

「総調査誤差」をめぐって

—ロバート M.グローヴス、ラース ライバーク論文

「総調査誤差 - 過去、現在、未来 -」を中心に—

大隅 昇（統計数理研究所）

鳩 真紀子（フリーランス翻訳家）

1. はじめに

周知のように国内外共に、回答率の低下とそれに伴う調査不能・無回答の増加など調査環境の悪化が指摘されて久しい。そして情報通信技術（ICT）の進歩による新たな調査用具や調査方式（とくに CATI やウェブ調査などのコンピュータ支援の調査方式）の登場による調査の多様化などで、調査方法論研究の内容がますます複雑になっている。

こうした調査実施環境にあって、あるいは調査方法論研究を進めるうえで、調査誤差（survey errors）の考察は重要な課題として注目を集めている。とくに近年は、“総調査誤差パラダイム”（TSE paradigm : total survey error paradigm）として語られるようになってきた。TSE の研究活動を促進する機構として総調査誤差国際ワークショップ（ITSEW : the International Total Survey Error Workshop）

（<http://niss.org/itsew>）が 2005 年から毎年開催されている [2012 年大会が 9 月 2 日～4 日に、オランダで開催された：[http://www.cbs.nl/en-GB/menu/organisatie/evenementen/itsew/information/default](http://www.cbs.nl/en-GB/menu/organisatie/evenementen/itsew/information/default.htm)

[htm](http://www.cbs.nl/en-GB/menu/organisatie/evenementen/itsew/information/default.htm)]。とくにここ数年、興味ある概説記事がいくつか登場しているので、これらを通して総調査誤差（以下、TSE と書く）についてわれわれの理解の及ぶ範囲で述べてみたい。

TSE に関わる類似の用語や考え方は古くから多々あった。実際に、関連記事を検索すると、過去の膨大な研究情報がヒットする。ここでは、とくに最近目にとまった以下の資料を対象として紹介したい。

1 つは、“*Public Opinion Quarterly*”（POQ）の 2010 年（74 巻、5 号）に TSE の特集号として掲載された 9 編の記事である。この特集号記事についてはオンラインで自由にダウンロードできる [<http://poq.oxfordjournals.org/content/74/5.toc>]。ここでは POQ の 9 編のうちの 2 編[表 1 の 1)、2)]に加えて数編の記事[表 1 の 3)～5)]を参考に、TSE の概要を述べる。

いずれも時宜を得た内容だが、POQ の 9 編のうち、上記の 2 編を除く 7 編の表題には、無回答誤差、測定誤差、調査方式効果、混合方式調査といった語句が登場しており、こうしたことが TSE のいまの関心課題ということがうかがわれる。1)、2)が総合報

表1 ここで取り上げた文献

1)	Paul P. Biemer: Total Survey Error (2010): Design, Implementation, and Evaluation, <i>POQ</i> , 74(5), 817-848.
2)	Robert M. Groves and Lars Lyberg (2010) : Total Survey Error - Past, Present, and Future, <i>POQ</i> , 74(5), 849-879.
3)	Tom W. Smith (2011): Refining the Total Survey Error Perspective, <i>International Journal of Public Opinion Research</i> , Vol. 23, No. 4, 464-484. (*) Groves ほかでは NORC リサーチメモ (2009 年) として引用、その後この論文となった。
4)	J. Bethlehem (2010): Selection Bias in Web Surveys, <i>International Statistical Review</i> , Vol. 78, No. 2, 161-188.
5)	J. Bethlehem and S. Biffignandi (2011): <i>Handbook of Web Surveys</i> , John Wiley and Sons.

告となっており、とくに 2) の Groves and Lyberg 論文は TSE 研究の歴史的経緯などが丁寧に紹介された内容となっている。ここではこの 2) を中心に、また必要に応じて他の文献も参照しながら、TSE を巡る話題を紹介する。なお、いずれの論文も紙数が多く内容が多岐にわたるので一部を抜粋紹介する。また 2) の Groves and Lyberg 論文については、完訳版を作成したので、関心のある読者はこれをお読みいただきたい (http://wordminer.comquest.co.jp/smr/02a_2010.html#new10)。

2. TSE の議論の要点

Groves and Lyberg 論文では、まず、総調査誤差パラダイムが歴史的にみても調査方法論の基礎的な概念として重要であること、しかしこのパラダイムは、調査現場で

みられる諸事象に照らすとさまざまな短所・欠点があること、しかし、こうした短所・欠点も総調査誤差の系統的な分解を行い類型化することで、諸統計量への誤差発生源の影響評価の可能性がみえてきたことが述べられている。そして今後考えられる総調査誤差の研究課題はなにか、といったことに触れている。本文は以下の 4 つの節に分かれるが、他の関連論文のキーワードを拾ってもう少し情報を細分化すると、およそ以下の①～⑥の項目を取り上げている。

