

回答指示の非遵守と反応バイアスの関連¹

増田 真也 坂上 貴之 慶應義塾大学 北岡 和代 金沢大学
佐々木 恵 北陸先端科学技術大学院大学

Relationships between ignoring instructions and response bias when completing questionnaires

Shinya Masuda, Takayuki Sakagami (Keio University), Kazuyo Kitaoka (Kanazawa University),
and Megumi Sasaki (Japan Advanced Institute of Science and Technology)

Certain participants are insincere, or careless when they respond to questionnaires. To identify such participants, we included three items in a questionnaire that instructed participants to choose a particular response category. Nurses ($N = 1,000$) responded to this questionnaire in a Web survey. One-hundred-twenty participants failed to follow the instructions for at least one item (non-followers). Analyzing their responses indicated the following: (a) non-followers were more likely to give identical, or midpoint responses; (b) the correlations between their responses to regular and reversed items were low or positive, and their responses to scales containing reversed items tended to show lower internal consistency; and finally, (c) the mean scores of non-followers were closer to the midpoint of the scale, regardless of whether the scale included reversed items. One reason that including reversed items lead to lower internal consistency could be because participants occasionally missed responding to these items. However, the results suggested that non-followers were not diligent in responding to regular items, and merely deleting reversed items from scales will be insufficient to ensure accurate results.

Key words: response instruction, straight line response, midpoint response, reversed item, Web survey.

The Japanese Journal of Psychology

2016, Vol. 87, No. 4, pp. 354–363

J-STAGE Advanced published date: July 9, 2016, doi.org/10.4992/jjpsy.87.15034

調査の回答者は、設問に必ずしもまじめに答えるとは限らないし、勘違いやミスをすることもある。こうした回答は分析結果に悪影響を与えるため、何を有効データとするかは調査研究において根本的に重要である。しかし、その判別基準は明確ではなく、仮に同じ調査データを扱っても、研究者によって用いられる分析データが異なって、結果が一致しないということが起こる可能性がある。

分析から除くことが検討されるような回答に、多くの項目に続けて同じ選択肢を選ぶという同一回答 (straight line response もしくは non-differential response) がある。Herzog & Bachman (1981) は高校生を対象と

した複数の調査から、10–23 項目からなる尺度の全項目で同じ回答をした人が 0–6.7% であったこと、また設問数の多い調査票の後半で同一回答が増える傾向があったことを報告している。このことは、疲労や関心の低下のために同一回答が生じることを示唆しており、無効とする根拠になると思われる。しかし、回答者が真摯に考慮した結果、同じ回答カテゴリーの選択を続けた可能性もあるし、どの程度連続して同一回答がなされたら無効とすべきかについて、はっきりとした基準が示されてきたわけではない。

Vannette & Krosnick (2014) は、同一回答や、黙従傾向 (acquiescence) や中間選択 (midpoint response) といった、項目内容と無関係に系統的な回答がなされるという反応バイアス (response bias: Paulhus, 1991) が、最小限化 (satisficing) によって生じると考えている。最小限化は、もともと Simon (1957) によって、人が限定された能力や時間の中で、最適でないとしても受け入れられる、最小限の基準を満たす選択肢を選ぶと

Correspondence concerning this article should be sent to: Shinya Masuda, Faculty of Nursing & Medical Care, Keio University, Endo, Fujisawa 252-0883, Japan. (E-mail: masuda@keio.jp)

¹ 本研究は、平成 26 年度科学研究費補助金 (基盤研究 (B) (一般)) 課題番号 26293486、研究代表者 北岡 和代) の助成を受けた。

いう事実を説明するのに用いられた。回答行動に関してこの語が用いられる場合には、回答教示や項目内容を理解したり、選択肢を選んだりする際に、関連情報を十分に、もしくは全く想起・統合しないなど、認知的コストや努力を十分に払わないことを指す。

最小限化が生起しているかどうかを判別するために、Oppenheimer, Meyvis, & Davidenko (2009) は、IMC (Instructional manipulation check) という方法を考案した。IMC では、Web 調査でのある設問で、通常の選択肢を選ぶのではなく、画面上の他の箇所をクリックするよう教示をして、教示文を読んでいるかどうかを確かめる。すると、例えばこの論文の研究 1 の回答者の中で、教示に従わない人は 46% もいた。

三浦・小林 (2015) では、IMC のような、回答の仕方に関する教示のある設問に加えて、個々の項目で「この項目は一番左 (右) を選択してください」といった特定の回答の指示を含む設問を設けた。すなわち教示文だけでなく、項目文を読んでいるかどうかも調べた。結果は、ある設問の回答の仕方に関する教示文を読んでいない人の方が、他の設問でのある項目に記された指示に従っていないことが多かった。しかし、その割合は最大でも 20% 程度であり、回答者は長い教示文は読み飛ばすが、項目文自体は読んでいることが多かった。

本研究では、三浦・小林 (2015) と同様に、特定の回答を指示するような項目 (以下、指示項目とする) を設け、回答の指示に従った遵守者と、従わなかった非遵守者を比較する。他の項目で同一の回答が続いていても、指示項目で指示に従って別の回答をしているのであれば、少なくとも項目文は読んでいると思われることから、真摯な回答がなされた可能性が高い。逆に指示を遵守した回答であっても、その直前の項目でも指示と同じ回答が続いている場合には、項目文を読まずになされた惰性的な回答が、たまたま指示と一致したことが疑われる (例えば、「一番左を選択してください」という指示通り、左が選ばれているが、それ以前の項目でも、続けて左の選択肢が選ばれているなど)。このように、同一回答の検討によって、指示項目が本当に役に立つかどうかの検証ができるものと思われる。

