

## Likert法における回答選択枝のレイアウトが 選択枝間の心理的距離に与える影響

脇田 貴文

### Effects of the layout of response options in Likert scale on psychological distance between options.

Takafumi WAKITA

#### abstract

In applying Likert scale to psychological measurement, various layouts are used in rating scales such as number line or numerical number formats. The present study examined whether the layout of options in the Likert Scale influences the responses. In this study, we focused on the psychological distance between options and considered what effect the difference of rating scale format has. For this aim, we used 4 formats of rating scale, and provided 940 undergraduate students with 4 types of questionnaires that were composed of the same items but with different layouts of options to assess these items, asking them to fill them in. We used the method suggested by Wakita et al. in Educational and Psychological Measurement (2012) for an evaluation of psychological distance. It was concluded that layouts did not make a difference in responses.

Keyword: Likert scale, option layout, psychological distance, Item Response Theory

#### 抄 録

心理測定場面で用いられる Likert 法では、選択枝に数直線や数値などさまざまな形式のレイアウトが用いられている。本研究では、選択枝のレイアウトが、回答者の回答行動にどのような影響を与えるかを、選択枝間の心理的距離の観点から検討することを目的とした。本研究では、同一尺度、同一内容でレイアウトのみが異なる 4 つのバージョンを 940 名の大学生にランダムに配布した。Wakita *et al.* (2012) の方法を用いて、選択枝間の心理的距離を検討した結果、レイアウトによって違いは見られなかった。

キーワード：リッカート法、選択枝のレイアウト、心理的距離、項目反応理論

問題と目的

心理学におけるデータ収集方法として最もよく用いられている方法が質問紙法である。そして、質問紙法の中でも、ある構成概念を測定する心理測定尺度に対して Likert 法を用いることが多い。Likert 法は、Likert (1932) が態度測定に用いたことがきっかけとなり一般的な方法となり現在に至っている。これは回答選択枝に等間隔の整数値を与え、その和（もしくは平均）をもって回答者の潜在特性値とする方法である。Likert (1932) では、Figure 1 に示した形式で回答を求めており、厳密な意味ではこのような形式のものを Likert 法と呼ぶべきだろう。しかし、現在はさまざまな回答段階数、選択枝の表現（anchor; 評定尺度表現）、回答形式が用いられており、どのような形式で回答を求めるかは調査者の考えに依存しているといつて良い。

9. Our army and navy should be used when necessary to defend American property rights in Mexico.

Strongly Strongly	Approve	Undecided	Disapprove	Approve Disapprove
(5)	(4)	(3)	(2)	(1)

Figure 1 Likert (1932) における形式 (例: Imperialism Scale)

Likert 法に関する記述は、さまざまな調査法の書籍に認められるが、その中でも統一した形式が示されているわけではない。例えば、Brace (2008) では、選択枝に数字を明示せず、チェックを行う形式 (Figure 2) および数字を用いる形式 (Figure 3) が示されている。また、鎌原・宮下・大野木・中澤 (1998) では、数直線に評定尺度表現を提示した形式 (Figure 4) や Brace (2008) とほぼ同様の形式 (Figure 5) が示されている。Thomas, Ostrom & Gannon (1996) では Figure 3 や 5 とほぼ同様の形式 (Figure 6) および、す

	Disagree		Neither	Agree	
	strongly	Disagree	Agree nor Disagree	Agree	strongly
Being a smart shopper is worth the extra time it takes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Figure 2 Brace (2008) における選択枝例 1

1=Excellent	2= Very Good	3=Good	4=Fair	5=Poor	
Cleanliness of your guest room upon entering	1	2	3	4	5

Figure 3 Brace (2008) における選択枝例 2

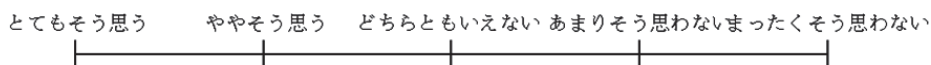


Figure 4 鎌原他 (1998) における選択枝例 1

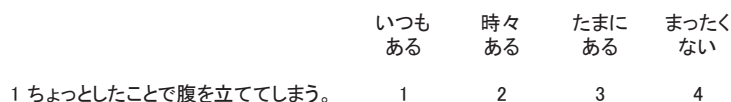


Figure 5 鎌原他 (1998) における選択枝例 2



Figure 6 Gannon (1996) における選択枝例 1



Figure 7 Gannon (1996) における選択枝例 2

すべての選択枝に評定尺度表現を示していないもの (Figure 7) が例示されている。

このように多種多様な回答形式が用いられているが、鈴木 (2011) による質問紙のデザインに関して著した書籍であっても、Likert 法の選択枝レイアウトに関してはほとんど触れられていないものもある。ここで例示した文献は、限られたものではあるが、この多様性は、回答形式がそれほど回答に影響しないと考えられているためである可能性が高い。

