

# ウェブ調査において無効回答選択肢を含めることの影響<sup>1)</sup>

吉村 治正\*・正司 哲朗\*\*

Effect of “Don’t Know” Option in Web Survey

Harumasa YOSHIMURA and Tetsuo SHOJI

## 要 旨

本稿では、ウェブ調査において「わからない」という選択肢を与えることで回答にどのような影響が現れるかを検討した。訪問面接調査と異なり、ウェブ調査では「わからない」「なんともいえない」「知らない」といった、想定された尺度上にない回答を許容することができない。項目非回答として集計される無効回答は、あらかじめ選択肢として与えてやらない限り、有効回答の中にランダムに紛れ込んでしまう危険が高い。だが無効回答を選択肢として与えることは、不正確な情報の排除が可能になるという期待と同時に、回答者の最小限化回答を誘発し情報の精度が損なわれるという危惧をも生む。本稿では、業者委託によるウェブ調査実験を通じ、この二つの可能性について検討した。結果としては、無効回答を選択肢として与えてもデータの正確さが向上したとも言い切れず、また最小限化回答を誘発したというエビデンスも得られなかった。

キーワード：ウェブ調査、項目非回答、測定誤差、最小限化回答

## I 問題の所在

「わからない (don't know)」・「場合による (it depends)」・「どちらともいえない (can't decide)」・「なんともいえない (no opinion)」・無言や空欄など、社会調査 (social survey) を実施する際に必ず直面するのが、想定外の回答、いわゆる無効回答の発生である。回答選択肢にない回答、調査者が設定した尺度上にない回答は、欠損すなわち項目非回答として扱われ、集計および分析の段階で除外される。これは、歯止めのかからない回答率の低下に苦悩する調査者にとって、きわめて切実で深刻な問題である。現状で回答率の低下を抑制する有効な手段は見つかっておらず、よほど経験豊富な調査者がそのノウハウを駆使しても、回答率は50%を超えるのがやっとというのが現実である。ところが、協力してくれた対象者が個別の調査項目に回答しないこと、あるいは無効とみなされる回答の発生率は、いとも容易に10%を超える。項目によっては25%を超えることもあるという (Weisberg 2005)。たとえば内閣府が平成31年2月に実施した「社会意識に関する世論調査」では、「国民の間に国を愛するという気持ちをもっと育てる必要があ

2019年9月10日受理 \*社会学部総合社会学科教授・\*\*社会学部総合社会学科准教授

と思うか」という質問で、実に13.9%の回答者が「わからない」と答えている<sup>2)</sup>。わずか数パーセントの回答率向上のために悪戦苦闘を繰り返す調査者にとって、これほど徒労感を感じることはない。なんとかして無効回答の発生を抑えたい。それは、調査者の切実な願いである。

しかしながら、社会学へのウェブ調査の導入を見すえると、無効回答の発生は多少とも複雑な様相を見せる。社会学者はこれまで、社会調査の方法として訪問面接法をスタンダードとしてきた。ところが訪問面接による社会調査の実施は、熟練調査員の不足や対象者との接触率の低下、拒否率の上昇など、次第に困難さを増し、この20年ほどの間に郵送調査へ、さらにウェブ調査への移行を余儀なくされつつある。しかしながら、こうした調査方法の移行を進めるには、これに先立って技術的問題点の解明と適切な対処の策定が必要となる。具体的な例として内閣府世論調査をあげる。内閣府世論調査は社会学にとどまらず日本の社会科学全体にきわめて重要な情報を提供してきた。そのうちの質問の一つ、「国民生活に関する世論調査」(平成30年2月実施)で、社会学者にはよく知られた以下のような質問が使われている<sup>3)</sup>。

Q9 お宅の生活の程度は、世間一般からみて、どうですか。この中から1つお答えください。

- (ア) 上
- (イ) 中の上
- (ウ) 中の中
- (エ) 中の下
- (オ) 下

総務庁のホームページ上では、この5つの選択肢が調査対象者に提示され、そしてその回答として、上が1.6%、中の上が13.6%、中の中が58.0%、中の下が21.1%、下が4.2%となっており、これに加えて「わからない」が1.4%と記載されている。この時、「わからない」には選択肢の番号がない。つまり回答者に対して提示されたのは、(ア)から(オ)の5つの選択肢であり、「わからない」は選択肢としては与えられていないが回答者が口頭で答えた内容を集計したもの、ということになる。

選択肢で与えていない回答が生じるというのは、訪問面接ではよく起こるし、またその発生もある程度まで想定されている。それ故に、この内閣府の調査でも選択肢にはないにもかかわらず集計されている。自記式の、たとえば郵送調査の場合は、回答者が自由記述でその旨を表記した場合に、および何も書かなかった(空欄)場合に、選択肢にない回答が発生したことがわかり、無効回答として集計される。ところが、ウェブ調査の場合は欄外の自由記述はできないし、選択肢を選ばないで次の質問に進むこともできない。ウェブ調査はルーティング構造、つまり前の質問での回答内容によって次に進む画面が異なることを強く意識して設計されている。したがって、質問に回答しないと次に進むべき画面が特定できなくなってしまうため、無回答をブロックする、すなわち与えられた選択肢のどれかを選ばない限り次の画面へ進めないようにデザインされている。これは実際に民間のウェブ調査業者に委託交渉してみるとわかるが、ほぼ全ての調査業者で無回答のブロックがデフォルトになっており、その変更を依頼しても認めない業者の方が多い。