さらに本研究では、指示の遵守の有無と黙従傾向や中間選択、同一回答との関連も検討する。黙従傾向や中間選択が重なることで、同一回答が生じる可能性もあることから、これらの関連を確認することが必要であろう。

最後に、指示に従わない回答者が含まれることで、分析結果がどう変わるのかを、特に逆転項目に注目して検討する。逆転項目として想定されていた項目が、尺度の内部一貫性を損ねることがあるが、その原因として、否定語を見過ごして逆転に気がつかないとい

た回答ミスが挙げられている (Swain, Weathers, & Niedrich, 2008; Woods, 2006)。指示に従わないということは、逆転項目も注意深く読まれていないことの裏付けになるだろう。そしてその結果、例えば逆転していない通常項目と逆転項目間に強い負の関連が見られず、尺度の内部一貫性が低下したり、想定していた因子構造が得られなくなったりすることがあるものと思われる。

方 法

対象者

本調査は、看護師の勤務継続意思に関する研究課題における予備調査として計画、実施された。そのため、調査会社にモニタ登録されている約 2,000 名の看護師に、調査会社を通じて調査参加の依頼をした。したがって研究者に個人情報には知らされない。

依頼に同意した対象者は、指定された Web サイトにアクセスして回答した。そして回答者が 1,000 名に達した時点で、調査を打ち切った。回答者には、調査会社の規定に基づく謝礼が支払われた。なお本研究は、金沢大学医学倫理審査委員会の承認を得て実施した。

回答者の性別は男性 14.6%、女性 85.4%、年齢の平均は 42.2 ($SD = 8.99$) 歳であった。また最終学歴は、71.1% が専門学校であり、短大卒以上が 24.2%、その他が 4.7% であった。勤務先は総合病院が 33.2%、個人病院が 31.5%、大学病院 8.3%、その他 27.0% であった。回答者の所在地は、人口の多い都道府県が多いものの、日本全国にわたっている。分析には R (version 3.2.2) を用いた。

質問項目

質問項目には、まず学歴等の 12 の属性項目があり、続けて複数の測定尺度があって、その中に 3 つの指示項目が挿入された。回答画面は、関連項目や尺度ごとにページが区切られている (ただし回答者の環境によっては、ページ内の全項目が一画面に収まらないことがある)。そして、ページ内の全項目に回答すると、次のページに進むことができる。すなわち、非回答は許されない。

測定尺度には、まず就業継続意思に関する 3 項目があり、「1. 全くそう思わない」、「2. あまりそう思わない」、「3. そう思う」、「4. とてもそう思う」から、最もあてはまる選択肢が尋ねられた。次に、病院に勤務する看護師の職務満足尺度 (撫養・勝山・青山, 2014) があり、「1. 非常にそう思う」、「2. かなりそう思う」、「3. まあまあそう思う」、「4. あまりそう思わない」、「5. そう思わない」の中で、最も気持ちを表す選択肢を選ぶよう求められた。これは「仕事に対する肯定的感情 (以下、仕事肯定感とする)」、「働きやすい労働

Table 1
指示項目において選択された回答カテゴリ

	回答カテゴリ						
	1	2	3	4	5	6	7
指示項目 1	923 ^a	9	45	14	9		
2	11	3	7	34	9	5	931 ^a
3	944 ^a	5	41	6	4		

^a 遵守者数を示す。

環境（以下、好労働環境とする）, 「上司からの適切な支援（以下、上司支援とする）, 「職場での自らの存在意義（以下、存在意義とする）」の4つの下位尺度からなる計28項目の尺度で、この中の23番目に、「ここでは一番左をチェックして下さい」という項目が加えられた（指示項目1）。したがって、見た目上は29項目が並んでおり、指示に従う回答者は、23番目の項目で「1. 非常にそう思う」を選ぶことになる。

次に、高木・石田・益田（1997）が作成した組織コミットメント尺度における愛着要素（組織への情緒的な愛着の程度）の8項目（ただし「会社」を「病院」とした）と、根付き（Job Embeddedness）尺度（Crossley, Bennet, Jex, & Burnfield（2007）が開発し、Peltokorpi, Allen, & Froese（2015）が作成した日本語版）が続く。この尺度は、「1. そう思う」、「2. どちらかといえばそう思う」、「3. どちらともいえない」、「4. どちらかといえばそう思わない」、「5. そう思わない」の5つから、あてはまる選択肢を選ぶ。

根付き尺度は本来会社勤務者を対象としているが、ここでは「会社」を「病棟」と「病院」に置き換え、それぞれについて7項目ずつ「1. 全くそう思わない」、「2. そう思わない」、「3. あまりそう思わない」、「4. どちらとも言えない」、「5. 少しそう思う」、「6. そう思う」、「7. 非常にそう思う」の7件法で尋ねた（以下「病棟への根付き」、「病院への根付き」とする）。そして「病院への根付き」について尋ねる設問の4番目に、「一番右をチェックして下さい」という項目が挿入された（指示項目2）。

指示項目3は「一番左をチェックして下さい」であり、職場環境について「1. 全く違う」、「2. 違う」、「3. わからない」、「4. そうである」、「5. 全くそうである」であてはまる程度を尋ねる38項目の尺度（北岡・増田・森河・中川（2015）により作成された日本版 Areas of Worklife Survey（AWS）28項目に、患者に関するストレスについて尋ねる独自に作成・追加した10項目：以下、患者ストレスとする）の22番目に挿入された。AWSは「仕事の負担」、「裁量権」、「報酬」、「共同体」、「公平性」、「価値観」の6つの下位尺度からなる。

