした「後できいたとき」には両者を支持する意見の%は拮抗する。すなわち、先にきいたときの回答には、自国政府の方を支持するという「素朴な」国民感情が色濃く反映されている。これに対して後できいたときの回答には、自国と他国の比較を前提とした「公平な」もしくは「理性的な」判断が反映されていると考えられる。自国政府の制限に懐疑的な意見や、相手政府の制限に賛成する意見が増えているからである。仮に、日本のマスコミが図5の先にきいたときの結果のみを、日本人向けに「米国の世論」として紹介・報道したとすればどうだろうか？後できいたときの「冷静な」意見を全く紹介しない報道は、いたずらに米国人への反感をおおる「偏向報道」という批判を受けかねないだろう。同じ意味合いで、図2の後できいたときの結果を全く無視してしまうことはためらわれる。こうした場合、先にきいたときの結果と後できいたときの結果を（合併せずに）並べて示し、双方の結果から結論を導く

のがよいだろう。

位置の効果、もしくは位置と質問の交互作用効果がみられる場合の結論の導き方についてまとめてみたい。まず、ふたつの調査票の結果を単純に合算して結論を導くことが許されるのは、位置の効果のみが存在し、かつ主たる問題関心が、ふたつの対象を相対的に比較することにある場合だけである。それ以外の場合には、順序効果を極力排除する観点から、先にきいたときの集計結果に基づいて結論を導くことが推奨される。ただし、回答者が複数の対象を比較評価すること自体に積極的意義があるならば、後できいたときの結果を、先にきいたときの結果と並列して示し、そこから何らかの結論を導くことも有益である。

5. 結 論

ここまで、様々な対象をある観点から評価させる場合に生じる質問の順序効果の2類型を示し、それぞれに対応した検定法、また効果がみられた場合の対象の比較評価の仕方を、2個の回答カテゴリーがついた2個の質問を行うケースにしぼって論じてきた。以下、結論を述べる。

過去の実証研究で示されてきた順序効果の様々な類型は、位置の効果と位置と質問の交互作用効果の2種類に分けられる。加算効果および減算効果は位置の効果に、同じく整合効果、対比効果、公平性の規範などは位置と質問の交互作用効果にあたる。位置の効果を検定する方法として、Mainland-Gart の検定が利用可能であり、同じく位置と質問の交互作用効果を検定する方法として、Hills-Armitage の検定が利用可能である。また、スプリット・バロット法による調査で諸対象を比較評価する際に、位置の効果「のみ」がみられ、主たる関心が各対象への評価を相対的に比較することにある場合は、異なる質問順序の回答を合算してケース数を倍加し、信頼性の高い結論を導く方法を推奨できる。しかしその他のケースでは、順序効果がみられたならば、特別な理由が無い限り各対象を先にきいたときの集計結果で評価を行うべきである。ただし、回答者が複数の対象を比較すること自体に、調査目的に沿った積極的な意義が見出せるならば、後できいたときの集計結果を、先にきいたときの結果と並列して示し、そこから何がしかの結論を導くことも考慮してよい。

本稿では、以上のようなことを、二者択一式のふたつの質問を行う場合にしぼって論じた。それ以外の質問形式でも、位置の効果と位置と質問の交互作用効果という類型は有効であり、それに対応した検定方法を示すことができると考えるが、詳細の検討は別の機会にゆずりたい。

〔注〕

(1) 研究例として、Schuman & Presser(1981), Schuman, Presser & Ludwig(1981), Schuman, Kal-

ton & Ludwig (1983), Schuman & Ludwig (1983), Bishop, Oldendick & Tuchfarber (1985), Sudman, Bradburn & Schwarz (1996), Tourangeau, Rips & Rasinski (2000), Moore (2002), Wilson, Moore, Mckay & Avery (2008) などがある。