1. 「総調査誤差」の概念の進展
(Evolution of the Concept of “Total Survey Error”)
2. 総調査誤差の枠組みの脆弱性
(Weaknesses of the Total Survey Error)

Framework)

3. 総調査誤差から見ることの長所
(Strengths of the Total Survey Error Perspective)
4. 将来の展望 (A Vision of the Future)

-
- ①TSE の“歴史的変遷”を再考し展望することの重要性、つまり調査誤差を調査方法論研究ではどう考えてきたか。
 - ② “TSE パラダイム”をどう考えるか、その概要と定義。概念的また知的枠組み(パラダイム)として捉えてきたが、これをより具体化できるのか、弱点(脆弱性)と利点を調べ、実現可能なことと、さらに研究が必要なこととがあること。
 - ③TSE の“類型化”、つまり TSE を誤差成分に分解し、それぞれの発生源と相互の関連性を実務レベルの事象と対応させて体系化すること。伝統的な見方は“標本誤差と非標本誤差に分けること”だが、これをさらに細分化し考察する必要があること。とくに、標本誤差だけではなく“非標本誤差の重要度”が増していること。
 - ④しかし同時に完全な類型化・分類も不可能である。なぜなら、新しい技術とくに情報通信技術の進歩や、方法論の発展、混合方式のデータ収集の使用といった戦略的な設計決定によって、新たな誤差構造が出現していること。
 - ⑤TSE をおもに回答者行動の視点から体

系化するための“モデル化”、とくに誤差伝播の因果モデル、測定誤差因果関係モデル、誤差発生源の相互作用(interaction)研究、認知心理学的アプローチなどがあること。

- ⑥調査品質と(投入)経費は表裏の関係にあるが、調査品質向上の観点から、適切な経費で優れた調査品質を確保するための留意点についての議論が重要なこと。このため、調査データの利用者側からみた“品質概念”とくに“使用適合性(fitness for use)”が欠かせないこと。これは製造における総合的品質管理(TQC)の考え方の援用であること、とくに、調査データの利用者観点から使用適合性のもとに“総合的調査品質”(total survey quality)として考えること[Biemer の論文の主眼点はここにあるようだ]。
- ⑦“混合方式”の導入に伴う調査設計の多様化に合わせ、これを多様な誤差成分の評価手段として利用する方法の模索。

標本調査では伝統的に調査誤差を「標本誤差」と「非標本誤差」に分けて考えてきた。この2つを併せて、たとえば浅井(1987)は「全誤差」(total error)とし、あらゆる誤差発生源で生起する誤差の合計、としている。統計学辞典(1989)ではこれを「総合誤差」として扱っている。これはここでの TSE とどう異なるのか。従来は「標本誤差」(母集団全体を調べずに標本抽出で

得た確率標本を用いることから生じる誤差」と「非標本誤差」（調査過程で生じるその他のさまざまな誤差）に分けて考えることで対応できた。前者は調査設計・実施者側に管理の主導権があり、誤差が比較的容易に把握できる。後者の非標本誤差については、回答者、調査方式、調査員など管理がむずかしい要素の影響下にあり、加えて調査環境がより複雑になったこと、さまざまな調査方式を使い分ける必要性（例：混合方式の適用）が増したことから、この非標本誤差の影響評価が無視できなくなり、調査誤差全体を総合的に精査することが必要になった、ということだろう。

紙幅の都合もあるので、本稿ではとくに「総調査誤差の主な過去研究」そして「総調査誤差の分解と分類の重要性」、その周辺に的を絞って説明する（上記の①～③あたり）。

3. TSEとは 一歴史的変遷と類型化、数理的展開の経緯一

Groves らは“TSE パラダイムとは、標本調査の統計量に関する統計的な誤差特性を説明する概念的な枠組みのことである”という。また Biemer は、“TSE パラダイムとは、経費制約下でデータ品質を最大化することで、調査を最適化するための理論的枠組みを提供するものである”としている。さらに Smith は、“TSE とは、調査の測定がうまく機能しない方向に進んだと

きの多数の方法・手段の総体である”とし、狭義には“目標母集団の実際の値（真値）と、調査で得た測定値との差”であるといえるが、これでは十分でなく、“測定に係する他の変動まで考慮してより拡張的に総調査測定変動（TSMV：total survey measurement variation）として包括的に考えるべき”としている。