これはウェブ調査の技術的特性によるものだが、これが無効回答の発生にきわめて大きな問題をもたらす。すなわち、ウェブによる社会調査に回答していた個人が、ある質問で自己の考えに相当する選択肢が存在しないことに気づいたとき、この回答者は、自己の意見や考えに対応しなくても、提示されている選択肢の中から何かを選択しなければならなくなる。この場合、この回答者がどのような選択肢を選択するかの予想がつかず、その結果データの信憑性に疑問が生じてしまうのである。これを回避するためには、一つ一つの質問の画面で「わからない」「知らない」などといった、いわゆる無効回答を選択肢として与える必要がある。ところが無効回答を選択肢として与えると、回答者に「安直な」回答を促すことになるという指摘がある（Krosnick et al. 2002）。この場合も、やはりデータの信憑性を低下させることになる。ウェブ調査への移行を見送るうえで、無効回答を選択肢として与えることの是非を検討しておく必要が生じるのは、こうした事情による。

## II 無効回答もしくは項目非回答の発生に関する既存研究

社会調査の回答率低下が世界中で問題となっているにもかかわらず、個別の項目の回答率が問題になることは稀である。だが、項目非回答の発生率は、一部の研究者の間で長く問題視されてきた。1966年、ファーバーは郵送調査の回答調査票を見直して、回答者の2/3になんらかの無回答項目（空欄）が存在することに気づいた（Ferber 1966）。調査への協力を引き受けてくれた人の中に、調査者が求める情報を全て提供している人は、わずかに1/3しかいないというのである。また、サッドマンとブラッドバーンは、主要な社会調査を調べ、当時の一般的な世論調査での「わからない」・「なんともいえない」といった、いわゆる無効回答の発生率をおおむね10%程度と見積もっている（Sudman & Bradburn 1974）。しかも、無回答あるいは無効回答は回答者の属性と密接に関係している。したがって何らかの対処策を講じない限り、集計や分析に偏りをもたらす危険がある。これらの事例報告は、無効回答とそれがデータ化された項目非回答の問題がいかに深刻であるかという警鐘をならすことになった。

これ以降、無効回答の抑制は社会調査の重要課題の一つとなった（Dale 2006; Lessler & Kalsbeek 1992; Mason et al. 2002）。スマスは、訪問面接調査における項目非回答の大半は調査員のミス、つまりインタビューを行っている途上でルーティングを間違え、聞くべき質問をスキップしてしまうことから生じていると指摘した（Smith 1984）。この場合、調査員のトレーニングを徹底することで項目非回答の発生を抑制できる。また、ファウラーとマンジョーニは、もしもインタビューの最中に回答者が回答に窮することが生じるならば、それは調査者の準備不足（質問があいまいで意図がわかりにくい、質問の文章に難しい専門用語が多用されている、適切な選択肢が用意されていない、調査員の発話が聞き取りにくい、など）に起因するものであり、したがって事前に十分な準備と調査体制の見直しを行うことで抑制が可能であると主張する（Fowler & Mangione 1990）。

他方で、この無効回答の発生については、異なる視点も提示されてきた。1964年、アーリックは、個々人に対して意見や態度を問うような質問をする場合に、回答者がその対象となる事柄

をよく知らない、あるいは、はっきりした見解を持っていないことが、ごく当たり前で生じていることに気づいた (Ehrlich 1964)。同時期、コンバースは、ニュースや新聞などで日常的に使われる「リベラル」という用語に着目した (コンバース 1968)。この世論調査で頻繁に使用されている常識的な単語の意味を回答者に聞いてみると、世論調査の対象となった一般の人々の間できちんと理解できている人がほとんどいないこと、大半はあいまいな知識か、およそ見当はずれな理解しか持たないことが明らかになった。こうした事例は近年でも報告されており、たとえばムーアらが所得についての質問であいまいな回答や不正確な回答をする回答者について調べたところ、税金の徴収を恐れて正確な回答を行わないというよりも、むしろ本当に正確な情報を知らないという人が多いことが明らかになった (Moore et al. 1999)。

知らないことについて賛否を問われた場合、あるいは数字を示すように求められた場合、回答者にとって適切な回答は「知らない」・「わからない」しか存在しない。この時に、調査者の側が、なんらかの方法で、わからないその他の無効回答を抑制する (たとえば「わからない」という回答に対して「もう一度よく考えてください」と返す、「わからない」という選択肢を調査票に記載しない、など) と、回答者は回答に窮してしまう。回答に窮した回答者がどのような選択肢を選ぶかは予想できず、その結果、有効とみなされる選択肢を選んだ回答者のうち、本当に賛成なり反対なりの意見を持っている人がどれぐらいの割合で存在するか、回答の精度に疑いが生じることになる。そして、アーリックやコンバースが指摘したのは、社会調査において回答者が聞かれた内容に相当する情報を持っていない (つまり「知らない」) ことや聞かれた内容そのものを理解できない (「わからない」) ことが、調査に携わる調査者が予想しているよりも頻繁に発生しているということである。したがって、「わからない」・「知らない」という回答は無理に抑制すべきではなく、むしろ選択肢として与え、その上で集計から除外する方が、精度の高いデータが得られる (Mondak 2001)。これが項目非回答に対する、もう一つの立場を形成してきた。

この論争は 30 年以上の間続いてきたが、1990 年代に一つの転換を迎えることになった。それがクロズニックによる最小限化 (satisficing) 回答理論の提唱である (Krosnick 1991)。クロズニックは、回答者が果たして正確な情報を提供しているのかという疑いを投げかける。調査で聞かれたことについて答えるのは認知心理学的に負荷の大きい作業である、とクロズニックは考える。聞かれた内容がなんであるかを正確に理解し、それに相当する情報を記憶を走査して抽出し、これら抽出された断片的な情報を整合的に統合し、自身の意見あるいは態度を確定する。さらに、この意見なり態度なりを表出することで相手 (調査者または調査員) からどのような反応が得られるかを予測し、表出する態度や意見に修正を加える。これだけの作業を数秒のうちに行い、しかもそれを質問ごとに繰り返さなければならない。疲れるのは当然であり、面倒になって手を抜いてしまうことも日常的に発生する。このように、調査への回答を円滑に行っているように装いながらも、回答に要する認知的作業を回避し最小限の労力で与えられた課題すなわち回答を終えようとするのを、クロズニックは最小限化回答と名付けた (Krosnick 2002; Krosnick et al. 2002)。

