

|                  |   |
|------------------|---|
| Title            | ウェブ調査における回答者のニューメラシーと質問の種類による中間回答傾向の違い  |
| Sub Title        | Effects of respondents' objective numeracies and various kinds of alternative questions on response patterns in web-survey  |
| Author           | 広田, すみれ(Hirota, Sumire)   |
| Publisher        | 三田哲學會   |
| Publication year | 2019  |
| Jtitle           | 哲學 (Philosophy). No.142 (2019. 3) ,p.221- 243   |
| Abstract         | This study explored the differences in response patterns, especially of middle-alternative response patterns, by respondents' objective numeracies and the kinds of questions in a web-survey. A web-survey company recruited 1000 respondents whose age ranged from 20s to over-60s, and half were male and the other half, female. They completed the web survey that contained 5 alternative questions including Likert-scale, attitudes questions, and knowledge questions. Respondents' numeracies were measured using Weller et al.'s (2013) numeracy scale, and they were divided by median (ME = 5; high n = 625, low n = 375). Results showed that low-numeracy respondents chose middle alternatives significantly more frequently than did the high-numeracy group. Respondents also chose the middle-alternative more often in the Likert-scale than in the other two kinds of questions. These were mainly attributed to the positioning effect, but partially to the language label of the middle-alternative. Cognitive processes of satisficing effects in the survey were discussed. |
| Notes            | 特集：坂上貴之教授 退職記念号#寄稿論文  |
| Genre            | Journal Article   |
| URL              | <a href="http://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN00150430-00000142-0221">http://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN00150430-00000142-0221</a>   |

# ウェブ調査における回答者のニューメラシーと 質問の種類による中間回答傾向の違い

— 広 田 す み れ\* —

## **Effects of Respondents' Objective Numeracies and Various Kinds of Alternative Questions on Response Patterns in Web-survey**

*Sumire Hirota*

This study explored the differences in response patterns, especially of middle-alternative response patterns, by respondents' objective numeracies and the kinds of questions in a web-survey. A web-survey company recruited 1000 respondents whose age ranged from 20s to over-60s, and half were male and the other half, female. They completed the web survey that contained 5 alternative questions including Likert-scale, attitudes questions, and knowledge questions. Respondents' numeracies were measured using Weller et al.'s (2013) numeracy scale, and they were divided by median ( $ME=5$ ; high  $n=625$ , low  $n=375$ ). Results showed that low-numeracy respondents chose middle alternatives significantly more frequently than did the high-numeracy group. Respondents also chose the middle-alternative more often in the Likert-scale than in the other two kinds of questions. These were mainly attributed to the positioning effect, but partially to the language label of the middle-alternative. Cognitive processes of satisficing effects in the survey were discussed.

---

\* 東京都市大学メディア情報学部

## 1. はじめに

従来、質問紙調査で複数選択肢の中から1つを選ぶような場面では、調査者が想定している選択肢の意味だけでなく、それ以外の要因による複数のバイアスの存在が知られている。その例として、選択肢のうちでの中間回答 (middle alternative) あるいは極端選択肢への選好 (extreme responding) か、内容に関係なく質問に常に肯定的に回答するという黙従傾向 (acquiescence)、「意見なし」選択肢の有無による影響 (no-opinion filter) といったものが広く知られる (Schuman, Kalton, and Ludwig, 1983)。このような傾向は複数の研究で指摘されており (例えば林, 1996; レビューは増田・坂上, 2014)、近年ではウェブ調査を利用し、質問紙の後半になると中間回答が増える現象 (増田・坂上・北岡, 2017; 増田・坂上, 2014) や、郵送調査とのモードによる比較では、ウェブ調査では4択のうちで2列目に表示され中間で肯定的選択肢である「まあ当てはまる」の回答割合が高い (萩原・村田・吉藤・広川, 2018) など、知見が増えている。

このうち中間選択傾向に関しては、回答者の教育水準による影響が指摘されている。Schuman and Presser (1981) は教育水準の一貫した影響を見出さなかったとしたが、Narayan and Krosnick (1996) は実験により、態度設問での中間回答傾向について2つの異なる傾向を明らかにした。彼らは実験条件として態度に関する設問で中間選択肢の有無を操作した。例えば政治的態度の設問では「中庸」の選択肢の有無、また「あなたの意見では、マリファナを使用した時の罰則はより厳しくすべきでしょうか、より軽くすべきでしょうか、または現在と同じ位にしておくべきでしょうか?」という設問では「現在と同じ位」(現状維持)の有無を操作した。その結果、全体としては教育水準の高い者よりも中程度または低い者で選択肢の追加の効果が有意に大きく、特に単純な中庸の選択肢の追加による効果は後者で大きかった。このことから Narayan and Krosnick (1996) は他の反応バイアスも検討したうえで教育水準による影響を強調し、同時にこの

ことは回答行動で中間選択肢を選ぶという反応バイアスが最小限化 (satisficing) によるものであるとする見解と一貫している, とした。

最小限化モデルは Krosnick ら (Krosnick and Alwin, 1987; Krosnick, 1991) による, 認知負荷の大きさにより回答行動に反応バイアスが生じるとするモデルである。回答行動における最小限化の影響は複数指摘されている (例えば Tourangeau, Conrad, and Couper (2013), Oppenheimer, Meyvis, and Davidenko (2009), 三浦・小林 (2016a, 2016b))。しかし最小限化が具体的にどんなヒューリスティックや認知的メカニズムと関わっているのかの手がかりになる研究はまだ限られている。その点を検討している Tourangeau, Couper, and Conrad (2004) は, 人々が回答行動でとる5つのヒューリスティックを提案した。すなわち 1) 中央は典型的, 2) 左とトップに位置するものは概念的にも「第一番目」, 3) 近いものは関連している, 4) 縦のリストでは上の選択肢はもっとも望ましい, 5) 視覚的に似ているものは概念的にも近い, とし, この証拠として実験的に選択肢の順序やデザインを変え, 回答行動の変化を検討した。最初の2つの実験では NO や DK と, 通常の内容が記載された設問との間に線を引くか否かで視覚的な中央の位置を変化させて効果をみた, 続いて実験3では選択肢のボタン間の距離を変化させて視覚的位置を変えた。その結果, いずれの場合にも視覚的に中央に位置する選択肢への反応が高いことから, 回答が視覚的要素に影響されていることを指摘した。さらに選択肢の順序を変化させた場合にも中央に位置する選択肢への反応時間が速いこと (実験4), 7つの選択肢のうちで同一の選択肢が3番目と7番目に位置した場合にも選択率が変化すること (実験5) やグループ化した時の効果を検討した (実験6)。これらの実験は, さまざまな方法で視覚的中央に位置することが影響することを明らかにした。