尺度の違いを考慮しなければ、属性項目以外に94の小問があり、指示項目は26, 51, 76問目に設けら

れている。そして回答指示を遵守した場合、それぞれ1, 7, 1の回答値が得られることになる。

結 果

指示非遵守の割合

非遵守者は、指示項目1で77名（7.7%）、2で69名（6.9%）、3で56名（5.6%）であった（Table 1）。Cochran Q 検定の結果は有意であり（ $\chi^2(2) = 7.326, p = .026$ ）、多重比較（Wilcoxon のサインランク検定, Bonferroni の調整）の結果、指示項目1と3に有意な比率の差（ $p = .029$ ）があった。

また、3つの指示項目のすべてで遵守しなかった人は2.8%、一度でも遵守しなかった人は12.0%（以下、有非遵守者とし、項目ごとの非遵守者と区別する）、逆にすべてで遵守したのは88.0%であった（以下、全遵守者とする）。全遵守者と有非遵守者の間で、性別（ $\chi^2(1) = 0.166, p = .683$ ）、年齢（ $t(152.928) = 1.850, p = .066$ ）、学歴（専門学校卒か短大卒以上か： $\chi^2(1) = 1.201, p = .273$ ）、勤務先（ $\chi^2(3) = 2.857, p = .414$ ）の有意な差は見られなかった。

指示の遵守と矛盾回答との関連

本調査では、現勤務先の勤務年数、看護師としての勤務年数、年齢を尋ねているが、現勤務先や看護師としての勤務年数よりも年齢の方が小さいというものはありえない。しかし、2名がこのような回答をしており、うち1名は全部の指示に従っておらず、もう1名はすべて遵守していた。次に、1名が年齢よりも看護師勤務年が大きいと回答していたが、指示はすべて遵守されていた。

看護師としての勤務年数よりも、現勤務先の勤務年数の方が大きいことはありうる（例えば、事務職として勤務した後、看護師として同じ病院に務めるなど）。しかし、看護師資格の取得後に、初めて働きはじめるのが普通であると思われることから、矛盾回答である可能性が高い。本調査では34名が現勤務先の勤務年数の方が看護師としての勤務年数よりも長いと回答しており、このうち7名（有非遵守者中の5.8%）は3つの指示の少なくとも1つに従わない有非遵守者で

あった。一方、全遵守者でこうした回答をしたのは27名（全遵守者の3.1%）であったが、フィッシャーの正確確率検定で有意な差はなかった（ $p = .172$ ）。このように、そもそも矛盾回答が少なかったこともあり、指示の遵守と明確な関連は見られなかった。

指示の遵守と黙従傾向との関連

黙従傾向があると、指示項目1で「1.非常にそう思う」、指示項目2で「7.非常にそう思う」が選ばれ、指示を遵守したとみなされることになる。しかし、指示項目3で指示に従うと否定的な回答カテゴリ（「1.全く違う」）を選ぶことになるため、黙従傾向による回答と区別することができる。

このような指示項目1、2で遵守、3で非遵守という回答は17名で見られた。そしてこの中で、指示項目3で肯定的な回答（「5.全くそうである」）を選んだ人は4名（全回答者の0.4%）で、弱い肯定（「4.そうである」）を選んだ人は1名であった。残りの12名はすべて、中間選択（「3.わからない」）であった。すなわち、黙従傾向と指示の遵守との関連はほとんど見られなかった。

指示の遵守と中間選択との関連

非遵守者のうち、指示された回答カテゴリではなく、中間カテゴリを選んだのは、指示項目1で77人中45人（58.4%）、2で69人中34人（49.3%）、3で56人中41人（73.2%）であった（Table 1）。5件法と7件法の違いがあるので単純に比較できないが、他の回答カテゴリは最大で18%しか選ばれておらず、非遵守者は指示項目で中間選択をしていることが多かった。

さらに、指示項目以外での中間選択傾向を検討したところ、5件法で回答を求めた74項目の平均中間選択数は、有非遵守者で40.7（ $SD = 18.8$ ）、全遵守者で25.0（ $SD = 11.6$ ）であり、有意な差があった（ $t(131.5) = 8.91, p < .001$, Hedgesの g （以下 g とする）= 1.24）。同様に7件法の14項目でも、有非遵守者の平均中間選択数が5.8（ $SD = 5.0$ ）であるのに対し、全遵守者で3.7（ $SD = 3.3$ ）と有意差が見られた（ $t(133.3) = 4.52, p < .001, g = 0.59$ ）。4件法での3項目に関しては、中央の2カテゴリの選択数で比較したが、有非遵守者（ $M = 2.4, SD = 1.1$ ）、全遵守者（ $M = 2.2, SD = 1.1$ ）で有意差はなかった（ $t(998) = 1.83, p = .067, g = 0.18$ ）。

指示の遵守と同一回答傾向との関連

項目文を注意深く読んでいる人でも、連続で同じ回答カテゴリを選ぶことがある。しかしそれが長く続くと、真の回答かどうか疑われることになる。そこで、指示項目の直前にどの程度続けて同一回答をしているかを確認した。本調査では、異なる尺度は異なるページに提示される。また尺度によって回答形式が異なっ

ていることから、回答行動が変わる可能性が高い。そのため、同一回答についての検討は、同じ尺度内に限定した。

指示項目1は、29項目が並べられている設問中の23番目に配置されている。すなわち、指示項目以前に22の項目がある。指示項目1の遵守者923人中、直前の20項目以上で連続して同じ回答値を選んだ人は12名（1.3%）、10項目以上では27名（2.9%）いた。しかし、この中で「1.非常にそう思う」を続けて選んでいた人はおらず、同一回答の一環として、指示された回答カテゴリを選んだことが疑われる人はいなかった。なお、20項目以上で同一回答をしていた12名中の10名（83.3%）、10項目以上で同一回答をしていた27名中の17名（63.0%）は、中間カテゴリを続けて選んでいた。