これらの記述はもっともなことであると同時に、やや概念的であり、分かりやすいとはいえない。そこで、筆者なりに読み替えてみると、これは、理想（つまり、確率標本から必要とする母集団特性値の確かな推定量が得られること、あるいはそうあって欲しいこと）と、現実（実際の調査実施で標本の測定で生じる変動 [variance：偶然誤差] や偏り [bias：系統誤差]）との間に生じるズレ（gap）を容認し、これがどこまで読み取れるかを知ることである、といえるだろう。さまざまな誤差発生源（error sources）があることを前提に、（母集団推論のための）調査推定量の統計的な特性を記述するための総合的かつ概念的枠組み（パラダイム）ということである。

しかし現状は、このパラダイムの体系化や“厳密な意味での定義”があるわけではない、ということも共通認識である。このことが、後述するように“さまざまな TSE の類型化”と、それを巡る多様な議論となることの原因でもある。明らかなことは、“TSE がうまく定義できたとして、これを

最小化するような調査設計を目標とする”ということである。つまり“TSEを最小化する、あるいは低減することが、調査品質の改善につながる”という総合的なアプローチをいう。

要は、Groves のいう“母集団を特定し、標本を記述し、標本内の回答者群にアクセスし、測定の対象となる構成概念 (constructs) の検証を可能にし、測定に対する回答を得て、なんらかの指定された母集団パラメータ (母数) を推定するデータを要約する、という調査設計の成分すべてに注意を払うことが、調査誤差という言葉に固有のこととしてある”ということである。

ところで、前述のように、伝統的に調査誤差を「標本誤差」と「非標本誤差」の2つに分けて考えてきた。実際、多くの過去研究でも基本はそう考えている。しかし Groves らは、これだけでは十分ではないこと、過去の先達がさまざまな試みを行ってきたことを、歴史の時間軸に沿ってうまく説明している。これはまた“TSE の類型化”をどう考えるか、ということにつながる。

Groves は、TSE の概念の進展・発展には、過去の偉大な研究者たちが、いずれも調査誤差を重要な課題とし、誤差発生源と誤差成分の分類に注目してきたことを挙げている。あの Deming はすでに 1944 年の論文で、“「調査は考えられる誤差発生源か

ら影響を受けるとして、どのようにして有用な結果を生み出すことができるか？」という問題”に取り組み、実際に「調査の有用性に影響を及ぼす 13 の要因」を整理分類してみせたが、ここでは総調査誤差という概念にまでは至っていないという。つまり、標本誤差の重要性・有効性を強調しこれをひろめるために、(ある意味で) 逆説的に非標本誤差にも注目すべき、と考えていたようだ と推測している。このほか、Hansen ほか (1953)、Cochran (1953) など多くの先達がさまざまな研究を展開したことが要領よく述べられているので、これらを表 2 のように一覧にしてみた。

調査における誤差発生源を列挙し分類することは、標本誤差発生源を示すだけではなく、より細かい吟味を行うこと、とくに測定誤差成分の検証の重要度を示すことである。ここで Groves らは、TSE の推定からみた統計的方法論の研究の進展についても述べている。ここでは、回答誤差に影響する変動のうち、とくに測定誤差の変動に注目し、真値とその推定値、その偏りといった概念の導入が必要なことを、著者の視点から歴史的にみて重要とされた、Hansen 他 (1961、1964)、Fellegi (1964)、Bailar と Dalenius (1969)、Biemer と Stokes (1991)、Sarıs と Andrews (1991) などの研究を引用している。

要は、いかに偏りなくまた変動の小さな母集団特性の真値の推定値を得るかという

表 2 総調査誤差の分類要約

提唱者と出处・出典	主な誤差発生源とした事項, 考え方
Deming (1944) の分類	誤差発生源を以下の 13 項目に大分類, さらにこれらを細分類した. また, ここでは TSE と名付けるまでには至っていない. 1. 回答 (内) の変動 2. さまざまな種類のさまざまなレベルの調査方式間の差異 3. 調査員に起因する偏りと変動 (調査員変動) 4. 調査依頼者・調査主体に起因の偏り 5. 調査票設計と集計計画における不完全性 (欠陥) 6. 集計結果が得られる前に母集団内に生じる変化 7. 無回答 (脱落を含む) から起こる偏り 8. 報告の遅れから起こる偏り 9. 調査目的にとって, 代表性のない日付や該当期間の選択に起因の偏り 10. 代表性のない回答者選出から生じる偏り 11. 標本誤差および偏り 12. 処理誤差 (例: コーディング, エディティング, 算出, 集計, 転記, 併合など) 13. (調査結果の) 説明・解釈における誤差
Hansen ほか (1953) 「標本調査の手法と理論」 “ <i>Sample Survey Methods and Theory</i> ”	・標本誤差の特性研究に重きをおくが, 非標本誤差を無視していたわけではない. ・主要な誤差発生源として以下を挙げている. 「調査計画あるいは定義の欠陥に起因する誤差」「回答誤差」「カバレッジにおける誤差」「分類における誤差 (層別の誤差)」「エディティングの誤差」「公表・公開の誤差」を挙げている. *ここで「エディティングの誤差」は Deming のいう「処理誤差」に近い.
Cochran (1953) 標本抽出の技術 “ <i>Sampling Techniques</i> ”の最終章 [注: 訳本「サンプリングの理論と方法」, 東京図書]	・無回答誤差, さまざまな測定誤差 ・エディティング, コーディング, および集計する段階で導入される誤差, などに触れている.
Kish (1965) (★) 「調査標本抽出法」 “ <i>Survey Sampling</i> ”	・「偏りと無回答誤差」と名付けた章で説明. ・完全に「偏り」に焦点を当てた議論となっている (図がある). ・観測誤差と非観測誤差とに分離するはじめての試み. ・標本誤差が本質的に非観測誤差であることには気付いていない.