自明のことではあるが, 紙による質問紙調査でもウェブ調査でも, 程度の差はあれ最小限化が影響したとされる回答行動, たとえば同一選択肢を

選り続けるような現象は共に存在していた。ただ紙による質問紙調査ではその現象自体を検討するというより、現実の調査の実施や分析上の対処法、すなわちほぼ一貫して同じ選択肢を選り続けた回答者を不良回答者として任意に除く、といったことに力点が置かれていたと推測される。それは、質問紙による調査では回答中の行動について測定したり、途中段階に実験的介入できる余地が少なかったことが一因であろう。しかし、ウェブ調査では反応時間やマウスの動きのようなパラデータの取得が可能になったことや、調査票のデザイン変更が比較的容易であることによって介入が容易になったことから、前述の Tourangeau et al. (2004) のように、詳細に検討できる余地が出てきた。

しかし紙での調査でも、そういった回答行動のバイアスについて認知的メカニズムとの関連を間接的に推測する研究がなかったわけではない。中間回答や極端反応傾向と人口動態属性など回答者属性との関係の検討はその1つである。これらと回答者の所属する文化やそこでのコミュニケーションスタイルとの関係を指摘するものがよく知られるが、他にも認知能力、特に読解力による影響 (Messick and Frederiksen, 1958)、国家間での比較研究 (マクロレベルでの研究) では IQ や教育水準による影響も指摘されている (例えば Rammstedt, Goldberg, and Borg, 2010)。文化と結びついた解釈では、前述の選択傾向を「文化アイデンティティ」「社会的望ましさ」「価値観」の影響とする。一方、教育水準や IQ による違いは前述の傾向を主に認知能力に帰属するもので、Narayan and Krosnick (1996) での教育水準の影響の指摘と同様に、回答行動のバイアスが認知的メカニズムと関連していることを間接的に支持するものとみられることもできる。

そこで本研究では最小限化に関わる認知能力のより詳細な検討をするため、IQ や教育水準から一歩踏み込んで、現在意思決定研究で注目されているニューメラシーを個人差要因として導入した。ニューメラシー

(numeracy) は本来英語の普通名詞で数的能力を指し、日本語で言う、「読み書き算盤」のうちの算盤部分に当たる。意思決定研究では、主に医療意思決定 (medical decision making) で患者の意思決定の違いを説明する目的から、2000年前後から尺度を作成し、その尺度に基づいて意思決定の個人差を検討する研究が行われてきた。医療意思決定でニューメラシーが特に注目された理由は、合意形成や双方向性が重要視される他の分野でのリスクコミュニケーションに比べ、送り手の医療者の専門性が重視され、送り手による数値を用いた科学情報の提供を受け手に理解してもらうことが重視されているためと推測される。そして結果として、受け手のニューメラシーにより数量を含んだ情報への判断が異なるという研究報告が複数行われている (例えば Peters, Västfjäll, Slovic, Mertz, Mazzocco, and Dickert, 2006)。

ウェブ調査で設問に回答するという場面は意思決定場面と共通性があり、実際、Tourangeau, Rips, and Rasinski (2000) はその著書 "The psychology of survey response" の5章で、事実に関する設問、特に頻度に関する質問 (例、「あなたが成長期のうち、あなたの父親はどのくらい頻繁に宗教的な会に参加していましたか?」) については Tversky and Kahneman (1973, 1974) の指摘した3つのヒューリスティックの関与を論じている。ただ、本研究では前述のような数量的情報を受け取る時の意思決定における数量的情報処理過程への影響ではなく、送り手として、回答者が複数の選択肢を用いて回答するときの影響として注目する。具体的にいうと、ニューメラシーが低いと、数値を幅広く使って自分の心理的強度を表出すること自体に制約がある可能性がある。特にリッカート尺度では、様々な事象に関する内観での個人の心理的強度を数値に置き直して表出することが前提とされ、このような状況ではニューメラシーの高低が影響する可能性は大きいと考えられる。実際、広田 (2014) は食品リスク等に関する110問の4段階のリッカート尺度に対する選択行動について、

ウェブ調査における回答者のニューメラシーと質問の種類による中間回答傾向の違い

Lipkus, Samsa, and Rimer (2001) のニューメラシー尺度の日本語版の得点の高低で回答者を分類して回答行動を比較した結果、得点の低い回答者では最も negative な選択肢（「4. 全くそう思わない」）の忌避が得点の低い回答者よりも有意に多いことを明らかにした。このようにニューメラシー低群で極端選択が忌避されることから見ると、ニューメラシーが回答行動に関わっている可能性は高い。

そこで本研究はまず客観ニューメラシーによる回答行動の違いを検討したが、加えて認知的能力や認知過程との関係を探索的に検討する目的で、性質が異なる複数の選択肢を用いた。具体的には同じ5つの選択肢であるが、「あてはまる」から「あてはまらない」までのようなカテゴリー間に順序性のあるリッカート尺度、順序性がなく、どのような選択も可能な態度に関するカテゴリー尺度、さらに順序性はないものの、特定のカテゴリーのみが「正しい」正解のある知識問題を用い、この3種類で中間回答の選択率に違いがあるかを検討し、中間選択傾向の背景となる認知的メカニズムを探索的に検討することを目的とした。さらに、視覚上の位置の影響を考慮して、リッカート尺度に関しては Tourangeau, Couper, and Conrad (2004) の用いた通常とは選択肢の位置の異なる条件を導入し、視覚的手がかりがニューメラシーレベルの違う回答者にどう影響するかを検討した。

## 2. 方法

ウェブ調査による。調査では、1) リッカート尺度の設問 24 問、2) 社会調査等で用いられてきた順序性のない態度に関する 5 択の設問 10 問、3) 正解のある知識問題 12 問、を尋ねた。リッカート尺度はユーモア態度尺度（上野, 1993; 宮戸・上野, 1996）の 24 項目で、「1. あてはまらない」「2. あまりあてはまらない」「3. どちらでもない」「4. ややあてはまる」「5. あてはまる」の 5 段階である。態度に関する質問は、内閣府国民生活