クロズニックによれば、「知らない」・「わからない」などの無効回答は、本当に「知らない」・「わからない」よりも、最小限化回答すなわち「回答するのが面倒」や「考えるのに疲れた」という

意思表示であるともみなされる。わかりやすく言えば、アーリックやコンバースらは無効回答を「can」（答えられるか否か）の問題としてとらえていたのに対し、クロズニックは「will」（答えようとするか否か）の問題としてとらえる（Beatty & Herrmann 2002）。したがって「わからない」や「なんともいえない」といった無効回答を選択肢として調査票に提示することは、情報の正確さを向上させることにはつながらず、逆に最小限化回答を誘発しデータの精度を低下させる。むしろ無効回答を抑制する方がデータの精度向上に貢献する。これがクロズニックの指摘である。

最小限化回答は回答に要する認知的負荷が大きいときに発生すると予想される。したがって、認知的能力の低い（同じ作業を行っても認知的負荷を大きく感じる）人、難しい（認知的負荷の大きい）質問、回答へのモチベーションの低い人に生じやすいと考えられる。ここから、もしもクロズニックの主張するように調査票の質問項目に無効回答を選択肢として与えることで最小限化回答が誘発されるとしたら

- ①無効回答を選択する人は認知的能力の低い人（低学歴者や高齢者）に多い
- ②難しい質問ほど無効回答の選択率が高くなる
- ③無効回答の選択率は、回答に疲れてくる後ろの質問ほど高くなる
- ④無効回答を選ぶ方が認知的負荷が軽いいため、回答に要する時間が短い

といった現象が観察されるはずである。反対に、アーリックらのように、無効回答選択肢をもうけることで回答に窮した回答者が集計から排除されるのであれば、

- ⑤回答者の分布で中央近くが低下し両極にある強い意見の選択率が上がる
- ⑥肯定的回答（「賛成」・「そう思う」など）の選択率が低下する
- ⑦同一の尺度を構成する他の質問項目との相関係数が高くなる

といった現象が観察されると予想される。⑤の中央部の回答の減少は、中間回答傾向（増田 他 2017; 福田 2018）と呼ばれるもので、自己の回答に合致する適切な選択肢が見つからない場合、中央に位置する回答を選びやすくなる傾向をいう。心理学や社会心理学の質問項目では、肯定（positive）と否定（negative）を両極に置く段階的評定が一般的に用いられる。そこでは中央に近づくと肯定・否定ともに弱い態度を示すことになり、中央に位置する選択肢はニュートラル、つまり肯定でも否定でもない回答とみなされる。この中央に、「わからない」・「知らない」という選択肢を選ばなかった回答者が集まると考えられる。したがって、「わからない」が選択肢として与えられ無効回答として集計から除外されると、中央およびこれに近い選択肢の選択率が低下し、両極に近い選択肢の割合が相対的に上昇すると予想される。⑥は黙従傾向（acquiescence）と呼ばれる現象で、はっきり理解していない事柄や関心のない事象について問われたとき、深く考えないまま「yes」と答えてしまう心理的傾向を指す（Ehrlich 1964）。これは対面的状況で「no」と発言することが、対面している相手を拒否するものと受け取られるという社会文化的な慣習から生じたものと理解される。

⑦はクロズニックが強調するもので、自身の意見に合致する選択肢がないために回答に窮した回答者はランダム (random) に回答すると仮定する (Krosnick 1999; Krosnick et al. 2002)。特定の位置に偏らずランダムに回答が分散すると、回答の分布の変化からその影響を把握することはきわめて困難となるが、測定誤差が大きくなることで項目同士の相関係数が低下する。つまり単変量では変化が見つからなくとも、多変量ではデータの精度が低くなっていることが観察されるはずだ、とクロズニックは主張する。実際、クロズニックが無効回答を選択肢として与えてもデータの精度が向上しないと批判した根拠は、この相関係数の変化がほとんど観察されないという点にあった。

### III 実験の設定

ウェブ調査実験は、2019年2月末に行われた。委託先は自社モニターを抱えるウェブ調査会社で、対比実験用にA・B二つのフォーマットを用意し、Aフォーマットへの回答者数を300、Bフォーマットへの回答者数を300として、同一のモニターがAとBの両方に回答しないように振り分けるように依頼した。なお、回答者については年齢・性別・居住地などの制約をつけず、割り付けも行わないことを委託の条件とした。これは、今回の調査があくまでも比較実験であり、標本の代表性を問題とする必要がないからである。

調査項目数は、性別・年齢・婚姻状態・学歴・就労状態および世帯所得に関する設問が9、これはA・B両フォーマットで共通である。そして態度に関する項目が、Aフォーマットでは外国について(8)、愛国心(7)、現在の世相(11)、家族および性役割分業(7)、階層意識もしくは生活の程度(5)で合計38項目となっている。Bフォーマットでは、外国について(4)、愛国心(7)、現在の世相(2)、家族および性役割分業(7)、階層意識もしくは生活の程度(5)で合計25項目、ただし現在の世相に関する設問はそれぞれ12選択肢の複数回答となっているため、カラム数では合計47となる。