Dalenius (1974) 「調査における誤差」プロジェクト (1967年～1974年)	・「総調査設計」(total survey design) という考え方を提案. ・定義された目標に最も近い総調査設計過程が適した設計であると定義. ・「総」とは重要な誤差発生源のすべてを総合的に制御・管理すべきことを意味している. ・「総」には調査業務の設計と同様に制御, 評価のシステムの総体が含まれること.
Anderson, Kasper, Frankel(1979) 「総調査誤差」 “Total Survey Error”, Jossey-Bass	総調査誤差の分解を提示. 変動(偶然誤差)と偏り(系統誤差), 標本抽出と非標本抽出, 観測と非観測と主な区分分類を強調.
Groves (1989) (★) 「調査誤差と調査経費」 “Survey Errors and Survey Costs”, New York Wiley.	・調査誤差の数理的定式化, 心理測定学や経済統計学から得た知識と調査誤差との結合化. ・「平均二乗誤差」を重視し「変動(分散): 偶然誤差」と「偏り: 系統誤差の二乗」の和として表記. ・「非観測誤差」(例: カバレッジ誤差, 標本誤差, 調査不能・項目無回答)と「観測誤差」[ある調査変数についての報告値と真値(true value)あるいは根拠のある値(underlying value)との差]とに分類. ・誤差構造の概念に, 心理統計学的, 調査統計学を併せた見方を導入. ・主要メッセージは, (a) 心理測定学の信頼性(reliability)の考え方の導入 (b) 妥当性(validity)の考え方と調査誤差の概念との対応を検証.
Lessler and Kalsbeek (1992) 「調査における非標本抽出誤差」 “Nonsampling Error in Surveys”	・枠誤差, 標本誤差, 無回答誤差, および測定誤差について議論. ・「総調査誤差」よりも「総調査設計」の考え方を強調. ・Daleniusの考え方の影響を受けている.
Biemer and Lyberg (2003) (★) 「調査品質入門」 “Introduction to Survey Quality”	・まず, 標本誤差と非標本誤差とに区分することの重要性を指摘. ・非標本誤差を設計仕様誤差, 枠誤差, 無回答誤差, 測定誤差, 処理誤差とし, この5つの非標本誤差の主な発生源をリスト化. ・処理の品質(process quality)と総調査誤差の考え方の統合を初めて試みる. ・心理測定学の理論的妥当性の考え方に類似. ・品質管理的な思想の導入(Deming, Juran, Taguchiなど). ・DemingやJuranのいう「使用適合性」(fitness for use), 「有用性」(usefulness)と「目的適合性」(fitness for purpose)といった“調査データの利用者側”の視点からのアプローチの重要性を強調. ・具体的にパラデータ(プロセス・データ), 管理図, 応答設計(responsive design)を導入してみせた.