一方、指示項目1の非遵守者77名においては、20項目以上連続で同じ回答値を選んだ人が25名（32.5%）で、そのうち中間カテゴリを選んでいたのが20名（80.0%）であった。また10項目以上で同一回答をした者は37名（48.1%）で、中間カテゴリを選んでいたのはその中の27名（73.0%）であった。

指示項目1の前での、同一回答の連続数別（22—3）に見た、同一回答者数の割合を非遵守者と遵守者別に算出した。Figure 1は、横軸に示された数以上の連続回答者の比率（累積比率）である。すると、遵守者でも10以下になると続けて同じ回答をする者が増えるが、非遵守者の方が遵守者よりも同一回答がはるかに多かった。

指示項目1の直前かどうかに限らず、22項目中で最大どれくらい連続して同一回答をしていたのかを比較したところ、遵守者の中央値は1.4、非遵守者は6.0で、Wilcoxonの順位和検定で有意であった（ $W = 36042.5, p < .001$ ）。さらに指示項目1の後に続く6項

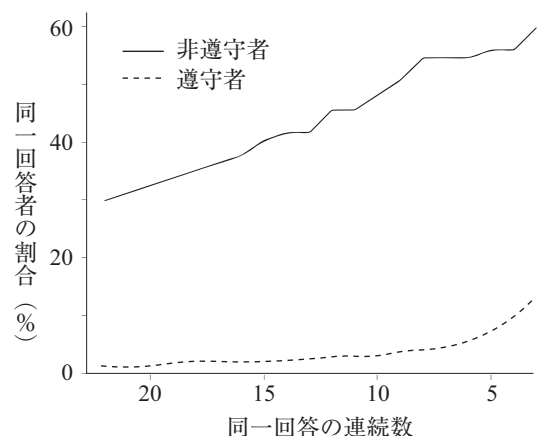


Figure 1. 同一回答の連続数と同一回答者の割合（指示項目1、累積比率）。

目に関しては、遵守者の 49 名 (5.3%) が同一回答で、その中の 30 名 (61.2%) が中間選択、非遵守者の 45 名 (58.4%) が同一回答で、その中の 32 名 (71.1%) が中間選択であった。指示項目を除く 28 項目での回答者内分散 (同一回答者における回答値のばらつき) の平均は、遵守者で .684、非遵守者で .340 であった。

指示項目 2 は、7 項目中の 4 番目に位置している。そこで指示項目 2 以前の 3 項目の回答について検討したところ、すべて同じ回答カテゴリを選んでいたのは遵守者 931 名中 220 名 (23.6%) で、その中で中間選択をしていたのは 68 名 (30.9%) であった。さらにこの 220 名の中で、「7. 非常にそう思う」を続けて選んでいた人は 11 名であった (5.0%)。すなわち指示項目 2 の指示に対し、同一回答によって見かけ上遵守したことが疑われる人は、やはり少なかった。非遵守者 69 名においては、3 項目のすべてで同一回答をしていた者は 41 名 (59.4%) で、その中の中間選択者は 24 名 (58.5%) であった。指示項目 2 以外の 7 項目での回答者内分散の平均は、遵守者で 1.820、非遵守者で .977 であった。

最後に指示項目 3 で同様の検討をしたところ、遵守者 944 名中、直前の 10 項目以上で同一回答をしていたのは 36 名 (3.8%) で、その中の 33 名 (91.7%) が中間選択をしていた。また同一回答者の中で、「1. 全く違う」を 10 項目以上続けて選んでいた人は 1 名であった。すなわち指示項目 3 でも、同一回答によって指示を遵守したことが疑われる人はほとんどいなかった。また非遵守者 56 名中 10 項目以上での同一回答者は 33 名 (58.9%) で、その中の中間選択者は 32 名 (97.0%) であった。指示項目 3 以外の 38 項目での回答者内分散の平均は、遵守者で 1.104、非遵守者で .571 であった。指示項目 3 で中間選択者の割合がより高くなっていったものの、非遵守者に同一回答が多く、また同一回答の多くが中間選択によって生じていて、回答者内分散が小さいという傾向は一貫していた。

全遵守者と有非遵守者の回答の違い

同一回答や中間選択が多いということは、非遵守者が指示項目以外でも不適切な回答をしていることを示唆している。それではこのような回答者を分析に含めることで、結果はどうなるであろうか。

Table 2 に、逆転項目を含む尺度での逆転項目と通常項目の合計との相関を示した。逆転項目が 2 つある場合は、それらの相関が正であった (.312 — .615) ことから、合計得点との相関係数を求めた。逆転項目は通常項目と負の相関を示すことが期待されるが、有非遵守者では正であったり、負であっても全遵守者より値が小さかったりすることが多かった。

次に逆転項目の得点を通常項目と同じ方向になるよう処理した上で、尺度の内部一貫性を検討した。こう

した目的の際に、最もよく用いられるのは α 係数であるが、 α 係数は不当に低く見積られることがあることから、因子分析モデルから算出される ω 係数も示した。すると、全遵守者だけで値を算出したときの方が、有非遵守者だけの値よりも概ね高く、かつ逆転項目を含む尺度でその差が大きかった。また、逆転項目を除いた通常項目だけだと、特に有非遵守者で値が高くなり、全遵守者を上回ることもあった。

通常項目と逆転項目との関連の仕方が異なることから、全遵守者と有非遵守者とは、尺度の因子構造が異なっているものと思われる。そのことを確認するために、R の lavaan パッケージ (0.5-19) を用いて、確認的因子分析を実施した (Rosseel, 2012)。本研究では 4 つの 1 次元尺度 (組織コミットメント、病棟への根付き、病院への根付き、患者ストレス) と、2 つの多次元尺度 (職務満足感、AWS) が含まれているが、複数の下位尺度がある場合は、すべての因子間に相関があるものとした。また識別条件を満たすために、因子の分散を 1 に固定した。