選好度調査（内閣府大臣官房政府広報室，2016）並びに JGSS（日本版総合的社会調査）（大阪商業大学 JGSS 研究センター，2016）で過去に用いられた価値観に関する設問で，回答の選択肢に順序性がなく名義尺度とみなせるもので，各調査の選択肢から「その他」を削って5択にするなど問い自体は変化しない範囲で一部改変した．また正解のある問題は大学生の就職試験の問題（SPI ノートの会，2016a, 2016b）から難易度が適度に高く，5 選択肢で，またニューメラシーと明らかに関連しているような数的な問題以外の言語，推論（文字列の規則性），論理問題をわずかに改変した問題に，Wason（1966）の4枚カード問題の具体的な場面の問題を加えた．1)～3) は，各グループ内の設問順序は同一だが，グループ間の提示順序を条件によって入れ替え，提示順序による効果がないようにした．

さらにリッカート尺度は1～5の選択肢に順序性があり，2) 3) とは異なり視覚的位置と順序が一致していることが回答行動に影響する可能性がある．そこで，本研究ではリッカート尺度については①通常の1～5の順序での提示（通常条件）（実際の提示の例は図1参照）に加えて，②中間左条件（「どちらでもない」が一番左側に提示され，続いて「あてはまらない」「あまりあてはまらない」…と続く．よって中央には，「あまりあてはまらない」が位置する．③中間右条件（「どちらでもない」が一番右側に提示される．よって中央には「ややあてはまる」が位置する．④ランダ


| <br>回答方向 |                      | あてはまらない | どちらでもない | ややあてはまる | あてはまる |    |
|---|----------------------|---------|---------|---------|-------|----|
| 1   | 笑いには多少毒があった方がおもしろい   | 1○      | 2○      | 3○      | 4○    | 5○ |
| 2   | 友人を軽く皮肉ったりして楽しむことがある | 1○      | 2○      | 3○      | 4○    | 5○ |
| 3   | 過激な冗談が好きだ            | 1○      | 2○      | 3○      | 4○    | 5○ |
| 4   | ブラックユーモアが好きだ         | 1○      | 2○      | 3○      | 4○    | 5○ |
| 5   | きついことを言って人を笑うのは嫌いだ   | 1○      | 2○      | 3○      | 4○    | 5○ |

Fig. 1. Screen shot of Likert scale questions.



ウェブ調査における回答者のニューメラシーと質問の種類による中間回答傾向の違い

ム条件（1～5の選択肢が実験参加者によってランダムな順序で提示される）、⑤スペース条件（1～5になるにつれ、選択肢間の感覚が徐々に広がる）。なお、いずれの場合も図1のように、画面には言語ラベルのみが提示され、数値が振られているわけではないので、回答者は数値を手掛かりにすることはできない。ニューメラシーについては以上の設問の前に属性に関する質問に続いて尋ねたが、客観的なニューメラシーを測定するWeller, Dieckman, Tusler, Mertz, Burns, and Peters (2013) による短縮版ニューメラシー尺度（9問）と認知的熟慮テスト（Frederick, 2005）、主観的ニューメラシーテスト（SNS: Fagerlin, Zikmund-Fisher, Ubel, Jankovic, Derry, and Smith, 2007）を用いた。ここでは客観ニューメラシーテストに基づく分析結果のみを報告する。

ウェブ調査は2018年2月下旬に(株)クロス・マーケティングに委託して行われた。回答者は同社のモニターのうち、一都三県に居住者する男女1000名で、20代、30代、40代、50代、60代以上の5層について各男女50名ずつである。なお、リッカート尺度に関する5条件は10層の各層に均等に5条件が割り振られるようにし、また3つの尺度群の呈示順序も同様にした。

### 3. 結果

#### (1) ニューメラシー尺度に関する回答

回答者についてWeller et al. (2013) のニューメラシー尺度の得点を算出したところ、9点満点で平均は4.99点、中央値5点、標準偏差2.35となった。そこでニューメラシー高群は5点以上、低群は4点以下とした。結果として高群が625名、低群が375名となり、これを以降の分析に用いた。

(2) 設問群別 中間回答の割合

各設問群及びリッカート尺度の各条件での中間回答の割合を示したのが図2である。なおカテゴリ尺度の2つの設問は全員が回答しているが、リッカート尺度は回答者全体をランダムに5条件に割り付けたため、データ数が相対的に少ない(通常条件  $N_H=133, N_L=67$ ; 中間左条件  $N_H=110, N_L=90$ ; 中間右条件  $N_H=125, N_L=75$ ; ランダム条件  $N_H=123, N_L=77$ ; スペース条件  $N_H=134, N_L=66$ )。

図2から明らかなように設問群別にみると、リッカート尺度のランダム条件(高群23.0 低群23.1, 有意差なし)以外はいずれもニューメラー低群(L)の方が高群(H)よりも中間選択の割合が有意に多かった。ランダム条件は視覚的位置の手掛かりの乏しいことが影響したとみられるが、このことから見ると、設問の種類に関わらず回答者の客観ニューメラーにより中間回答選択率は異なり、低群の方が高いと言える。

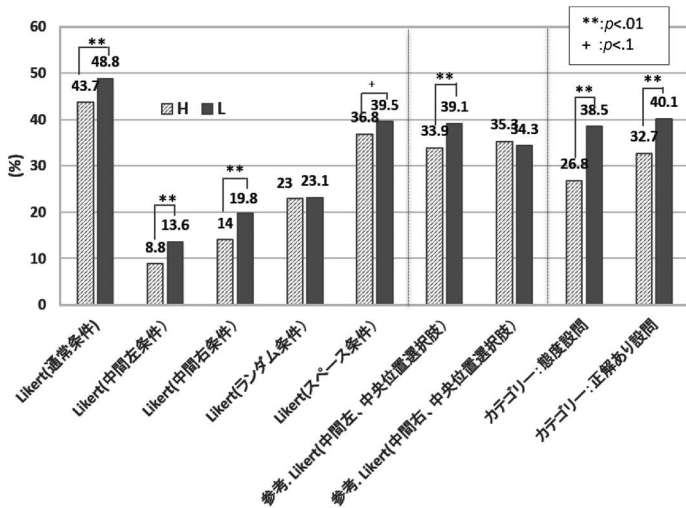


Fig. 2. Percentages of the participants who chose the mid-category in various questions