Groves ほか (2004, 2009) (★) 「調査方法論」 “<i>Survey Methodology</i>” [注:訳本「調査法ハンドブック」 (2011), 朝倉書店]	・総調査誤差の枠組みを中心に編成された書。 ・調査設計, 収集, および推定の段階を誤差発生源に結び付けようと試みた。 ・調査実施過程のどの段階で, どのような誤差発生源を調べ, 低減を図るかを提示。[訳注:この考え方の流れは(図2)を示す]
Herbert Weisberg (2005) 「総調査誤差アプローチ」 “<i>The Total Survey Error Approach</i>”	・生起するさまざまな種類の誤差を中心にして調査研究を行う方法を提示。 ・倫理規定や適時性といった制約事項についても議論。 ・「処理誤差」について異なる見方をした。「調査管理上の問題点」(Survey Administrative Issues)として分類。
Smith (2009, 2011) の論文 (★) Refining the Total Survey Error Perspective, <i>International Journal of Public Opinion Research</i>. (*) Groves は 2009 年発表の NORC 研究レポートを引用している。その後, 2011 年に論文となった。	・国際調査にまで総調査誤差のパラダイムを拡大し, 誤差発生源に「比較誤差」(comparison error)を追加。 ・パネル調査に特有の「調整誤差」(conditioning error)も追加。 ・類型化はできる限り詳細かつ包括的でなければならないと主張。 ・TSE を (a) ランダム変動 (偶然誤差) と (b) 偏り (バイアス) あるいは系統的誤差に分ける。TSE はこれら 2 つの成分の合成。 ・TSE の“Total”は完全であること (completeness) を保証するとの含意 (現実の達成は困難)。その理由は, 第 1 に, 調査は多数の成分からなる装置であるからであり, また, 第 2 に, 誤差を分類するさまざまな方法があるからとした。 ・「標本誤差」「非標本誤差」と分けてさらに細分類すること。
Bethlehem の分類 (2010, 2011) の考え方 (★)	・情報通信技術 (ICT) を多用するコンピュータ支援型の調査方式, たとえばウェブ調査や CATI などでは, 総調査誤差をよりきめ細かく分けて考える必要があること。 ・従来の分類を総括的に整理した, より一般的な分類を提案。[訳注:図1にここでの分類図を示した] ・標本誤差と非標本誤差に分ける, ここまでは同じ。 ・標本誤差を「推定誤差」「調査設計誤差」に分け, 非標本誤差を「観測誤差」「非観測誤差」に分ける。 ・観測誤差を「オーバーカバレッジ」「測定誤差」「処理誤差」発生源に分ける。 ・非観測誤差を「アンダーカバレッジ」「無回答誤差」に分ける。 ・パラデータの活用で回答者行動分析と誤差 (とくに測定誤差, 無回答誤差) を評価。

*Groves and Lyberg 論文、その他の論文から要約した。(★)を付けたものには TSE の分類図 (樹木図) も添えられている。

ことであるが、一般的にはこの真値を得られるとは限らない。そこで、真値に替わる代用としての外部情報源 (external sources)、たとえば既存の確かな記録、優れた基準 (gold standard) を用い、相互貫入的な設計と再面接の設計の利用、回答の心理測定的なアプローチ (構成概念、回答と測定誤差のモデル化) を適用し、よく知られた推定値の平均 2 乗誤差 (MSE: mean squared error) を変動 (偶然誤差) と偏り (系統誤差) の 2 乗の和とする関係を考察すること、さらに、ある推定値の MSE を $\text{抽出分散} + \text{回答変動} + \text{回答と抽出偏差の共分散} + \text{偏りの 2 乗}$ などに分解することの重要性 (Hansen 他) が紹介される。

上に述べたような、調査方法論研究で日常的に用いられる概念が登場するのだが、記述情報がかなり圧縮化されており、これらの全体は引用の多くの論文や著作物をフォローアップする必要があるだろう [注: Groves (1987), *Survey Errors and Survey Costs* などに詳しい]。

4. TSE の分類と調査実施過程における誤差発生源

ここで、表 1 に要約した Groves and Lyberg らの情報に共通したこととして以下がある。

- TSE を巡る議論はさまざまであるから、当然のことではあるが、分類は一意には

記述できないことがある [注: 各誤差を皆版的にわけることへの懸念もある (Couper 氏私信)]。一方、誤差成分をなるべく系統分類学的に「樹木図」で示そうとする試みがある。

- TSE を「標本誤差」「非標本誤差」に分け、さらに非標本誤差を「観測誤差」と「非観測誤差」に分ける、ここまでは、ほぼ全体に共通している。
- ただし、はじめに「観測誤差」と「非観測誤差」と分ければよいとする意見、そして (前述の) MSE を「変動」と「偏りの二乗」の和として考察することを重視すべきとする意見もある (Groves and Lyberg, Biemer など)。
- 前述のように共通認識として、非標本誤差の重要度が高まり、結果としてこれをどう細分類するかに議論が集まる。
- 一方、標本誤差については、標本抽出論を始めとする過去研究の実績、成果が、正しく反映されたいうえでの議論としている (母集団の設定、適切な標本抽出、確率標本を前提とする確率的アプローチは当然のこととする)。
- 情報通信技術に依存する新たな調査方式の登場で、誤差構造が複雑となったことを共通認識として、従来は想定する必要のなかったあらたな誤差発生源をどう評価するか、たとえばパラデータ (プロセス・データ) を用いた回答者行動の分析、さらに混合方式などの適用での誤差発生