基本的に上記のような論文・書籍における記述は、研究者がこのように用いているという例示であり、必ずしも検証されたものとはいえない。Likert 法のレイアウトに関する研究はそれほど多くはないが、例えば Dillman & Christian (2005) では、選択枝を線形 (上から下に順) に配置した場合と、そうではない場合 (2 行にわたって配置) で比較を行っ

た。その結果、回答者の選択行動に違いが認められたことを報告している。また、チェックボックス間の距離が等しい場合とそうではない場合の比較では、順序尺度水準にある選択枝の場合には有意な違いが認められなかったと報告している。Hartley & Betts (2010) では、選択枝の提示順と数字が、回答者にどのような影響を与えるか検討し、その影響を示唆している。しかし、これらの研究は、必ずしも心理尺度で行われておらず、一般化できる結果とは言えない。

Likert 法における選択枝のレイアウトの違いが測定においてどのような影響があるかについて、測定の誤差の観点から次のように考えることができる。測定の誤差は系統誤差と偶然誤差に分けて考えられる。通常、測定の誤差を評価する際に用いる信頼性係数は、偶然誤差の部分の評価していることになる。しかし、複数項目による心理尺度を用いた測定においては、同一の選択枝を用いているため、選択肢によって生じる誤差は系統誤差であると考えることができる。この誤差にもさまざまなものが考えられるが、その1つとして選択肢間の心理的距離が挙げられるだろう。本研究では、Likert 法の重要な仮定の1つである選択枝間の心理的距離の等間隔性の観点から選択枝のレイアウトについて検討することを目的とする。

Wakita, Ueshima, & Noguchi (2012) では、Likert 法における選択枝間の心理的距離に関して Item Response Theory (IRT) の IRT モデルの1つである Generalized Partial Credit Model (Muraki, 1992) を用いて検討する方法を提案した。そして、同一尺度に対して4件法、5件法、7件法の場合で、選択枝間の心理的距離にどのような違いが認められるかを検討している。そこで本研究では Wakita *et al.* (2012) の方法を用いて、選択枝間の心理的理の観点から4つの選択枝レイアウトを比較する。

## 方法

**質問紙の構成** Big Five 尺度 (和田, 1996) の下位尺度である Extraversion (外向性) と Neuroticism (情緒不安定性) の各12項目を用いた。なお、外向性では12項目中6項目の逆転項目が含まれ、情緒不安定性では1項目の逆転項目が含まれる。脇田 (2004) では、項目内容がポジティブかネガティブかにより評定尺度の間隔が異なることが示されており、本研究においても両者を区別する必要がある。そこで、外向性の社会的にポジティブな内容の6項目 (以下、BF-EN と表記する)、ネガティブな内容の6項目 (以下、BF-ER と表記する)、そして逆転項目1項目を除いた情緒不安定性の11項目 (以下 BF-N と表記する) を分析対象とした。