ウェブ調査における回答者のニューメラシーと質問の種類による中間回答傾向の違い

ただし設問の種類では、中間回答の選択率は大きく異なっていた。リッカート：通常条件が最も中間回答の割合が大きく、低群では5割弱(48.8%)、高群でも43.7%であり、続いてリッカート：スペース条件(高群で39.5%、低群で36.8%)、カテゴリー：正解あり設問とカテゴリー：態度設問が大よそ同じ程度のレベルとなっている(前者が低群32.7%、高群40.1%;カテゴリー態度設問 低群26.8%、高群38.5%)。この結果は2017年に行った調査(広田・佐藤, 2017)での主観的ニューメラシー尺度での差でも同様であり、全く内容が異なる2つのリッカート尺度(評価過敏性-誇大性自己愛尺度(中山・中谷, 2006); 二次元レジリエンス要因尺度(平野, 2010))でも同様の傾向で、前者では低群45.4、高群38.1、後者では48.3、41.6、カテゴリー：態度設問では18.4、20.2、カテゴリー：正解あり設問では23.7、20.2となっており、やはりリッカート尺度ではそれ以外のカテゴリー設問に比べ遥かに中間回答が高かった。

なお3つの設問の種類を通してニューメラシーによる差が最も大きいのはカテゴリー：態度設問で低群38.5に対し高群26.8、次がカテゴリー：正解あり設問(40.1 vs. 32.7)と、カテゴリー設問間で差が大きかった。

### (3) リッカート尺度での選択肢の順序や空間配置の操作の効果

次にリッカート尺度での選択肢の順序や視覚的配置を操作した5条件での違いを見た結果(図2)、通常条件以外は順序が通常条件と比較的近いスペース条件を除くと著しく中間回答「どちらでもない」の選択率が下がった(スペース条件 低39.5、高36.8; ランダム条件23.1, 23.0; リッカート中間右条件19.8, 14.0; リッカート中間左条件13.6, 8.8)。

なお、視覚上中央に位置している選択肢の選択率を算出したのが図2のスペース条件右側の参考の2つの数値である。これで見ると、中間左条件で中央に位置した選択肢の選択率は低群で39.1%、高群で33.9%、中間右条件では低群が34.3%、高群35.3%といずれもかなり高く、カテゴリー：

態度設問やカテゴリー：正解あり設問に近いレベルになっている。このことから見ると、回答者がリッカート尺度で中間回答を選ぶとき、「どちらでもない」という中庸な意味を持つ言語ラベルの影響よりも、視覚上の位置が反応の大きな手掛かりになっていると考えられる。これに対し、位置の手がかりを除いた中間左条件や中間右条件でのカテゴリー3に対する10%前後の選択率が「どちらでもない」という言語ラベルの影響と考えられる。中間という「視覚的位置」の手がかりによる回答率がカテゴリー設問と同じくらいのレベル（30～35%）であると仮定すると、リッカート尺度での中間回答選択率の高さ（44～49%）はこれらの加算結果と大よそ一致している。

#### (4) ニューメラシーの高さとリッカート尺度での中間回答

次に回答者のニューメラシーの高さが尺度の種類別での中間回答の選択率にどのように影響しているかを見る目的で5段階尺度の各選択肢の選択率の比較を行った。カイ二乗検定の結果では、いずれの尺度でもニューメラシーとカテゴリーの関係は有意であった。各図中の矢印と星は、それぞれ残差分析で有意な差が見られた部分で、星は有意水準を表している。

まず通常条件（図3.  $chisq=53.2, df=4, p<.001; Cohen's w=.11$ ）では

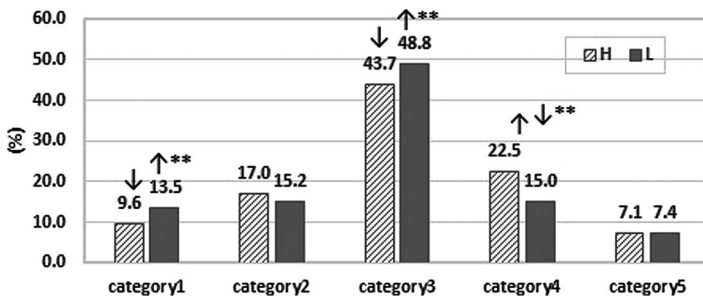


Fig. 3. Response frequencies by the levels of objective numeracy: Likert\_control condition

ウェブ調査における回答者のニューメラシーと質問の種類による中間回答傾向の違い

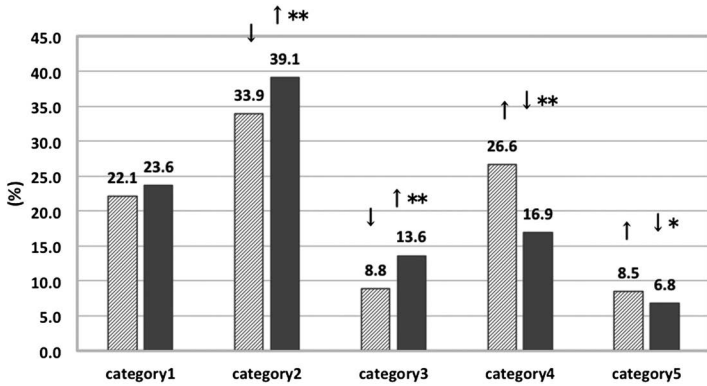


Fig. 4. Response frequencies by the levels of objective numeracy: Likert\_middle left (3-1-2-4-5)

すでに述べたように中間回答と極端回答，加えてカテゴリー 4 で差が見られた。

一方，中間選択が一番左，すなわちカテゴリーのトップに位置していた中間左条件（図 4， $chisq=89.8$ ， $df=4$ ， $p<.001$ ； $Cohen's w=.14$ ）ではカテゴリー 1 以外の全てで有意差が見られ，特に視覚的に中間に位置していたカテゴリー 2 と通常中間に位置する「どちらでもない」のカテゴリー 3 で低群が高く，逆にカテゴリー 4 では高群が低群より有意に多かった。なおカテゴリー 3 は通常の実験に比べるとかなり選択率は低い。