などをどう評価するか。

- ・他の研究分野、たとえば、心理測定学・認知心理学、経済統計学、品質管理とその周辺技術などの知識を併せたアプローチを、誤差評価やモデル化に援用するという見方がある。

ここでは、上の諸事項が比較的うまく反映された、典型的な誤差分類の例として、Bethlehem (2010) を挙げよう (図 1)。ここで点線囲みの<A>の部分は、どの研究にも (ほぼ) 共通する区分であり、一方、の部分については、いろいろと議論がなされてきた (いる) 部分といえる。また、カバレッジ誤差を「アンダーカバレッジ (調査漏れ、過小カバレッジ)」「オーバーカバレッジ (過大カバレッジ)」の2つにわけ、それぞれを観測誤差と非観測誤差の下に置いていることが特徴である。

Bethlehem によると、コンピュータ支援型の調査方式、とくに CATI やウェブ調査では、この2種を分けて考えることが重要だとしている。

ところで Groves らは、調査実施過程と誤差発生源との関係を“Survey Methodology” (2004、2009) でうまくまとめている。この書自体が、そもそも TSE パラダイムの観点から調査方法論をどう考えるかを述べたテキストであり、調査の実状に合った分かりやすい説明となっている。Groves and Lyberg 論文でもこの書から引用しているが、ここではこの書の翻訳版「調査法ハンドブック」(2011) からの引用として図 2 を用意した。ここでは、カバレッジ誤差、標本誤差、無回答誤差、調整誤差、測定誤差、処理誤差が、それぞれ調査実施過程のどの段階で考慮すべきかがわかりやすく整理されている。

図 1 総調査誤差の分類の例

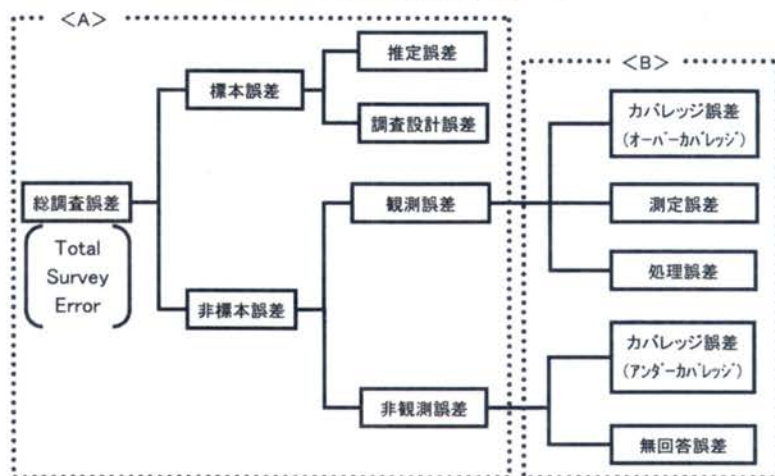
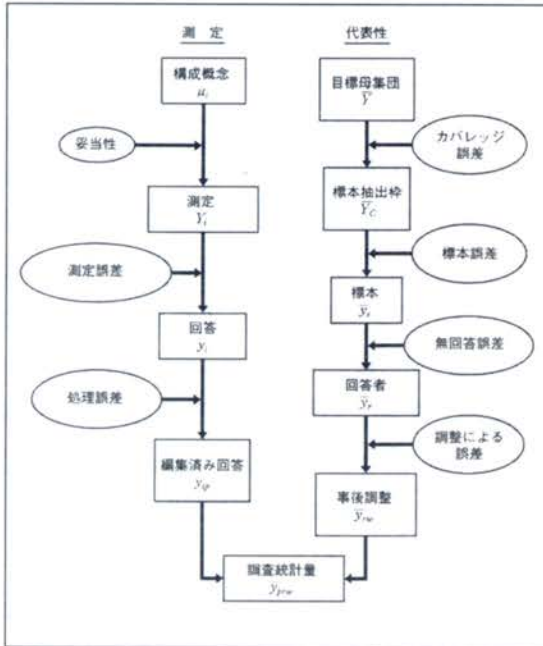


図2 調査実施過程と誤差発生源の関係



ではこれらの各誤差は、具体的にどのような内容をいうのだろうか。これについては、Groves らは、個々の誤差発生源に合わせて随所で例を挙げているのだが、やや分かりにくい。また、いままで述べたことから明かなように意味解釈が未だ不確定な部分もある。そこでここでは1つの試みとして、Groves、Biemer、Smith、Bethlehem 他を参考に、「非標本誤差」の各誤差成分を表3のように整理してみた。表にみるように、非標本誤差の誤差発生源は多岐にわたることが分かる。ここに挙げた諸事項は、調査研究のエキスパートの読者にとっては、いわずもがなの事柄であるが、こうして体系的に分けてあらためて見直すことは意味があるだろう。また、他分野の読者も意識