その結果、全遵守者では CFI が .772 — .895、TLI が .658 — .881、RMSEA が .252 — .062、SRMR が .053 — .116、有非遵守者では CFI が .596 — .884、TLI が .544 — .851、RMSEA が .268 — .099、SRMR が .070 — .134 となり、どの尺度のどの指標においても、全遵守者の方が有非遵守者よりも適合していた。

因子負荷 (標準化解) に関しても、全遵守者ではすべての通常項目で、正の有意な値であったと同時に、12 の逆転項目のすべてで、想定されていた通りに有意な負の値 (-.669 — -.276) を示した (有意水準はすべて 0.1%)。一方で有非遵守者では、通常項目に関してはすべて 0.1% 水準で有意な正の値であったが、逆転項目が通常項目と同じ正の値になっていたり、負であっても常に全遵守者より値が小さく、多くは 0 に近かった (-.400 — .491)。有意な推定値を得たのは 12 項目中 3 項目あったが、そのうちの 2 つは通常項目と同方向の正の値であった。すなわち、有非遵守者では、ほとんどの逆転項目で期待されているような値を示しておらず、全遵守者の方が理論的内容的に想定されているのと一致した因子構造を示していた。

因子構造が同じであるかどうかについて、さらに多母集団同時解析を実施して確認した。多母集団同時解析では、複数の母集団において因子構造が同じであることを示す配置不変モデルを基本とし、その上で因子負荷、観測変数の切片や誤差分散が等しいとの制約をおくなどして、測定不変であるかどうかを検証する。しかし、6 つの尺度に関する配置不変モデルの結果は、CFI が .765 — .894、TLI が .648 — .864、RMSEA が .071 — .266、SRMR が .052 — .118 となり、RMSEA や SRMR において良好な値を示すことはあるものの、複数の指標を鑑みると適合しているとは言い難かった。

Table 2
各尺度の α 係数及び逆転項目と通常項目との関係

	項目数 (逆転項目数)	全体	全遵守者	有非遵守者
仕事肯定感	11 (1)	.91 (.92)	.92 (.92)	.86 (.91)
		.91 (.92)	.92 (.92)	.89 (.91)
		-.37	-.45	.36
好労働環境	6 (0)	.90	.90	.87
		.90	.90	.87
上司支援	6 (0)	.81	.81	.84
		.82	.82	.85
存在意義	5 (1)	.68 (.73)	.69 (.73)	.53 (.81)
		.72 (.76)	.72 (.75)	.69 (.82)
		-.18	-.24	.42
組織コミットメント	8 (1)	.91 (.91)	.91 (.91)	.90 (.92)
		.91 (.91)	.91 (.91)	.91 (.92)
		-.57	-.60	-.32
病棟への根付き	7 (1)	.79 (.77)	.79 (.77)	.74 (.83)
		.80 (.78)	.80 (.78)	.78 (.84)
		-.46	-.53	.17
病院への根付き	7 (1)	.79 (.78)	.79 (.77)	.79 (.84)
		.80 (.79)	.80 (.79)	.81 (.85)
		-.43	-.47	-.06
仕事の負担	5 (2)	.73 (.75)	.74 (.76)	.31 (.64)
		.74 (.76)	.76 (.76)	.43 (.65)
		-.42	-.46	.06
裁量権	4 (0)	.86	.86	.85
		.86	.86	.85
報酬	4 (2)	.78 (.76)	.81 (.78)	.38 (.54)
		.78 (.80)	.81 (.82)	.40 (.55)
		-.48	-.51	.01
共同体	5 (1)	.87 (.89)	.88 (.89)	.73 (.83)
		.88 (.89)	.89 (.90)	.78 (.84)
		-.51	-.55	.09
公平性	6 (2)	.74 (.73)	.75 (.72)	.50 (.69)
		.75 (.73)	.75 (.73)	.57 (.70)
		-.45	-.48	-.02
価値観	4 (0)	.79	.79	.72
		.79	.80	.73
患者ストレス	10 (0)	.93	.93	.91
		.93	.93	.91

注) 一段目：逆転項目の逆転処理後 α 係数 (逆転項目除去後 α 係数)。

二段目：逆転項目の逆転処理後 ω 係数 (逆転項目除去後 ω 係数)。

三段目：逆転項目 (2 項目のときは合計, 逆転未処理) と通常項目合計との相関係数。

(Hooper, Coughlan, & Mullen, 2008)。したがって、全遵守者と有非遵守者とで、各尺度の因子構造が同じであるとは言えない。

最後に、有非遵守者に中間選択が多いのであれば、素点合計得点の平均値も、その尺度が取りうる得点の中心値に近づき、ばらつきが小さくなることが予想される。すると、全 14 尺度中 11 尺度で、また有意差の

あった 9 尺度ではすべてで、有非遵守者の平均値の方が全遵守者よりも中心値に近く、標準偏差が小さかった (Table 3)。ただし、必ずしも逆転項目を含んだ尺度で有意差が見られたわけではなく、通常項目だけの合計得点の方が、効果量 (g) が大きいこともあった。