この実験調査は内閣府世論調査の測定誤差の検討を目的に実施されたもので、Bフォーマットは基本的に内閣府世論調査の調査項目を質問文および回答選択肢をそのまま転載したものであり、Aフォーマットはこれに改良を加えたもの、つまり質問文の文言や回答選択肢を変更したものである。ただし、愛国心に関する質問では辻(2008)の行った調査を参考に6項目を、家族および性役割に関する質問では他の社会調査を参考に6項目を、階層意識に関する項目はSSM調査の階層意識調査項目とオリジナルの項目を3項目、A・Bフォーマットともに追加してある。

本稿で取り上げたいのは、この態度に関する項目のうち、愛国心に関する7項目、家族および性役割分業に関する6項目、そして階層意識に関する4項目である<sup>4)</sup>。これら17項目はAフォーマットとBフォーマットで質問文は全く同じになっているが、回答選択肢についてAフォーマットは「わからない」という選択肢が加えられており、Bフォーマットにはこの選択肢が存在しない。両フォーマットともに、無回答のまま次画面へ進むことはできない(ブロックされる)ように設計されているため、Bフォーマットの場合は、わからなくても何かを選択せざるを得ないようになっている。ここで無効回答選択肢の影響を測定しようというわけである。

両フォーマットの回答者のバランスを見ると、Aフォーマットの回答者の女性割合が41.7%、未婚者率24.7%、高等教育（専門学校を含む）進学者率70.6%、就労率65.3%、平均年齢56.2歳となり、Bフォーマットでは女性割合38.7%、未婚者率24.0%、高等教育進学者率72.3%、就労率67.3%、平均年齢55.8歳となった。つまり二つのフォーマットの回答者の属性については有意な差が存在しない。世帯所得についてはAフォーマットの平均が620万円とBフォーマットの平均が641万円と、多少の差がみられるが、これは外れ値の影響、つまり世帯所得を2000万円以上と回答した者がBフォーマットで3名多かったことによるものと判断される。実際、中央値で見ると両フォーマットとも550万円と差が存在していない。つまり、態度に関する項目での差が回答者の属性に関する構成の違いから生じたと考える必要はないと判断できる。

#### IV 結果

無効回答選択肢を与えることは最小限化回答を誘発するというクロズニックの仮説から検証していきたい。表1は態度項目に関して個々の回答者が無効回答（図表中ではDK選択肢と表記されている）を選択した回数を従属変数とした回帰分析の結果である。まず注目すべきは決定係数であり、0.134とかなり小さい数値を示している。つまり無効回答の選択と基礎属性との関係は統計的には有意とは認められるが、必ずしも強いものではない。そして、有意と判断されたのは婚姻状態と年齢であった。つまり未婚者と若年者ほど無効回答を選択しやすい。クロズニックの仮説では認知能力に直接関係する学歴<sup>5)</sup>と年齢が影響するはずだが、学歴は有意差を示しておらず、さらに年齢についてはクロズニックの予想とは逆に高齢者ほど無効回答の発生率が低くなっている。さらに、難しい質問、回答するのに高い認知的負荷を要求する質問ほど無効回答の選択率が高くなるというクロズニックの予想について、回答に要した時間<sup>6)</sup>を質問の難解さのインデックスとして無効回答選択率との関係を見たところ（図1）、回答に要する時間と無効回答選択率との間に明確な関係性は全く認められなかった。つまり、回答に要する認知的負荷と無効回答の選択率との間には関係が認められない。

表1：無効回答発生回数を従属変数とする回帰分析結果

	係数	標準偏差	t	p
未婚者* <sup>1</sup>	1.644	0.428	3.84	0.000 <sup>†</sup>
女性* <sup>1</sup>	0.662	0.369	1.79	0.074
就労者* <sup>1</sup>	-0.478	0.414	-1.15	0.249
学歴* <sup>2</sup>	-0.592	0.387	-1.53	0.127
年齢	-0.046	0.019	-2.37	0.019 <sup>†</sup>
平均回答時間* <sup>3</sup>	-0.105	0.098	-1.07	0.288
(定数)	5.47	1.279	4.27	0.000
n	296			
F値	7.47			
p	0			
R <sup>2</sup>	0.134			