して、このような要約情報を作ってみた。

5. Groves and Lyberg が考える今後の課題

調査における誤差をどう考えるか、総調査誤差 (TSE) とは何か、これについての欧米研究のごく一端をかなり圧縮した形で、またわれわれの主観・解釈も交えて紹介してみた。Groves and Lyberg 論文では、この他、総調査誤差パラダイムの弱点、長所、それに将来の展望について述べている。要は、各論文とも今後の調査方法論研究の軸足が、調査誤差をより広い範囲にわたって考察することの必要性を訴えている。とくに、今後の展望として以下のようなことを指摘している。

- ①今後の総調査誤差パラダイムの精緻化に向けて今世紀 (21 世紀) の調査方法論研究の方向・特徴として、i) 調査の適用分野の拡大が進んでいること、ii) 研究対象とする母集団は測定に適した要求内容に応じて変化すること、iii) 調査以外の外部要因 (例：ウェブ、電子調査システム、電子化ビデオ情報など) の絶え間ない変化による測定環境の変化・可能性などがあること。
- ②過去に提案されてきた、さまざまな調査誤差の因果モデルあるいはこれに類似のモデルの改善。とくに、調査票設計 (質問文構造、ワーディング、回答選択肢、

表3 「非標本誤差」の要約

誤差の種類	考えられる誤差の内容・簡単な例
調査設計誤差 (specification error)	● 調査企画・仕様策定時の調査概念・構想（構成概念）の食い違いや不適切な調査目的・設計に起因の誤差。 ・ 標本抽出枠（枠）の不備で設計通りに抽出ができない。 ・ PC 非保有者にまで PC の性能評価を質問する。
カバレッジ誤差 (coverage error)	● 枠から脱落、逆に誤って含めてしまう、重複など。 ・ 想定リストと実際のリストのズレ（例：電話調査のシステム上の RDD と実際の番号）。 ・ オーバーカバレッジ（過大カバレッジ）とアンダーカバレッジ（過小カバレッジ、調査漏れ）ほか無数のカバレッジの影響。 ・ 名簿の陳腐化、パネルの疲労など。 ・ 枠の誤差に関連すること。
無回答誤差 (nonresponse error)	● 選ばれた適格な調査対象者から調査の要求する回答が得られない、あるいは回答ができないことで生じる誤差の総称。調査実施側では制御ができない誤差の1つ。 ・ 接触不能、調査不能・回収不能、一部無回答（項目無回答）、欠測。 ・ 回答拒否（例：微妙な質問に対して、回答者が回答を提供しない）。 ・ 回答不能：回答する意思があるのに回答ができない（例：ウェブ調査で回答送信エラーや入力操作不具合；聴力障害者の聴取）。 ・ 自由回答質問での不完全情報。回答内容が質問文の主旨・意図に應えるほど十分な記述量がないとき（例：職業記入でポストコーディング時に該当コードが見つからず欠測や不明扱いとなる）。
測定誤差 (measurement error)	● 回答者が、真の回答とは異なる回答を提供したときに生じる誤差。 ・ 用いる調査方式、調査用具、調査員変動などで生じる誤差の総称。 ・ 測定過程で生じるほとんどの事象。 ・ 調査方式の差違の影響、面接員・調査員、調査用具（調査票、提示カード、PC など）の影響。 ・ 回答選択順の影響（初頭効果、新近性効果）、黙従傾向、社会的望ましきなど。 ・ 一部を観測誤差と括ることもある。観測誤差は観測・測定に関係する誤差。非観測誤差は観測・測定には直接は関係しない誤差。 ・ 回答者が質問文の内容を理解できない、あるいは意図的に自分の意見と異なる選択肢を選ぶなど。 ・ 質問文の影響：ワーディング、微妙な質問、文脈効果・順序効果などで回答に偏りを生じる。
調整誤差 (adjustment error)	● 調査の事後調整（post adjustment）で生じる誤差。 ・ 重みづけ（加重補正）や補定（imputation）で生じる誤差。
処理誤差 (processing error)	● 調査後に生じる加工・処理上のさまざまな事象に起因の誤差。 ・ エディティング、データ入出力、作表処理、…。 ・ コーディング、ポストコーディング、…。

*Biemer and Lyberg (2003)、Groves ほか (2004, 2009)、Bethlehem (2010, 2011) などを参考に再編集した。