これと視覚的に逆の条件である中間右条件（図 5， $chisq=32.9$ ， $df=4$ ， $p<.001$ ； $Cohen's w=.08$ ）では全体で有意差があったが，通常条件と位置が変わらないカテゴリー 1, 2 では高群の実験は低群より有意に高くなり（いずれも  $p<.05$ ），中央に位置していた 4 では群間差はなく，逆に一番右に移動した「どちらともいえない」は通常条件よりも選択率はかなり低いものの，低群が高群を有意に上回っている（ $p<.01$ ）。このことから見ると，「どちらともいえない」の言語ラベルの付いたカテゴリー 3 は位置が移動することで選択率は下がるものの，位置が変わることの効果は高群で

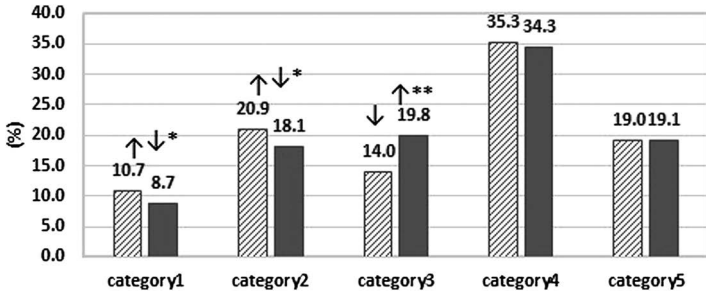


Fig. 5. Response frequencies by the levels of objective numeracy: Likert\_middle right (1-2-4-5-3)

大きく、低群ではむしろ「どちらでもいえない」を根強く選んでいる回答者が一定割合いることがわかる。

ただ、中間左条件では中央に位置していたカテゴリー 2 でやはり低群が高群より選択率が高かった反面、中間右条件で中央に位置していたカテゴリー 4 で高群・低群の間で差がなかったことは、カテゴリーの言語ラベルが影響している可能性が考えられる。カテゴリー 2 は「あまりあてはまらない」、カテゴリー 4 は「ややあてはまる」であり、また中間右選択 (図 5) ではカテゴリー 3 の減少が中間左選択 (図 4) よりも比較的小さいことからみると、中間右条件でのカテゴリー 3 「どちらともいえない」は一種積極的に選ばれているとみることができる。このことを併せて考えると、もしも位置を手掛かりに比較的認知負荷なく中央を選ぼうとしたとき、カテゴリー 2 「あまりあてはまらない」がある場合 (中間左条件) はそのまま選ばれることが多い一方、カテゴリー 4 「ややあてはまる」がある場合 (中間右条件) には、低群の回答者はそこを選ばず、「どちらでもない」を積極的に選んでいることになる。単に、視覚的手がかりにより認知負荷を下げて中間選択をしているならばこういったことは生じないだろう。このことから見ると、むしろ低群で行われるリッカート尺度での中間選択は必ずしも最小限化効果のみではなく、意味性を持った選択である可能性がある。

ウェブ調査における回答者のニューメラシーと質問の種類による中間回答傾向の違い

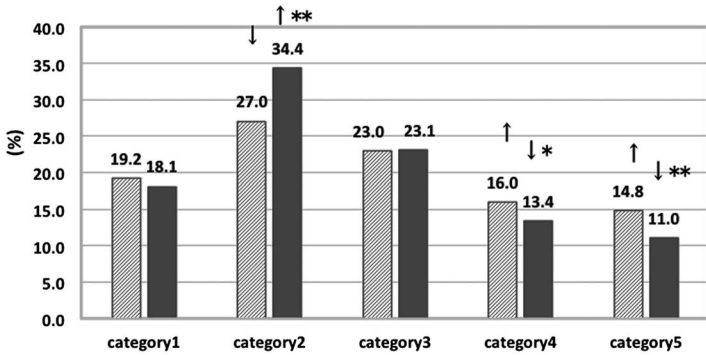


Fig. 6. Response frequencies by the levels of objective numeracy: Likert\_random condition

図6は5つのカテゴリーを位置をランダムに提示した条件 ( $chisq=38.1$ ,  $df=4$ ,  $p<.001$ ;  $Cohen's w=.09$ ) で、ここでは「どちらでもない」のラベルの付いたカテゴリー3では有意差はなく、カテゴリー2で低群が高く ( $p<.01$ )、カテゴリー4と5では逆に高群が相対的に高く (カテゴリー4  $p<.05$ ; カテゴリー5  $p<.01$ ) となった。このことも先述の結果と一貫しており、低群でカテゴリー2が、高群ではカテゴリー4が相対的に多く好まれていることになる。

なお単純にカテゴリーの幅を右に行くにつれて広くしていったスペース条件では (図7)、中央に位置するカテゴリー3ではわずかに低群が高く (ただし有意傾向)、また一番左に位置するカテゴリー1で低群が高群より有意に高く ( $p<.01$ )、カテゴリー2と4では逆に高群が有意に高くなっており (いずれも  $p<.01$ )、全体的な傾向は通常条件と同様である。

これらリッカート尺度での条件別の結果をまとめると、ランダム条件を除くと「どちらでもない」の言語ラベルの付いた中間選択肢はスペース条件で有意傾向と弱いとはいえ、いずれでもニューメラシー低群で高群よりも選択率が高いことが明らかになった。このことから見ると、視覚的位置の手がかりが中間選択にとって重要ではあるものの、「どちらでもない」

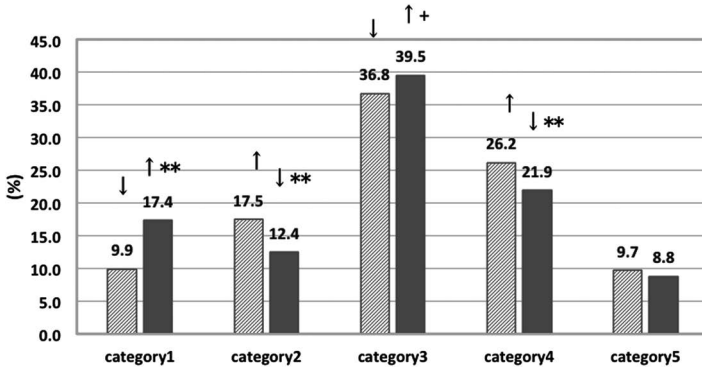


Fig. 7. Response frequencies by the levels of objective numeracy: Likert\_spacing