\*<sup>1</sup> ダミー変数

\*<sup>2</sup> ダミー変数、高卒以下を0、専門学校・短大以上を1

\*<sup>3</sup> 各態度項目の回答時間（秒）の平均値

<sup>†</sup>  $\alpha < 0.05$

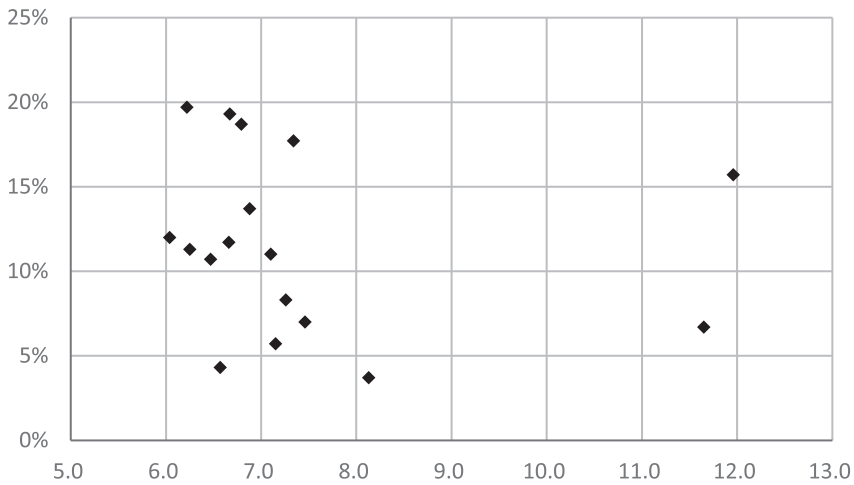


図1：平均回答時間（横軸）とDK 選択率（縦軸）

クロズニックの仮説でもう一つ問題とされる回答モチベーションの問題について、後ろの質問ほど無効回答選択肢の選択率が上がるかどうかを検証したところ（図2）、予想とは逆に、後ろの質問ほど無効回答の選択率が低くなるという現象が見られた。つまり、認知的負荷が高いときと回答に対するモチベーションが低いときに最小限化回答が生じやすくなるというクロズニックの仮説を支持する結果は、得られなかった。負荷が大きいときの安直な逃げ道として無効回答選択肢が選ばれるのであれば、無効回答を選択する回答者の方が回答時間が短いはずだが、表2をみても無効回答を選択する回答者は回答に要する時間が短くなるという傾向はみられない。こうした結果を見る限り、無効回答選択肢が与えられることで最小限化回答が誘発されるとは言い得ない。

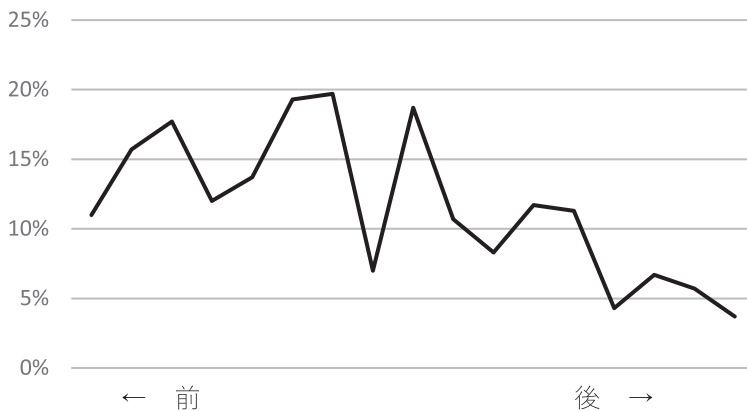


図2：質問の順序とDK 選択率



表 2：有効回答者と無効回答者の回答時間（分散分析）

	有効回答者	無効回答者	<i>f</i>	<i>p</i>
愛国心は強いほうか	7.08	7.26	0.12	0.732
伝統や慣習が軽視されすぎ	12.20	10.68	3.01	0.084
国を愛する気持ちを育てる必要がある	7.02	7.14	0.07	0.785
国に愛着を持つのは自然なこと	6.14	5.27	6.72	0.010 <sup>†</sup>
オリンピックで日本を応援するのは当然	6.90	6.70	0.20	0.655
学校で国旗掲揚・国歌斉唱は当然	6.79	6.16	2.11	0.147
小中学校で愛国教育が必要	6.32	5.80	1.96	0.162
家族は誰にとっても大事なもの	7.45	7.62	0.07	0.789
一生独身でも幸せに生きられる	6.73	7.04	0.48	0.490
既婚女性は家事と育児に専念すべき	6.45	6.60	0.11	0.737
既婚男性は定職について家族を養う	7.34	6.26	2.83	0.094
女性の目標は家庭を築くこと	6.69	6.41	0.32	0.572
男性の目標は仕事で成功すること	6.31	5.85	0.98	0.324
階層意識（内閣府形式）	6.61	5.83	1.07	0.302
階層意識（SSM形式）	11.70	10.98	0.39	0.535
主観的生活水準	7.16	6.92	0.10	0.750
物的不足感	8.13	8.09	0.00	0.976

<sup>†</sup>  $\alpha < 0.05$

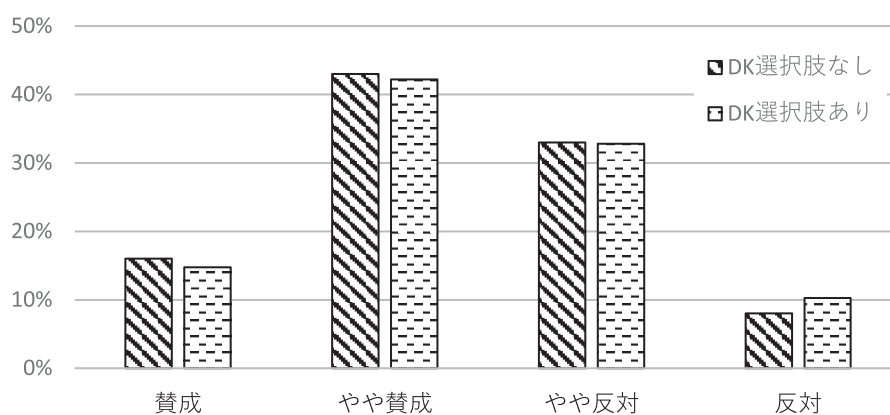


図 3：「一生独身でも幸せな人生を送れる」への賛否

次にアーリックやモンダクの主張する、無効回答を選択肢として与えることで回答の正確さが向上するという仮説について、まず中間回答が減少するかという点から検証していきたい。図 3 では、「一生独身でも幸せな人生を送れるという意見に賛成かどうか」という質問項目をとりあげたが、この質問は無効回答の選択率が 18.7% ときわめて高いにもかかわらず、無効回答が与えられない場合と与えられた場合とで、分布の形がほとんど変化していないことがわかる。これは他の項目についても同様で、無効回答を選択肢として与えると中央に位置する回答の発生率が低下するという中間回答傾向は、少なくとも本実験では認められなかった。