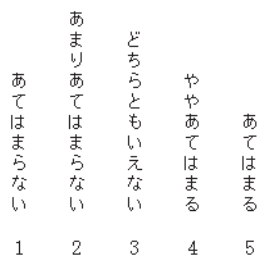


Figure 8 選択枝 Ver.B

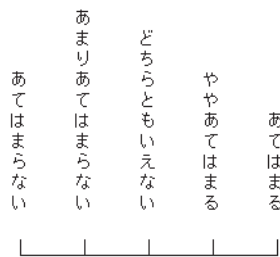


Figure 9 選択枝 Ver.C

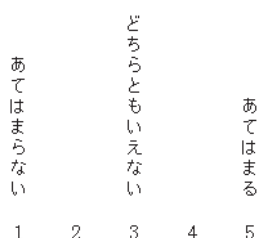


Figure 10 選択枝 Ver.D

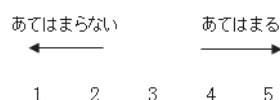


Figure 11 選択枝 Ver.E

本研究で用いた4つの選択枝レイアウトのバージョンはすべて5件法であった。Ver.Bは比較の際に基準となるもので、一般的な評定尺度表現を用いた（Figure 8）<sup>1)</sup>。Ver.Cは、選択枝には数字を明示せず、数直線で示したものである（Figure 9）。Ver.Bのように数字を示すことで、回答者に選択枝間の等間隔性をイメージさせることができるといわれているが、良い回答や悪い回答（得点が高いもしくは低い）をイメージさせるため望ましくないという意見もある。従って、Ver.Cは等間隔性をイメージさせつつも、数字の影響をなくしたものとして設定した。Ver.Dは、数字を明示しているが、両端に5。“あてはまる”と1。“あてはまらない”，そして中央に3。“どちらともいえない”という評定尺度表現を明示し、2と4には評定尺度表現を示さないパターンを用意した（Figure 10）。これはVer.Bにあるように“やや”や“あまり”という言葉によって等間隔性に影響があると考えたためである。Ver.Eは、選択枝の意味として“あてはまる”という方向，“あてはまらない”という方向のみを示し、評定尺度表現によって選択枝を定義しないものである（Figure 11）。この場合、回答者は自身であてはまる～あてはまらないという回答次元を分

1) Ver.Bの結果は、Wakita, Ueshima, & Noguchi (2012) において報告しているが、比較基準とするために再度記述した。

割して回答をすることが求められる。そして数字の情報から、等間隔に分割するというイメージを作りやすいのではないかと考えられる。

**実施時期** 2002年9月～11月に実施した。

**実施方法** 大学生2071名を対象に、本研究で検討する3つのバージョンを含む9つのバージョンをランダムに配布した。本研究で扱う各バージョンに対する回答者はVer.Bが254名、Ver.Cが218名、Ver.Dが252名、Ver.Eが216名であった。

**分析方法** はじめに尺度得点の平均値および標準偏差を求め、尺度得点の平均値にどのような影響が認められるかを検討した。次に一次元性の確認を行った。さらにPARSCALE (Muraki & Bock, 2002) を用いて、IRTのGeneralized Partial Credit Model (Muraki, 1992) により分析を行い、各項目パラメタ値を算出した。その後、Wakita *et al.* (2012) の方法を用いて、各選択枝の尺度値を求めた。

## 結果と考察

### 尺度得点の平均および標準偏差

バージョンごとに平均および標準偏差を求めた (Table 1)。その結果、最も大きな差が生じているBF-NではVer.BとVer.Dで0.150、BF-ENではVer.CとVer.Eで0.128、BF-ERではVer.CとVer.Dで0.078であった。この差を大きな違いとみるか、大差ないとみるかは判断の分かれるところであろう。しかし、特定のバージョンにおいて常に高くなるというような規則性も認められないため、誤差の範囲であると考えられる。また、標準偏差に関しては、いずれの尺度においても0.1程度の違いが生じている。これに関しても平均値と同様に積極的な解釈をすることはできないと考えられる。なお、いずれの尺度においてもVer.DとVer.Eの間でVer.Dの標準偏差が0.1ほど大きくなっているが、BF-Nでは逆の結果となっている。このように尺度得点の平均値および標準偏差の観点からは選

Table 1 各バージョン、尺度ごとの平均および標準偏差

	BF-N			BF-EN			BF-ER		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Ver.B	252	3.517	(0.799)	254	3.317	(0.803)	254	2.552	(0.791)
Ver.C	218	3.374	(0.698)	218	3.433	(0.797)	216	2.483	(0.730)
Ver.D	214	3.367	(0.790)	214	3.315	(0.839)	214	2.561	(0.799)
Ver.E	251	3.390	(0.743)	252	3.304	(0.739)	252	2.538	(0.691)

択枝レイアウトの影響は少ないと判断して良いと考えられる。

### 各選択枝の尺度値の算出

IRTモデルを尺度に適用するための前提が尺度の一次元性である。そこで、因子分析（主因子法）を行い、その固有値の減衰状況から尺度の一次元性を検討した。各尺度、各バージョンの第1固有値と第2固有値およびその比をTable 2に示した。いずれにおいても一次元性が満たされていると判断した。