という中間的な言語ラベルによる意味性は低ニューメラシーの回答者が中間選択する傾向の一因になっているといえる。

一方、中間左条件と中間右条件で高ニューメラシーの回答者が低群よりも有意に多く選択した選択肢はいずれも3が実験条件として視覚的に位置したものの逆側の極の2つの選択肢（例えば中間左条件では一番右、逆は一番左）である。これから見ると、3の位置が一種のアンカーになっている可能性が考えられる。

### (3) 態度質問及び正解のある知識問題でのニューメラシーによる差

次にカテゴリカルな態度に関する質問と、同様にカテゴリカルだが正解のある知識問題でのニューメラシーの高低による違いを見たのが図8、図9である。

図8の態度に関する質問では低群で中間選択の選択傾向が強いことは明白であり ( $chisq=203.7, df=4, p<.001; Cohen's w=.14$ )、他の設問に比べてカイ二乗値も大きくなっている。中間選択では1%水準で低群の方が高く、カテゴリ5でも若干高い ( $p<.05$ ) が、あとはいずれも高群の方が高い (カテゴリ1  $p<.01$ ; カテゴリ2  $p<.01$ ; カテゴリ4 有意傾向)。



ウェブ調査における回答者のニューメラシーと質問の種類による中間回答傾向の違い

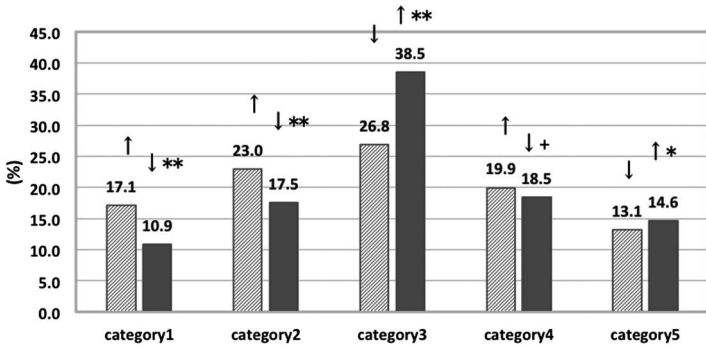


Fig. 8. Response frequencies by the levels of objective numeracy: Attitudes

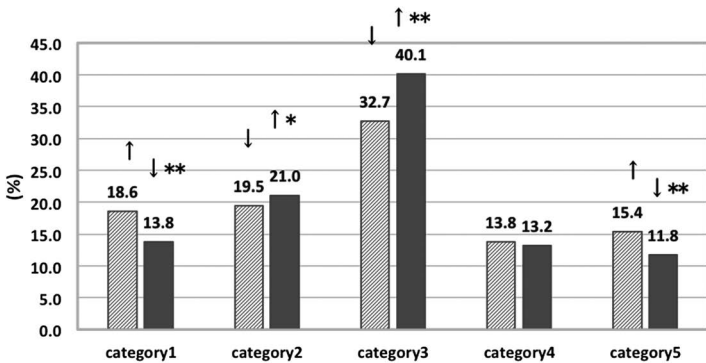


Fig. 9. Response frequencies by the levels of objective numeracy: Questions

このような中間選択が低群で高い傾向は正解のある設問でも見られ ( $\chi^2=112.6, df=4, p<.001; Cohen's w=.10$ ), 中間選択では1%水準の有意差が見られた (図9)。

#### (5) ニューメラシー別 質問の種類間の相関

最後に質問の種類間での中間選択の数の相関を、ニューメラシー別にまとめたものが表1である。第1に明らかなことは、態度の質問と正解のあ

る問題での中間選択数の相関が高く、特にニューメラシー低群でそれが相対的に高く、高群で.29~.54に対し低群では.48~.66となっていることで、例外的にニューメラシーの高低群で相関の高さが逆転している中間右条件でも.06程度の違いに過ぎない。このことから見ると、ニューメラシー低群では少なくとも質問の性格がかなり異なる態度設問と正解のある設問でも中間選択傾向はかなり共通しているといえることができる。

一方、リッカート尺度についてはやや結果は複雑である。通常の並びで

Table 1. Correlation between the number of mid-response for each kinds of questions by numeracy levels

| 客観<br>numeracy level |                    | 3 (どちらでもない) への反応 |                | ※中央位置の<br>選択肢への反応 |                | 言語ラベル  |               |
|----------------------|--------------------|------------------|----------------|-------------------|----------------|--------|---------------|
|                      |                    | カテゴリー—<br>態度     | カテゴリー—<br>正解あり | カテゴリー—<br>態度      | カテゴリー—<br>正解あり |        |               |
| High control         | リッカート              | 0.280            | 0.268          |                   |                |        |               |
|                      | カテゴリー__態度          |                  | 0.332          |                   |                |        |               |
|                      | 中間左<br>(3-1-2-4-5) | リッカート            | -0.072         | 0.019             | -0.011         | -0.085 | 2「あまりあてはまらない」 |
|                      | カテゴリー__態度          |                  | 0.290          |                   |                |        |               |
|                      | 中間右<br>(1-2-4-5-3) | リッカート            | -0.230         | -0.197            | 0.271          | 0.244  | 4「ややあてはまる」    |
|                      | カテゴリー__態度          |                  | 0.536          |                   |                |        |               |
|                      | ランダム               | リッカート            | -0.086         | -0.091            |                |        |               |
|                      | カテゴリー__態度          |                  | 0.476          |                   |                |        |               |
| スペース                 | リッカート              | 0.027            | 0.066          |                   |                |        |               |
|                      | カテゴリー__態度          |                  | 0.290          |                   |                |        |               |
|                      | Low control        | リッカート            | 0.312          | 0.306             |                |        |               |
|                      |                    | カテゴリー__態度        |                | 0.624             |                |        |               |
|                      | 中間左<br>(3-1-2-4-5) | リッカート            | -0.125         | -0.310            | 0.041          | 0.036  | 2「あまりあてはまらない」 |
|                      | カテゴリー__態度          |                  | 0.610          |                   |                |        |               |
|                      | 中間右<br>(1-2-4-5-3) | リッカート            | -0.206         | -0.227            | -0.045         | -0.070 | 4「ややあてはまる」    |
|                      | カテゴリー__態度          |                  | 0.475          |                   |                |        |               |
| ランダム                 | リッカート              | -0.028           | 0.043          |                   |                |        |               |
| カテゴリー__態度            |                    | 0.659            |                |                   |                |        |               |
| スペース                 | リッカート              | 0.123            | 0.126          |                   |                |        |               |
|                      | カテゴリー__態度          |                  | 0.592          |                   |                |        |               |