通常項目だけの合計得点も、9 尺度中 7 尺度で有非遵守者の平均値の方が中心値に近かった。また有非

Table 3
各尺度における平均値と中心値との差

	尺度得点の 中心値	全体	全遵守者	有非遵守者	全遵守者と 有非遵守者の差	
		平均 (SD)	平均 (SD)	平均 (SD)	<i>t</i>	<i>g</i>
仕事肯定感	33	-0.17 (7.52)	-0.11 (7.72)	-0.61 (5.80)	0.84	0.07
	30	-0.85 (7.11)	-0.83 (7.24)	-1.00 (6.05)	0.28	0.02
好労働環境	18	-2.22 (4.83)	-2.38 (4.93)	-1.07 (3.88)	-3.37 ***	0.27
上司支援	18	-1.52 (4.41)	-1.64 (4.49)	-0.64 (3.69)	-2.70 **	0.23
存在意義	15	-0.71 (3.30)	-0.75 (3.40)	-0.40 (2.40)	-1.42	0.11
	12	-1.11 (2.97)	-1.17 (3.00)	-0.65 (2.63)	-1.81	0.18
組織コミットメント	24	-2.31 (7.04)	-2.49 (7.18)	-1.00 (5.77)	-2.57 **	0.21
	21	-2.81 (6.30)	-3.01 (6.40)	-1.34 (5.37)	-3.12 ***	0.27
病棟への根付き	28	-0.45 (7.08)	-0.55 (7.24)	0.30 (5.70)	-1.47	0.12
	24	-0.60 (6.30)	-0.71 (6.36)	0.14 (5.78)	-1.39	0.14
病院への根付き	28	-1.29 (7.10)	-1.40 (7.23)	-0.48 (6.08)	-1.52	0.13
	24	-1.43 (6.35)	-1.53 (6.41)	-0.65 (5.86)	-1.43	0.14
仕事の負担	15	2.15 (3.66)	2.36 (3.74)	0.55 (2.42)	7.12 ***	0.50
	9	0.81 (2.65)	0.86 (2.72)	0.39 (2.02)	2.28 *	0.18
裁量権	12	-0.59 (3.47)	-0.71 (3.55)	0.28 (2.65)	-3.67 ***	0.29
報酬	12	-0.01 (2.60)	0.00 (2.71)	-0.07 (1.58)	0.41	0.03
	6	0.18 (1.48)	0.17 (1.52)	0.19 (1.12)	-0.18	0.01
共同体	15	1.77 (3.66)	1.87 (3.79)	1.04 (2.36)	3.31 ***	0.23
	12	1.55 (1.48)	1.64 (3.14)	0.93 (2.30)	3.02 ***	0.23
公平性	18	-1.45 (3.65)	-1.64 (3.75)	-0.07 (2.47)	-6.07 ***	0.43
	12	-1.34 (2.74)	-1.51 (2.77)	-0.07 (2.11)	-6.73 ***	0.53
価値観	12	-0.98 (2.76)	-1.12 (2.84)	0.01 (1.88)	-5.75 ***	0.41
患者ストレス	30	7.53 (7.70)	8.02 (7.72)	3.92 (6.54)	6.30 ***	0.54

注) 上段: 逆転項目を含めた全項目の合計得点と中心値の差。

下段: 通常項目だけの合計得点と中心値の差。

*** $p < .001$, ** $p < .01$, * $p < .05$

守者を分析データに加えた全体でのデータでは、全遵守者よりも平均値が中心値に近づき、ばらつきが小さくなった。

考 察

結果のまとめと中間選択の影響

3つの指示項目における非遵守者は5.6—7.7%であった。指示の非遵守が疲労効果で生じるのであれば、質問票の後半で非遵守が増えることが予想される。しかし、後半になるにつれてむしろ非遵守者は減っていた。これは一度、特定の指示があることに気がつくと、その後の項目内容に気をつけるようになるためと思わ

れる。

また、黙従傾向や同一回答傾向の結果として、指示を遵守したように見える人はほとんどいなかった。一方で非遵守者には同一回答が多く、また中間選択がなされた結果、非遵守となっていることが多かった。さらに、非遵守者は指示項目以外でも中間選択が多かった。

評定尺度において、中間の回答カテゴリがよく選択されることは広く知られているが、中間選択は問われている内容に対して中立の態度や意見を示すときだけでなく、「わからない」、「関心がない」、「答えたくない」といった、いわば非回答の意味で生じることがある (Raaijmakers, van Hoof, 't Hart, Verbogt, & Vollebergh,

2000; Sturgis, Roberts, & Smith, 2014)。このような場合に中間回答を得点化すると系統誤差を含むことになるが、回答者がどちらの意味で中間を選んだのかを判別するのは難しい(増田・坂上, 2014)。しかし本研究の非遵守者は、項目文を精読しないで回答していると思われることから、中立の意味で中間カテゴリを選んだのではないだろう(ただし、逆に、遵守者によってなされた中間選択が、中立の意味であるとは限らない)。

非系統誤差は合計得点尺度間の相関係数を低下させるが、複数項目で尺度を構成すると、各項目の非系統誤差が相殺されることから、測定信頼性が高くなりやすい。そのため、心理尺度においては、多くの項目を用意することが望ましいとされてきた。しかし、多数の項目を合計すると、反応バイアスがもたらす系統誤差の影響が加算されて大きくなる可能性がある(Schuman & Kalton, 1985)。本研究でも、特に非遵守者で非回答の意味での中間選択が蓄積され、こうした回答者を含めた全データでの合計得点の平均が、中心値に近づいたものと思われる。

このように、得点がある方向に歪めることから、尺度の妥当性を低めるにもかわからず、中間選択のような系統誤差が結果に影響を与えることはほとんど知られていないという(Baumgartner & Steenkamp, 2001)。しかしながら、このことは適切な調査研究を実施するために留意されなければならない重大な問題である。