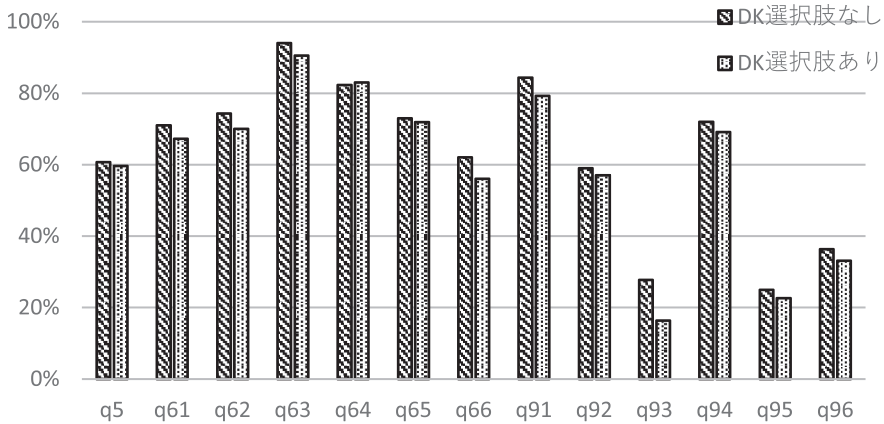


図4：賛成の割合

黙従傾向については、図4にみられるように、各項目とも、無効回答選択肢を与えた場合は、肯定的な選択肢の選択率がわずかに低下する傾向がみられる<sup>7)</sup>。平均すると4%ほど低下しているが、ただし、これは統計的には有意差とは判定されない。つまり黙従傾向はおそらく存在していて、それが無効回答選択肢を与えることで軽減されていると推測はされるが、それは統計学的に顕著といえるほどはっきりしたものではない。なお、この点については、アーリックが全く同じ結果を得ている (Ehrlich 1964)。最後に、愛国心に関する7項目、性役割分業に関する6項目、階層意識に関する4項目をそれぞれ一つの尺度とみなし、信頼性係数を算出したところ、愛国心に関する尺度のクロンバッハの $\alpha$ 係数が無効回答選択肢がないときで0.846、無効回答選択肢を与えたときが0.890と、ほとんど変化がみられなかった。これは性役割分業の尺度も同様で、無効回答選択肢がないときが0.773、無効回答選択肢を与えると0.771となった。階層意識の尺度では0.911 (無効回答選択肢がない場合) から0.902 (ある場合) と、きわめて微細な変化しか見られず、選択肢を与えることで信頼性係数が上がったとは言い得ない結果となった。

上記の内容をまとめると、アーリックらの主張するように無効回答を認めることで有効回答のデータの正確さが向上するという明瞭な結果は得られず、同時にクロズニックらの無効回答を認めることは最小限化回答を誘発するという主張を支持する結果も得られなかった。

## V 考察

同じように無効回答の問題を論じていても、政治学者や社会学者は、個人によって程度の差はあるものの、「わからない」という回答を許容する傾向があるのに対し、心理学者や社会心理学者は尺度上に置かれない回答に対してきわめて厳しい。この理由の一つには、同じ調査票を用いても、対象とする標本の特性が学問ごとに異なっていることがあげられる。標本の代表性にこだわる政治学や社会学は、調査者と面識のない人々が調査票に回答することを前提としている。したがって、やる気があるかないかも含め、ある程度まで回答者の予想外の行動も、実際に対応

できているかは別としても、想像の範囲の内にある。これに対し心理学で調査票を用いた調査を行う場合、主に学生、それも多くの場合は調査者の教えている学生を被検者とするのが想定されている。合理的思考と知的活動に慣れた回答者が意識的あるいは無意識に起こすミスは、決して頻度の高いものではない。どうしても、厳しい眼差しが向けられてしまう。

こうした点から見れば、今回のウェブ調査実験の回答者は、社会学者にとっても心理学者にとっても、従来とは異なる特性を持つ。社会学者の想定と異なり、ウェブ調査の対象者は全くの未経験者ではない。調査会社の誘いに応え、これまでも何度もマーケティング調査に回答してきた人々である。聞かれることに、それなりに慣れているとあってよい。同時に、心理学者が今まで主たる対象としてきた大学生とも異なり、学歴も年齢もまちまちで、質問されている内容に対する理解度も決して高くはない。今回の実験調査が政治学者・社会学者の予想とも、心理学者・社会心理学者の予想とも異なる結果を示したのは、このウェブ調査の回答者特性と関わっている可能性がある。しかしながら、自ら進んで登録モニターとなった回答者の特性は、まだわかっていない部分が多く、今後の研究の展開を待つ必要がある。

ただ、今回の実験の中で少々気になる点も観察されている。態度項目に対する無効回答の選択率は未婚者で高いと記述したが、この傾向は性役割分業に関する6項目で顕著に表れた。つまり家庭内の性役割分業に関する質問に対し、賛否両方を含めた態度を明らかにした人の中には未婚者は20%ほどしかいないが、「わからない」と答えた人のうちの未婚者の割合は50%に達する(図表省略)。これは直感的には理解できる。つまり家庭内で男性あるいは女性に何が異性から期待されているかというのは、結婚し日常生活を営んでいく中で経験し実感していくという一面を持っている。個々人の態度とはその人の生活経験の産物であるというのはトーマス以来の社会学の基本命題であり、その点からいえば、個々人の生活経験の内容によって、たとえ自分自身のことであっても「わからない」という回答が生じ得るということになる。