ある通常条件の場合、中間選択の数は、高群で態度設問との相関がそれぞれ .28 と .31、正解あり設問との相関は .27 と .31 であり、態度設問と正解のある設問ほどの相関ではないが多少の相関は見られた。ただし、呈示順序を様々に変えた場合、低群でも高群でも相関はほとんどなくなり、ところによっては逆相関がみられた。

#### 4. 考察

まず全体として注目されるのは、質問の性質にかかわらずいずれの場合も中間選択率が高かったという点である。本研究の3つの質問群に対して用いられる認知能力は性質が異なっていると推測され、知識問題の場合は問題自体を解く知的能力が必要になる一方、同じカテゴリカルな設問でも態度設問ではまず設問に関連する事象や態度の記憶を再生する、あるいはそれまで考えなかった問題については知識や評価を集め、特定のカテゴリに符号化していることが想定される。これらに対し、リッカート尺度もまた態度設問で必要な能力に加え、最終的には心的強度を1~5の順序カテゴリに符号化することが求められる。だがそれらを通じて中間選択率が高かったことから見ると、5選択肢の設問の場合、正解があるような問題であっても、記憶や態度を振り返る場合でも、対象に関する心的強度を表出する場合でも、性質にかかわらず中間選択肢を選ぶ人が一定割合存在していることを示唆している。特に知識問題では日常の試験におけるストラテジー、すなわち「正解はランダムに位置しているはず」が影響している可能性があり、その場合ニューメラシーの高低に関わらず中間選択傾向は弱いはずだが、本研究の結果(図9)で見るとやはり中間選択傾向は明白である。リッカート尺度での視覚的配置や順序への操作の影響を見ると、通常条件ではほかの設問に比べても最も中間選択傾向が強かったが左や右、ランダムにすると中間選択率は大幅に減少し、逆に意味的に中間ではないが位置が中央である選択は通常条件ほどではないが非常に高くなっ

ていた。このことから見ると、「視覚的位置が中央」であることが中間選択を引き起こす強い要因になっていることがわかる。この結果は Tourangeau, Couper, and Conrad (2004) のいう『中央が典型的』ヒューリスティックの主張と一致するものである。

一方ニューメラシーによる差に着目すると、中間選択傾向はニューメラシー低群でリッカート尺度のランダム設問の条件以外ほどの設問の種類でも高群よりも有意に高かった。このことはニューメラシーで測られる数的能力が中間選択傾向と関連のあることを示唆している。本研究で用いている Weller et al. のニューメラシー尺度 (2013) は教育水準や IQ とも一定の相関があることが報告されていることから、記憶や問題を解く知的能力が低群の方が劣っている可能性もある。ただし、態度に関する質問と正解のある問題での中間選択率とはそれほど大きな違いがないこと、一方でリッカート尺度の通常条件での中間選択率は高いとはいえそれ以外の視覚的位置を変化させた条件での選択率は逆に大幅に低下した。この結果から見ると、質問の種類による違いはそれほど小さくなく、むしろ視覚の手がかりが大きいというべきである。また設問間の中間選択傾向の相関を見ると、ニューメラシー低群の方が高群よりも相関が高くなっており、ニューメラシー低群は視覚の手がかりをより強く使っている可能性がある。

ただし、1つ注目すべきなのは中間左条件、中間右条件とも視覚的中央位置の選択肢の選択率が通常条件よりは下がっている点である。このことから見ると、回答者は単純に5択の中央を選んでいるわけではなく、10%前後はリッカート尺度の5択の言語ラベル「どちらでもない」の意味の中庸性が位置の手がかりによる中間選択に追加され、その傾向は低ニューメラシー群で強かった。その点では、認知負荷を抑えるという要素だけではなく、意味性の要素も部分的に影響していることがわかる。これは中間左条件と中間右条件での結果の非対称性からも裏付けられる。中間左条件、すなわち (3-1-2-4-5) と、中間右条件、すなわち (1-2-4-5-3) ではいず

ウェブ調査における回答者のニューメラシーと質問の種類による中間回答傾向の違い

れも3「どちらでもない」の選択率は相対的に下がっているがその低下量は後者の方が少なく、逆に中央の位置にある選択肢の選択率は中央左条件では低群の方が高いが、中央右条件では2群の中央選択肢の選択率の差はなくなっている。つまり、中間左条件では中央位置にきた2「あまりあてはまらない」についての選択率が2群で差があるが、中間右条件での中央位置にある4「ややあてはまる」では差がなくなっている。つまり、後者のように、肯定的な意味を持った選択肢が中央に位置する場合は低群も高群でも位置の手がかりに基づいてそれを見ても積極的に「どちらでもない」を探索して回答する行動は起きにくい、否定的な意味を持った選択肢が中央の場合、低群では積極的にこれを回避して「どちらでもない」を選ぶ傾向が生じたため、両群の間で中央位置の選択肢に対しての選択率に差が生じなかったと推測される。このことから見ると、意味に基づいた一種積極的な中間選択傾向は、特にニューメラシー低群で強いことが推測される。

本研究の結果からは客観ニューメラシーが質問の性質に限らず、中間選択傾向と関連していることが明らかになった。ただし特に低ニューメラシー群ではリッカート尺度において一部で言語ラベルによる影響も生じている可能性が明らかになった。客観ニューメラシーは教育水準やIQとも関連が高いことから、これらの間の関係については今後検討を進める必要がある。

※本研究は科研費補助金「多肢選択肢における回答行動の統合的研究：質問紙・ウェブ調査法の設計と妥当性の検討」(A)16H02050、代表者：坂上貴之、平成28～30年度)の助成に基づいて行われた。