逆転項目への回答

心理尺度にはしばしば、概念と一致する方向からなる項目だけでなく、反対方向で表現された逆転項目が設けられる。逆転項目には、概念内容を反映する広範囲の項目が含まれることになって尺度の妥当性が向上したり、黙從傾向等の影響を減らすことができるといったメリットがあるとされる(Paulhus, 1991)。一方で、逆転項目だけで別の因子が構成されたり、尺度の内部一貫性が低下したりするといった問題が生じることがある。例えば、自尊感情尺度(Self-Esteem Scale: Rosenberg, 1965)は、しばしば自尊感情と同方向のポジティブな状態を尋ねる項目と、ネガティブな状態を尋ねる逆転項目とが異なる因子に分離する(Motl & DiStefano, 2002)。

こうした結果が得られるのは、「満足していないからといって不満なわけではない」といったように、ポジティブな状態とネガティブな状態が対称でなく、別概念であるためかもしれない(Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Colston, 1999)。また先述したように、逆転項目では否定語の見過ごしのような不注意による回答ミスが生じるためかもしれない(Swain et al., 2008; Woods, 2006)。

本研究の全遵守者では、逆転項目を除くことで内部

一貫性が高くなることもあれば、低くなることもあった。しかしながら、有非遵守者では、逆転項目を除くと常に内部一貫性が高くなった。これは、逆転項目が通常項目と負の相関があっても小さいか、ときに正の相関が見られたためと思われる。確認的因子分析でも、有非遵守者では逆転項目が想定していたような因子負荷を示さず、適合度が低かった。Woods (2006)はシミュレーションの結果から、23項目中10項目が逆転項目であるような1次元尺度で、回答者の10%以上が逆転項目で回答ミスをする、2因子解を見出すようになる、と述べている。本研究で用いた尺度は逆転項目が少ないため、こうした点の詳しい検討はできなかった。しかし、逆転項目と通常項目の相関係数や確認的因子分析の因子負荷量から、非遵守者のような回答者が多いと、本来1次元であるような尺度であっても、逆転項目が異なる因子として抽出されることがあるものと思われる。

では、逆転項目では回答ミス等の問題が起きやすいが、通常項目では適切に回答されているのだろうか。しかし、非遵守者では逆転項目かどうかにかかわらず、同一回答や中間選択が多かった。そして有非遵守者の通常項目だけの合計得点も、全遵守者より中心値に近いことが多かった。したがって、非遵守者は通常項目でも項目文を精読しておらず、同一回答や中間選択のために項目間相関が高くなって、高い内部一貫性や因子負荷が見られるようになったものと思われる。

逆転項目があることで内部一貫性が低くなったり、逆転項目が想定されたような因子負荷を示さなかったりすると、その項目を除いて尺度を構成することが多い。しかしながら、こうした結果の原因は必ずしも逆転項目自体にあるとは限らないのである。さらには同じ尺度を用いた異なる研究で、因子構造が一致しないといった結果が得られたとしても、回答者集団ごとに異なる因子構造があるのではなく、非遵守者のような回答者の割合が違っているためである、といったことが起こりうると思われる。こうした尺度の内部構造については、尺度内の項目数とその中に含まれる逆転項目数を系統的に操作した実験をすることで、詳細に検討する必要があるだろう。

また、Marsh (1996)やTomás & Oliver (1999)では、因子分析で2因子解が得られることのある自尊感情尺度において、逆転項目に特有の回答傾向があることを想定したモデルで確認的因子分析を実施すると、それが無い場合よりも適合度が高くなる他、通常項目と逆転項目のどちらへも負荷が高い1つの因子が認められるようになることを示した。すなわち、2因子解が得られるのは、ネガティブな状態を尋ねる項目に特定の反応バイアスがあるためだというのである。このように反応バイアスの影響を推定して、内部構造の検討をすることも試みるべきであろう。

今後の展望と本研究の限界

非遵守者には、同一回答や中間選択が多いことから、指示項目を設けて非遵守者を見出すことで、分析データから結果を歪めるような回答をかなりの程度排除できるものと思われる。しかしながら、不適切な回答者を除くよりも、できるだけ多くの回答者に真摯な回答をしてもらう方がより望ましい。この点については非遵守と中間選択の関連が強いことから、特に中間選択の発生の観点から検討をしていくことが有益であろう。増田・坂上 (2014) は、中間選択の増減に影響する主要因として、項目文や回答形式の曖昧性や困難性の要素について整理した。これらが中間選択そのものの増減だけではなく、真摯な回答と関連するのかどうか、言い換えるならば質問項目や回答形式等の要因によって、非遵守者を減らすことができるのかどうかを検討されなければならない。

本研究の限界として、真の意見や態度を反映しない回答としての、同一回答以外の規則的な回答 (例えば、1 から 5 までの選択肢の選択を繰り返す) や、ランダムな回答について検討しなかったことがある。ただし、非遵守者において同一回答や中間選択が非常に多かったことから、これら以外の不適切な回答は少なかったものと思われる。

次に、回答者は項目文を読んでも、その上で本心と異なる回答をすることがあり、特に見せかけ (faking) 回答や社会的望ましさ (social desirability) によるバイアスが知られている。しかし、指示項目では項目文を読んだ上でこうした回答をするような回答者を見つけることはできない。本研究においても、遵守者の中に指示項目 1 の直前で 10 以上の同一回答をした者が 27 名いたが、項目を読んで指示に従っているものの、その他の項目では嘘やいい加減な回答をしていたという可能性がある。

こうした回答者は、虚偽尺度 (lie scale) を用いて検出が試みられてきた (岩脇, 1973)。虚偽尺度では、通常はありえない回答をするかどうかで、不適切な回答を見出そうとする。しかし実際には、虚偽尺度への回答に、項目文を読んだ上で見せかけの回答をした場合と、読まずにたまたまそのような回答をした場合の両方が含まれているものと思われる。したがって、虚偽尺度と指示項目の関係を明らかにすることも今後の研究課題となろう。