- (2) スプリット・バロット法 (split-ballot method) とは、ワーディングや質問の順序などを変えた数種類の調査票を、ランダムに分割した複数の回答者集団に割り当て、それぞれの調査票への回答を比較する方法である。2個の質問の順序効果を観察したければ、順序を入れ替えた2種類の調査票を作り、それをランダムに分割されたふたつの回答者集団に割り当てる。
- (3) 順序効果を生み出す社会心理学的プロセスの詳細については、Sudman, Bradburn & Schwarz (1996), Tourangeau, Rips & Rasinski (2000)などを参照。
- (4) クロスオーバー法は、複数の治療法を同一の被験者に施すところから、複数の被験者集団のそれぞれに別々の治療法を施す方法よりも、サンプル数が少なくて済み、被験者の個体差の影響を小さくできることが利点とされる (Hills & Armitage, 1979; Dawson-Saunders & Trapp, 1994)。
- (5) 本稿の位置の効果に相当するのが、クロスオーバー法の文脈では、時期の効果 (period effect) と呼ばれるもので、治療 (薬剤投与など) の効果が時期によって一定方向に変化することをさす。そして位置と質問の交互作用効果に相当するのが、時期と治療の交互作用 (interaction between period and treatment) で、同じく治療の効果が時期によって異なるが、変化の方向性が治療の種類によって異なる場合をさす。
- (6) Mainland (1963) が先に提案したが、それに理論的な裏づけを与えているのは Gart (1969) である。
- (7) 「A→B」のとき、「はい (A)」は少なくなるが「はい (B)」は逆に多くなるから、「はい・はい」という組み合わせは、多くなるとも少なくなるともいえない。同様に「B→A」のときにも「はい (A)」は多くなるが「はい (B)」は少なくなるから「はい・はい」という回答は、多くなるとも少なくなるともいえない。「いいえ・いいえ」という回答についても、同じ理由によって、質問順序が「A→B」であろうと「B→A」であろうと、多くなるとも少なくなるともいえない。
- (8) 理由としては、①カイ二乗検定は数表を使えば電卓による手計算も可能だが、小標本の場合には信頼性に欠けること、②正確確率検定はカイ二乗検定よりも信頼性の高い方法だが、大標本の場合、計算が非常に煩雑になるので専用の統計ソフトが必要になることが挙げられる。
- (9) 「A→B」のとき、「はい (A)」は少なくなるが「いいえ (B)」は多くなるから、「はい・いいえ」という組み合わせは多くなるとも少なくなるともいえない。同じく「B→A」のとき、「はい (A)」は多くなるが「いいえ (B)」は少なくなるから、やはり「はい・いいえ」の組み合わせは多くも少なくもならない。「いいえ (A)・はい (B)」という組み合わせについても、同様の理由から、順序による違いはみられないはずである。
- (10) ここで見いだされた比率の差を検定する方法としては、通常のマクネマー検定 (McNemar test) ではなく、Mainland (1963) と Gart (1969) が、本稿で紹介した順序効果の検定法とあわせて示している治療効果の差の検定法に従うのがよい。具体的には、表4の「A→B」「B→A」の欄はそのままにして、「質問 A」のところを「前の質問」に、「質問 B」のところを「後の質問」に置き換えたクロス表を作成し、独立性の検定を行う。

〔参考文献〕

- Bishop, G. F., Oldendick, R. W., and Tuchfarber A. J., 1985, The importance of replicating a failure to replicate: Order effects on abortion items. *Public Opinion Quarterly*, 49: 105-114.
- Gart, J. J., 1969, An exact test for comparing matched proportions in crossover designs. *Biometrika*, 56: 75-80.
- Hills, M., and Armitage, P., 1979, The two-period cross-over clinical trial. *British Journal of Clinical*

- cal Pharmacology*, 8 : 7–20.
- Mainland, D., 1963, *Elementary Medical Statistics 2nd ed.* Saunders. (邦訳 : D. メインランド著, 柏木力訳, 増山元三郎監修, 1971, 『医学統計の基礎』岩波書店)
- Moore, D. W., 2002, Measuring new types of question-order effects, *Public Opinion Quarterly*, 66 : 80–91.
- Dawson-Saunders, B., and Trapp, R. G., 1994, *Basic & Clinical Biostatistics 2nd ed.* Appleton & Lange. (邦訳 : 森田茂穂監訳, 1997, 『医学統計データを読む : 医学・医療に必要な統計学活用法 第2版』メディカル・サイエンス・インターナショナル)
- Schuman, H., Kalton, G., and Ludwig, J., 1983, Context and contiguity in survey questionnaires. *Public Opinion Quarterly*, 47 : 112–115.
- Schuman, H., and Ludwig, J., 1983, The norm of even-handedness in surveys as in life. *American Sociological Review*, 48 : 112–120.
- Schuman, H., and Presser, S., 1981, *Questions & Answers in Attitude Surveys.* Academic Press.
- Schuman, H., Presser, S., and Ludwig, J., 1981, Context effects on survey responses to questions about abortion. *Public Opinion Quarterly*, 45 : 216–223.
- Schwarz, N., and Hippler, H.-J., 1995, Subsequent questions may influence answers to preceding questions in mail surveys, *Public Opinion Quarterly*, 59 : 93–97.
- Sudman, S., Bradburn, N. M., and Schwarz, N., 1996, *Thinking about Answers : The Application of Cognitive Processes to Survey Methodology.* Jossey-Bass.
- Tourangeau, R., Rips, L. J., and Rasinski, K., 2000, *The Psychology of Survey Response.* Cambridge.
- Wilson, D. C., Moore, D. W., McKay, P. F., and Avery, D. R., 2008, Affirmative action programs for women and minorities : Expressed support affected by question order. *Public Opinion Quarterly*, 72 : 514–522.

(やまぐち よう 現代社会学科)
2010年10月8日受理