※謝辞 本研究は慶應義塾大学文学部坂上貴之教授を代表とする科研グループでの成果に基づくものである。坂上教授には学部生時代から長年にわたって指導を受け、後には意思決定を中心とした研究会「行動的意思決定研究会(略称 U研)」を結成して共同研究を行い、「心理学が描くリスクの世界」他複数の共編著を執筆させていただいた。長年にわたるご指導と共同研究に参加させていただいたことに深く

感謝いたします。

## 5. 引用文献

- Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (2007). Measuring numeracy without a math test: Development of the Subjective Numeracy Scale (SNS). *Medical Decision Making*, **27**(5), 672-680.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, **19**(4), 25-42.
- 萩原潤治・村田ひろ子・吉藤昌代・広川裕 (2018). 住民基本台帳からの無作為抽出による WEB 世論調査の検証②: 郵送調査との回答分布の比較. *放送研究と調査*, **68**(9), 48-79.
- 林知己夫 (1996). *日本らしさの構造: こころと文化をはかる*. 東洋経済新報社.
- 平野真理 (2010). レジリエンスの資質的要因・獲得的要因の分類の試み: 二次元レジリエンス要因尺度 (BRS) の作成. *パーソナリティ研究*, **19**(2), 94-106.
- 広田すみれ (2014). ニューメラシー, 個人属性と回答傾向の関係. 慶應義塾大学「思考と行動判断」研究拠点「社会的行動判断・意思決定のダイナミズムに関する人文科学, 行動科学, 神経科学の分野横断的統合研究」プロジェクトシンポジウム『質問紙の科学: その可能性と展望』.
- 広田すみれ・佐藤章子 (2017). Numeracy が Web 調査の異なる質問形式での回答傾向に与える影響. 日本行動計量学会第 45 回大会特別セッション話題提供.
- Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, **5**(3), 213-236.
- Krosnick, J. A. & Alwin, D. F. (1987). An evaluation of a cognitive theory of response-order effects in survey measurement. *Public Opinion Quarterly*, **51**(2), 201-219.
- Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making*, **21**(1), 37-44.
- 増田真也・坂上貴之 (2014). 調査の回答における中間選択: 一原因, 影響と其对策一. *心理学評論*, **57**(4), 472-494.
- 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

ウェブ調査における回答者のニューメラシーと質問の種類による中間回答傾向の違い

- 択の増加. *行動計量学*, **44(2)**, 117-128.
- Messick, S., & Frederiksen, N. (1958). Ability, acquiescence, and "authoritarianism". *Psychological Reports*, **4(3)**, 687-697.
- 三浦麻子・小林哲郎 (2016a). オンライン調査における努力の最小限化 (Satisfice) 傾向の比較: IMC 違反率を指標として. *メディア・情報・コミュニケーション研究*, **1**, 27-42.
- 三浦麻子・小林哲郎 (2016b). オンライン調査における努力の最小限化 (Satisfice) を検出する技法: 大学生サンプルを用いた検討. *社会心理学研究*, **32(2)**, 123-132.
- 宮戸美樹・上野行良 (1996). ユーモアの支援的効果の検討—支援的ユーモア志向尺度の構成—. *心理学研究*, **67(4)**, 270-277.
- 内閣府大臣官房政府広報室 (2016). 内閣府国民選好度調査. <http://warp.da.ndl.go.jp/infondlj/pid/10361265/www5.cao.go.jp/seikatsu/senkoudo/senkoudo.html>
- 中山留美子・中谷素之 (2006). 青年期における自己愛の構造と発達的变化の検討. *教育心理学研究*, **54(2)**, 188-198.
- Narayan, S., & Krosnick, J. A. (1996). Education moderates some response effects in attitude measurement. *Public Opinion Quarterly*, **60(1)**, 58-88.
- 大阪商業大学 JGSS 研究センター (2016). 日本版総合的社会調査 (JGSS: Japanese General Social Surveys). [http://jgss.daishodai.ac.jp/surveys/sur\\_top.html](http://jgss.daishodai.ac.jp/surveys/sur_top.html).
- Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology*, **45(4)**, 867-872.
- Peters, E., Västfjäll, D., Slovic, P., Mertz, C., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science*, **17(5)**, 407-413.
- Rammstedt, B., Goldberg, L. R., & Borg, I. (2010). The measurement equivalence of Big-Five factor markers for persons with different levels of education. *Journal of Research in Personality*, **44(1)**, 53-61.
- Schuman, H., Kalton, G., & Ludwig, J. (1983). Context and contiguity in survey questionnaires. *Public Opinion Quarterly*, **47(1)**, 112-115.
- Schuman, H., & Presser, S. (1981). *Questions and answers in attitude surveys: Experiments on question form, wording, and context*. CA: Academic Press.
- SPI ノートの会 (編著) (2016a). 必勝・就職試験! WEB テスティングサービス・CUBIC・TAP・TAL・ESP・CASEC 対策用 8 割が落とされる「Web テスト」完全突破法 3 2018 年度版. 洋泉社.

- SPI ノートの会 (編著) (2016b). 必勝・就職試験! Web-CAB・GAB Compact・IMAGES 対応 CAB・GAB 完全突破法 2018 年度版. 洋泉社.
- Tourangeau, R., Conrad, F. G., & Couper, M. P. (2013). *The science of web surveys*. Oxford university press.
- Tourangeau, R., Couper, M. P., & Conrad, F. G. (2004). Spacing, position, and order: Interactive heuristics for visual features of survey questions. *Public Opinion Quarterly*, **68**(3), 368-393.
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). Chapter 5. Factual judgments and numerical estimates. In Tourangeau, R., Rips, L. J., & Rasinski, K. *Psychology of survey response*. 136-164. Cambridge university press.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, **5**(2), 207-232.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and Biases. *Science*, **185**(4157), 1124-1131.
- 上野行良 (1993). ユーモアに対する態度と攻撃性及び愛他性との関係. *心理学研究*, **64**(4), 247-254.
- Wason, P. C. (1966). Reasoning. In B. Foss (ed.). *New horizons in psychology*. London: Penguin.
- Weller, J. A., Dieckmann, N. F., Tusler, M., Mertz, C. K., Burns, W. J., & Peters, E. (2013). Development and testing of an abbreviated numeracy scale: A Rasch analysis approach. *Journal of Behavioral Decision Making*, **26**(2), 198-212.