こう考えると非常に興味深いのが、愛国心に関する質問項目への回答である。愛国心に関する項目は「愛国心が強い方だと思うか」という内閣府世論調査の質問1項目と、もっと具体的な場面(例えば小学校での愛国教育の義務化など)を想定した6項目からなる。このうち漠然と聞いている内閣府世論調査の項目では無効回答(「わからない」)を選択した人と何らかの有効回答を選択した人の間に年齢の差が見られないが、それ以外の6項目では、無効回答を選択した人の方が5歳ほど平均年齢が低いのである(表3)。これを先ほどの「生活経験」というキーワードで考えるならば、高齢世代は愛国心を問われる具体的な経験を持っているのに、若年者になるとそうした具体的な経験が乏しいということになるだろう。これも生活実感としては理解できる。小泉政権以降、政治家が愛国心を公然と口にすることが日常化した反面で、愛国心をめぐる騒動(例えば卒業式で掲げられた日本国旗を式典に参加した教員や生徒が引きはがすシーンがテレビで報道されたり、街頭右翼の宣伝カーと警察車両がもみ合いになった光景など)はめっきり姿を消した。こうした社会的風潮の変化があるとすれば、若年者は愛国心が具体的に問われる場面がイメージしにくくなっているという可能性があるかもしれない。

表3：愛国心測定項目でのDK回答者・有効回答者の平均年齢（分散分析）

	有効回答者	DK回答者	<i>f</i>	<i>p</i>
愛国心は強いほうか	56.0	54.6	0.46	0.496
伝統や慣習が軽視されすぎ	56.9	50.1	15.06	0.000 <sup>†</sup>
国を愛する気持ちを育てる必要がある	56.9	51.0	12.67	0.000 <sup>†</sup>
国に愛着を持つのは自然なこと	56.7	49.2	15.09	0.000 <sup>†</sup>
オリンピックで日本を応援するのは当然	56.6	51.2	8.45	0.004 <sup>†</sup>
学校で国旗掲揚・国歌斉唱は当然	56.9	51.3	12.47	0.000 <sup>†</sup>
小中学校で愛国教育が必要	56.7	52.1	8.28	0.004 <sup>†</sup>

<sup>†</sup>  $\alpha < 0.05$

表4：年齢別愛国心得点（分散分析）

年齢	DK 選択肢なし	DK 選択肢あり
-34	15.82	16.33
35-49	17.20	17.78
50-64	17.56	17.87
65-	18.73	18.71
<i>f</i>	4.36	0.77
<i>p</i>	0.01	0.51
( <i>n</i> )	300	182

この解釈は、現時点ではまだ憶測にすぎない。だが、この6項目を一つの尺度として合算してみると、無効回答を与えなかった場合は愛国心得点と年齢の間に明らかな関係がみられる（高齢になるほど愛国心が強くなる）のに対し、無効回答を選択肢として与えて集計から除外した場合は愛国心得点と年齢の間に関係はみられなくなる（表4）。つまり無効回答選択肢を与えることで異なる現実が見えてくる可能性を、この事例は示唆していると考えられないだろうか。だとすると、無効回答選択肢を与えることでデータの正確さは増すかという問題は、信頼性（reliability）の問題に留まらず妥当性（validity）の問題に立ち入って議論していく必要があるということになる。

## 注

- 1) 本研究はJSPS 科研費 18H00940 の助成を受けて行われた。
- 2) 内閣府世論調査については、現在、インターネット上での公表が原則となっている。引用の31年度2月実施の「社会意識に関する世論調査」の記載ページは、<https://survey.gov-online.go.jp/h30/h30-shakai/index.html> にある。
- 3) 「国民生活に関する世論調査」（平成30年2月実施）の記載ページは、<https://survey.gov-online.go.jp/h30/h30-life/index.html>
- 4) 本稿でとりあげる態度項目の質問文は、以下のとおりである。  
（愛国心項目）

Q5 あなたはご自分で、「国を愛する」という気持ちが強い方だと思いますか、それとも弱い方だと思いますか。

次にあげる日本についての意見に、あなたは賛成ですか、それとも反対ですか

- Q6-1「今の日本では伝統や慣習が軽視されすぎている」
- Q6-2「国を愛する気持ちをもっと育てる必要がある」
- Q6-3「日本という国に愛着を持つのは自然なことだ」
- Q6-4「オリンピックや国際スポーツ大会で日本を応援するのは当然だ」
- Q6-5「小中学校で国旗掲揚や国歌斉唱を行うのは当然だ」
- Q6-6「小中学校でもっと愛国心を育てる教育が必要だ」

(家族および性別分業項目)

家族についてよく言われる、次のような意見について、あなたはどのように思いますか

- Q9-1「家族は誰にとっても必要なものだ」
- Q9-2「一生独身でも幸せな人生を送れる」
- Q9-3「女性は結婚したら仕事をやめて家事と育児に力を入れるべきだ」
- Q9-4「男性は結婚したら定職について家族を養わなければならない」
- Q9-5「女性にとって人生の目標は結婚して家庭を築くことにある」
- Q9-6「男性にとって人生の目標は仕事で成功することにある」

(階層意識項目)

- Q10 あなたの生活の程度は、世間一般から見て、次のどれにあてはまりますか。(回答選択肢：上・中の上・中の中・中の下・下)
- Q11 では、同じくあなたの生活の程度について、次のようにかがたら、あなたはどこに入りますか。(回答選択肢：上・中の上・中の下・下の上・下の下)
- Q12 他の人と比べて、あなたとあなたのご家族の収入は高い方だと思いますか、それとも低い方だと思いますか。
- Q13 お給料や年金・投資収益などを含め、あなたとあなたのご家族の収入でゆとりのある生活ができていると感じますか、それとも生活が厳しいと感じますか。