Table 2 各バージョン、尺度の第1、2固有値とその比

		BF-N	BF-EN	BF-ER
Ver.B	$\lambda_1$	5.362	3.562	3.013
	$\lambda_2$	1.008	0.779	0.870
	$\lambda_1/\lambda_2$	5.319	4.576	3.464
Ver.C	$\lambda_1$	4.715	3.727	2.835
	$\lambda_2$	1.204	0.751	0.850
	$\lambda_1/\lambda_2$	3.916	4.966	3.337
Ver.D	$\lambda_1$	5.375	3.591	2.749
	$\lambda_2$	1.108	0.800	0.984
	$\lambda_1/\lambda_2$	4.849	4.490	2.794
Ver.E	$\lambda_1$	5.266	3.817	3.087
	$\lambda_2$	1.161	0.709	0.875
	$\lambda_1/\lambda_2$	4.537	5.381	3.528

続いて、各項目パラメタを算出した。そして、カテゴリ・パラメタから選択枝の尺度値（ $\mu$ ）を求め、さらにその尺度値が1から5になるように変換を行った（Table 3）。Table 3に示した変換後の尺度値を数直線に表した（Figure 12, 13, 14）。Figure 12から、BF-Nにおいては、Ver.Cの第3、4選択枝の位置が左に寄っている。また第4選択枝はいずれのバージョンにおいても左に寄っている。これはBF-Nは、社会的にネガティブな項目内容であり、これらの項目に肯定的な回答をすることの影響であると考えられる。

項目内容がポジティブなBF-EN（Figure 13）においてはVer.Eが全体的に左に寄っているが、いずれのバージョンも、選択枝間の心理的距離は等しくなっていると判断して良いだろう。特に、第2、4選択枝に表現を与えていないVer.Dにおいては、ほぼ整数値に近い位置にあり等間隔性が満たされている。BF-Nと同様にネガティブな内容の項目で構

Table 3 各バージョン、尺度のカテゴリ・パラメタ、尺度値 ( $\mu$ )、変換後尺度値

	Ver.B			Ver.C			Ver.D			Ver.E		
	カテゴリ パラメタ	尺度値 ( $\mu$ )	変換後 尺度値	カテゴリ パラメタ	尺度値 ( $\mu$ )	変換後 尺度値	カテゴリ パラメタ	尺度値 ( $\mu$ )	変換後 尺度値	カテゴリ パラメタ	尺度値 ( $\mu$ )	変換後 尺度値
BF-N	-1.176	-1.668	1.000	-1.675	-2.089	1.000	-1.359	-1.819	1.000	-1.686	-2.098	1.000
	-0.354	-0.723	2.078	-0.510	-0.977	1.999	-0.431	-0.833	2.059	-0.323	-0.862	2.159
	0.150	-0.100	2.790	0.197	-0.150	2.742	0.329	-0.049	2.902	0.242	-0.039	2.930
	1.379	0.674	3.673	1.988	0.843	3.634	1.461	0.805	3.818	1.767	0.830	3.746
	1.836	5.000		2.363	5.000		1.905	5.000		2.168	5.000	
BF-EN	-1.641	-2.059	1.000	-1.615	-2.037	1.000	-1.575	-2.003	1.000	-1.794	-2.192	1.000
	-0.589	-1.018	1.986	-0.684	-1.070	1.922	-0.641	-1.031	1.960	-0.805	-1.200	1.868
	0.470	-0.054	2.899	0.541	-0.063	2.881	0.591	-0.022	2.957	0.590	-0.091	2.837
	1.761	0.973	3.873	1.758	1.018	3.912	1.625	1.015	3.981	2.009	1.105	3.884
	2.163	5.000		2.160	5.000		2.046	5.000		2.381	5.000	
BF-ER	-1.419	-1.870	1.000	-1.705	-2.115	1.000	-1.397	-1.852	1.000	-1.926	-2.308	1.000
	-0.698	-1.014	1.892	-0.676	-1.092	1.966	-0.696	-1.005	1.911	-0.794	-1.226	1.910
	0.582	-0.051	2.896	0.668	-0.004	2.993	0.679	-0.008	2.984	0.639	-0.065	2.887
	1.535	0.982	3.972	1.714	1.088	4.024	1.415	1.001	4.069	2.081	1.152	3.911
	1.968	5.000		2.122	5.000		1.867	5.000		2.446	5.000	