さらに、Web 調査の方が最小限化が促進されやすく、対面式のインタビュー調査よりも中間選択率が高いという指摘がある (Heerwegh & Loosveldt, 2008)。また三浦・小林 (2015) は、調査会社によって最小限化の生じやすさが異なるという、House effect があることを報告している。1 社での Web 調査だけの検討のため確認はできないが、こうしてみると本調査で特に最

小限化が強く生起し、そのために特有の回答が得られたという可能性がある。

しかしながら、どのような調査方法であっても、回答者は必ずしも真摯に回答してくれるとは限らず、調査者は常に、非遵守者のような回答者の影響に留意する必要があるだろう。例えば、Aust, Diedenhofen, Ullrich, & Musch (2012) は、Web 調査における不真面目回答 (nonserious response) を検知するために、本論文で触れた IMC や矛盾回答の他に、回答時間の長短や、単刀直入にまじめに回答したかどうかを尋ねるといった方法があることを記している。このような様々な手段を用いたり、その有効性を検討したりすることで、調査データの質の向上のための知見を得ていかなければならない。

引用文献

- Aust, F., Diedenhofen, B., Ullrich, S., & Musch, J. (2012). Seriousness checks are useful to improve data validity in online research. *Behavior Research Methods*, 45, 527-535.
- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, 5, 323-370.
- Baumgartner, H., & Steenkamp, J. (2001). Response styles in marketing research: A cross-national investigation. *Journal of Marketing Research*, 38, 143-156.
- Colston, H. (1999). "Not good" is "bad," but "not bad" is not "good": An analysis of three accounts of negation asymmetry. *Discourse Processes*, 28, 237-256.
- Crossley, C. D., Bennet, R. F., Jex, S. M., & Burnfield, J. L. (2007). Development of a global measure of job embeddedness and integration into a traditional model of voluntary turnover. *Journal of Applied Psychology*, 92, 1031-1042.
- Heerwegh, D., & Loosveldt, G. (2008). Face-to-face versus web surveying in a high-internet-coverage population: Differences in response quality. *Public Opinion Quarterly*, 72, 836-846.
- Herzog, A. R., & Bachman, J. G. (1981). Effects of questionnaire length on response quality. *Public Opinion Quarterly*, 45, 549-559.
- Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modelling: Guidelines for determining model fit. *Journal of Business Research Methods*, 6, 53-60.
- 岩脇 三良 (1973). 心理検査における反応の心理 日本文化科学社
- 北岡 和代・増田 真也・森河 裕子・中川 秀昭 (2015). 日本版 Areas of Worklife Survey (AWS) ——個人と職場環境 6 つのミスマッチ診断—— 経営行動科学, 28, 53-63.
- Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifacts? *Journal of Personality and Social Psychology*, 70, 810-819.

- 増田 真也・坂上 貴之 (2014). 調査の回答における中間選択——原因, 影響とその対策—— 心理学評論, 57, 472-494.
- 三浦 麻子・小林 哲郎 (2015). オンライン調査モニタの Satisfice 行動に関する実験的研究 社会心理学研究, 31, 1-12.
- Motl, R., & DiStefano, C. (2002). Longitudinal invariance of self-esteem and method effects associated with negatively worded items. *Structural Equation Modeling*, 9, 562-578.
- 撫養 真紀子・勝山 貴美子・青山 ヒフミ (2014). 病院に勤務する看護師の職務満足測定尺度の信頼性・妥当性の検討 社会医学研究, 31, 37-44.
- Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, 45, 867-872.
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). San Diego, CA: Academic Press.
- Peltokorpi, V., Allen, D. G., & Froese, F. (2015). Organizational embeddedness, turnover intentions, and voluntary turnover: The moderating effects of employee demographic characteristics and value orientations. *Journal of Organizational Behavior*, 36, 292-312.
- Raaijmakers, Q. A. W., van Hoof, A., 't Hart, H., Verbogt, T., & Vollebergh, W. (2000). Adolescents' midpoint responses on Likert-type scale items: Neutral or missing values? *International Journal of Public Opinion Research*, 12, 208-216.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48, 1-36.
- Schuman, H., & Kalton, G. (1985). Survey methods. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (Vol. 1, pp. 635-697). New York: Random House.
- Simon, H. A. (1957). *Models of man: Social and rational*. New York: John Wiley and Sons.
- Sturgis, P., Roberts, C., & Smith, P. (2014). Middle alternatives revisited: How the neither/nor response acts as a way of saying "I don't know"? *Sociological Methods & Research*, 43, 15-38.
- Swain, S. D., Weathers, D., & Niedrich, R. W. (2008). Assessing three sources of misresponse to reversed Likert items. *Journal of Marketing Research*, 45, 116-131.
- 高木 浩人・石田 正浩・益田 圭 (1997). 実証的研究——会社人間をめぐる要因構造—— 田尾 雅夫 (編)「会社人間」の研究——組織コミットメントの理論と実際—— (pp. 266-296) 京都大学学術出版会
- Tomás, J., & Oliver, A. (1999). Rosenberg's Self-Esteem Scale: Two factors or method effects. *Structural Equation Modeling*, 6, 84-98.
- Vannette, D. L., & Krosnick, J. A. (2014). Answering questions: A comparison of survey satisficing and mindlessness. In A. Ie, C. T. Ngunoumen, & L. E. Langer (Eds.), *The Wiley Blackwell handbook of mindfulness* (pp. 312-327). Chichester, West Sussex: John Wiley & Sons.
- Woods, C. M. (2006). Careless responding to reverse-worded items: Implications for confirmatory factor analysis. *Journal of Psychopathology and Behavioral Assessment*, 28, 189-194.