回答者行動)と調査員変動の関係やメカニズムをどうモデル化するか。

- ③今後の研究課題で重要なことは、調査行動や回答者行動を調べる社会科学や行動科学の知見、認知心理学的アプローチなど、異分野研究の成果を取り入れた統合的なモデルの構築を行う必要があること。
- ④誤差発生源は個別の吟味検証も必要だが、複数の誤差発生源間の差違や相関(共分散)が無視できない。よって、それぞれの誤差発生源が互いにどのように影響しあうか、いわゆる相互作用の考察が重要であること。
- ⑤変動と偏りが、互いにどのように関連しあうのかをさらに理解すること、両者を併せて検証する研究の重要性。
- ⑥社会科学分野などで発展した因果推論モデルに、調査方法論における調査誤差推定のモデルを加味した、調査品質の評価モデルの仕様と設計の検討の必要性。
- ⑦品質の測定および品質確保と投入経費の関係は重要。品質保証あるいは品質測定の妥当な経費の検討、限られた資源の配分と調査品質の関係の検証。とくに、(品質管理的に)調査結果の利用者側に立った品質評価における「使用適合性」(fitness for use)の考え方を総調査誤差パラダイムと合わせ統合的に考察すること。
- ⑧複数の枠、複数の調査方式(混合方式: mixed-mode)あるいは多重方式

(multiple-mode)、多段設計の調査設計への移行はすでに進んでいるが(例:電話調査と面接調査の混用、自記式と調査員方式の混用、回答率向上のためにウェブ調査、郵送調査、電話調査など複数の調査方式を併用すること)、これと総調査誤差との関係をどう考え発展させるか。

6. おわりに

Groves and Lyberg 論文を中心に、いくつかの論文から、TSE に関する研究の変遷、考え方、誤差分類、現状、今後の課題、といったことを抜粋して紹介した。各論文の著者らは、ほとんどが官庁統計あるいはそれに関連のある機関・機構(センサス局をはじめ各省庁の統計部門など)とも密接な関係にあり、また、マーケティング・リサーチや他分野との共同研究にも積極的である。つまりは産官学の共同体制が整ったなかでの研究成果である[参考文献 3]。研究論文もこうした土壌の中から生まれた成果であろう。

日本国内での調査方法論の分野の過去研究を振り返ってみたとき、どちらかというところ率標本を前提とする調査設計の標本誤差に重きをおいたアプローチが多かったように感じる(もちろん重要な要素である)。また、われわれが調査を実施するうえで、回答者行動に注目した調査不能・無回答、調査票設計(質問文作成、ワーディング、文脈効果、…)、調査員変動など非標本誤差

にかかわる事象の観察や実証研究に無関心であったわけではない。しかし、欧米研究のように体系的かつ組織的に進めてきたかという、かすかな疑問がなくもない。いずれの文献も指摘するように、情報通信技術などの急速な進歩で調査環境は確実に変容している。とくに調査方式の主流となりつつあるコンピュータ支援の調査方式（CATI、IVR（自動応答装置）やウェブ調査）と従来方式（面接調査、郵送調査など）の比較研究は不可欠であるが、いまだ十分とはいえない。

ここで取り上げた論文では、主として確率標本を基本とした確率的アプローチを前提とした議論であるが、商業ベースではむしろ、定性型調査（フォーカス・グループ、アクティブ・インタビュー、CLT（会場テスト））や便宜的標本を用いる場合も多用されている。こうした調査方式を非確率的アプローチであるから不確か、信頼できないと批判するだけではなく、前向きに研究を進めることも必要である。すべてに通底することは「調査の品質向上、改善、評価をどう考えるか」である。

追記

いくつかの論文をもとに、近年、欧米の調査方法論研究で関心を集めている「総調査誤差」に関連の話題の紹介を試みた。編集ご担当のご厚意により、かなりの紙数をいただいたのだが、読み返してみると、内容がいささか多岐にわたったきらいもある。加えて、おそらくは国内ではあまり馴染みのない用語・語句を多数用いることにもなったが、これらに細かい説明を付けることにも限界があった。結果として、やや読みにくい内容となったことはお許しいただきたい。欧米の研究者の調査の考え方、研究目的や関心の焦点は日本国内とは異なるであろうし、調査環境面の基盤の違いもある。それでもなお、欧米の調査方法論研究は、研究者層が厚く、実務・理論の両面ともに着実に進展していることを知っておくことも意味があるだろう。

【参考文献】

- 1) 浅井晃（1987、1992）：「調査の技術」、日科技連出版社
- 2) 竹内啓他（編）（1989）：「統計学辞典」、東洋経済新報社
- 3) 大隅昇（監訳）（2011）：「調査法ハンドブック」、朝倉書店〔Groves 他（2004）、“Survey Methodology”（John Wiley & Sons）の全訳版〕