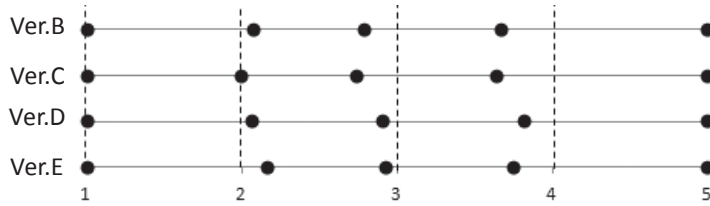


Figure 12 BF-Nの選択肢間の心理的距離

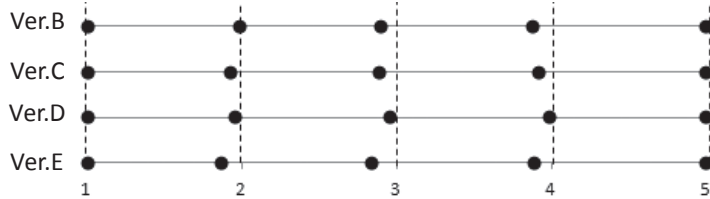


Figure 13 BF-ENの選択肢間の心理的距離

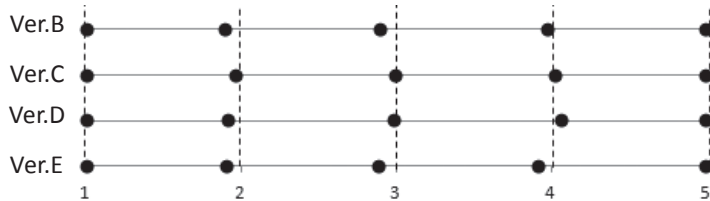


Figure 14 BF-ERの選択肢間の心理的距離



成される BF-ER でも、いずれのバージョンにおいても等間隔性が満たされていると考えて良いだろう。

これらの結果から、BF-N においてはバージョンによるばらつきが認められたが、BF-EN および BF-ER ではいずれのバージョンにおいても選択枝間の心理的距離は等しいという結果が得られた。BF-N も BF-ER も項目内容がネガティブであるという点はほぼ同様であるにもかかわらず結果の傾向に違いが認められた点について、他の尺度の分析を通してさらに検討する必要がある。いずれの尺度においても安定していたレイアウトは Ver.D であった。実際の質問紙調査においてこの形式が用いられることは少ないと思われるが、選択枝間の等間隔性という観点からは推奨できるものであると考えられる。ただし、この形式は Likert 法による回答に慣れていない回答者にとっては、回答の負担が大きい可能性がある。

また、Ver.B と Ver.C は数字と数直線の比較であったが、大きな違いは認められなかった。先述したとおり、数字を明示することで回答者の回答に影響を与えるという指摘がなされることがあるが、本研究の結果からは特に違いはなく問題はないと考えられる。これらの結果を通して、選択枝間の等間隔性という観点からは選択枝のレイアウトに関して大きな影響は認められなかった。Nicholls, Okubo & Andrea (2006) が検討しているような、“あてはまる” に 5, “あてはまらない” に 1 を与えた場合と、“あてはまる” に 1, “あてはまらない” に 5 を与えた場合の違いや、Hartley & Betts (2010) が検討しているような選択枝を配置する位置 (水平に配置するか、垂直に配置するか) 等についても同様に検討していくことが必要であろう。

#### 引用文献

- Brace, I. (2008). Questionnaire design: How to plan, structure and write survey material for effective market research (2<sup>nd</sup> edition) Kogan Page, London.
- Hartley, J., & Betts, L. R. (2010). Four layouts and a finding: the effects of changes in the order of the verbal labels and numerical values on Likert-type scales. *International Journal of Social Research Methodology*, *13*, 17-27.
- 鎌原雅彦・宮下一博・大野木裕明・中澤潤 (1998). 心理学マニュアル 質問紙法 北大路書房
- Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, *16*, 159-176.
- Muraki, E. & Bock, R. D. (1996). PARSCALE: parameter scaling of rating data [computer program]. Chicago, IL: Scientific Software, Inc..
- Nicholls, M. E. R., Orr, C. a, Okubo, M., & Loftus, A. (2006). Satisfaction guaranteed: the effect of

- spatial biases on responses to likert scales. *Psychological science*, **17**, 1027-1028.
- Ostrom, T. M. and Gannon, K. M. (1996) Exemplar Generation: Assessing How Respondents Give Meaning to Rating Scales, Schwarz, N. Sudman, S. edit. *Answering Questions: Methodology for determining cognitive and communicative processes in survey research*. Jossey-Vass Inc. California.
- 鈴木淳子 (2011). 質問紙デザインの技法 ナカニシヤ出版
- Wakita, Ueshima & Noguchi (2012). Psychological Distance between Categories in the Likert Scale: Comparing Different Numbers of Options. *Educational and Psychological Measurement*.

—2011.12.8受稿—