- 5) 学歴については、国勢調査にならない、中退者は卒業学校をもって学歴とした。なお、現在就学中の者および「その他」を選択した者は除外した。回帰分析の標本数 ( $n$ ) が 300 ではなく 296 となっているのは、この理由による。
- 6) ウェブ調査では回答者が自宅で回答するため、途中で席を外したり、中断して後に再開するといったことが頻繁に発生する。したがって回答時間をデータとして取得する場合、一つの質問に対する回答時間が極端に長くなることがある。実際、本実験ではほとんどの質問への回答時間が 6 秒から 12 秒ほどで収まっているが、まれに 1000 秒を超えるような回答時間が現れている。こうしたデータ特性への対応として外れ値 (outlier) の処理が必要になるが、これほど極端な値が出てくると平均値を用いて外れ値を特定するのは不適切となる。そこで本稿では平均値ではなく中央値に着目し、median absolute deviation test (MAD) を用い、最大値  $\text{Max} = 4$  を超えるものを外れ値として除外した。MAD については Sheskin (2011) を参照のこと。
- 7) 17 項目のうち、「オリンピックや国際スポーツ大会で日本を応援するのは当然だ」という質問だけが、無効回答選択肢を与えた方が肯定的回答の発生率が高くなっている。ただし、高くなっているといっても、その差は 0.7% しかない。

## 文献一覧

- Beatty, Paul, & Douglas Herrmann. (2002). To Answer or Not to Answer: Decision Processes Related to Survey Item Nonresponse. in Robert Groves et al. eds. *Survey Nonresponse*.
- De Vaus, David. (2014). *Surveys in Social Research, 6th ed.*

- Ehrlich, Howard J. (1964). Instrument Error and the Study of Prejudice. *Social Force*, 43:197-206.
- Ferber, Robert. (1966). Item Nonresponse in a Consumer Survey. *Public Opinion Quarterly*, 30: 399-415.
- Groves, Robert M. (1989). *Survey Errors and Survey Costs*.
- Lessler Judith T. & William D. Kalsbeek. (1992). *Nonsampling Error in Surveys*.
- Kaczmirek Lars. (2009). *Human-Survey Interaction: Usability and Nonresponse in Online Surveys*.
- Krosnick, Jon A. (1999). Survey Research. *Annual Review of Psychology*, 50: 537-67.
- Krosnick, Jon A. (2002). The Causes of No-Opinion Responses to Attitude Measures in Surveys. In Robert Groves et al. eds. *Survey Nonresponse*.
- Krosnick Jon A., Allyson L. Holbrook, Matthew K. Berent, Richard T. Carson, William Hanemann, Raymond J. Kopp, Robert Cameron Mitchell, Stanley Presser, Paul A. Ruud, V. Kerry Smith, Wendy R. Moody, Melanie C. Green, & Michael Conawayet. (2002). The Impact of "No Opinion" Response Options on Data quality. *Public Opinion Quarterly*, 66:371-404.
- Mason, Robert, Virginia Lesser & Michael W. Traugott. (2002). Effect of Item Nonresponse on Nonresponse Error and Inference. In Robert Groves et al. eds. *Survey Nonresponse*.
- McNabb, David E. (2014). *Nonsampling Errors in Social Survey*.
- Mondak, Jeffery J. (2001). Developing Valid Knowledge Scales. *American Journal of Political Science*, 45(1): 224-238.
- Moore, Jeffery, Linda L. Stinson, & Edward J. Welniak, Jr. (1999). Income Reporting in Surveys: Cognitive Issues and Measurement Error. In Sirken, Monroe G. et al. eds. *Cognition and Survey Research*.
- Poe, Gail S., Isadore Seeman, Joseph McLaughlin, Eric Mehl & Michael Dietz. (1988). Don't Know Boxes in Factual Questions in a Mail Questionnaire. *Public Opinion Quarterly*, 52: 212-222.
- Sheskin, David J. (2011). *Handbook of Parametric and Nonparametric Statistical Procedures*.
- Smith, Tom. (1984). No Attitudes. In Tuner, Charles F. & Elizabeth Martin eds. *Surveying Subjective Phenomena. Vol.2*.
- Sudman, Seymour & Norman Bradburn. (1974). *Response Effects in Surveys*.
- Weisberg Herbert F. (2005). *The Total Survey Error Approach*.
- 石田賢示 (2016). どちらともいえないとわからないは何を意味しているのか、『東京大学社会科学研究所パネル調査プロジェクトディスカッションペーパーシリーズ』、97号
- コンバース、フィリップ (1968). 国民大衆における信条体系の性格、デヴィット・アプター編『現代政治とイデオロギー』、慶應義塾大学地域研究グループ訳、慶應通信
- 辻大介 (2008). 『インターネットにおける右傾化現象に関する実証研究：調査結果概要報告書』
- 土屋隆裕 (2005). 電話調査における調査所要時間と回答傾向について、『統計数理』、53(1):83-101.
- 福田昌史 (2018). 0-100 尺度質問への中間回答傾向とその影響、『政策と調査』、14:19-28.
- 増田真也・坂上貴之・北岡和代 (2017). 多くの項目に回答することによる中間選択の増加、『行動計量学』、44(2):117-128.
- 三浦麻子・小林哲郎 (2015). オンライン調査モニタの Satisfice に関する実験的研究、『社会心理学研究』、31(1):1-12.
- 吉村治正 (2017). 『社会調査における非標本誤差』、東信堂

## Summary

The effect of giving “don’t know” option in Web survey is examined in this paper. Web survey, due to its technical property, does not accept “don’t know” response, unless researcher sets an overt DK --- i.e., “don’t know” --- option for each item. This will be a very embarrassing situation for those who really don’t know what they are asked. Giving the DK option, accordingly, will contribute to the improvement of data-quality through excluding invalid answers mingled with other convincing responses, on the one hand. This treatment, on the other hand, has a potential risk to deteriorate data-quality, for it might invite satisficing behavior among respondents. Our web-survey experiment, however, supports neither hypotheses; giving the DK option does not enhance data-quality, nor induce satisficing. The effect of the DK option is rather minimal, and no conspicuous change in distribution is observed.

**Keywords:** Web survey, item nonresponse, measurement error, satisficing.

