

Skala Likert, Bias Penggunaan dan Jalan Keluarnya

Bilson Simamora

Program Studi Ilmu Administrasi Bisnis, Institut Bisnis dan Informatika Kwik Kian Gie, Jakarta

*Correspondence address: bilson.simamora@kwikkiangie.ac.id

Abstrak: Many researchers refer to any graded scale as a Likert scale. In addition to misnaming, this assumption can lead to errors in the data scale and analysis technique. This paper aims to explain the instrument scales in the survey and the data types they generate. Thus, the position of the Likert scale compared to other measurement scales becomes apparent so that researchers can give the right name to the scale they use and choose the proper data analysis technique according to its type.

Keywords: Skala Likert, bias penggunaan, bias posisi, skala grafis, jenis data

Cite: Simamora, B. (2022). Skala Likert, Bias Penggunaan dan Jalan Keluarnya. *Jurnal Manajemen*, 12(1), 84–93.
<https://doi.org/10.46806/jman.v12i1.978>

Copyright © Jurnal Manajemen.

All rights reserved



1. Latar Belakang

Banyak peneliti menyatakan setiap skala yang pilihannya bertingkat sebagai skala Likert. Selain salah penamaan, anggapan demikian juga dapat menyebabkan kesalahan dalam memperlakukan data sebagai data interval dan memilih teknik analisis yang digunakan.

Selain menjelaskan kesalahan skala Likert, tulisan ini juga memberi alternatif skala pengukuran pengganti skala Likert apabila peneliti ingin memperoleh data interval yang diolah dengan teknik analisis yang membutuhkan data metrik, seperti regresi dan SEM. Kalau tetap menggunakan skala Likert, tulisan ini memberi alternatif teknik analisis yang sesuai.

2. Skala Likert

2.1. Bentuk

Skala Likert ditemukan Rensis Likert tahun 1932 untuk mengukur sikap. Dalam skala tersebut diajukan pernyataan atau pertanyaan dan meminta persetujuan (*agreement*) responden atas pertanyaan atau pernyataan yang diajukan (Hair et al., 2003). Skala ini banyak digunakan sesuai bentuk aslinya. Namun, banyak juga peneliti yang melakukan penyalahgunaan dengan menyatakan setiap pilihan yang berjenjang sebagai skala Likert. Contoh berikut ini adalah *itemized scale*, bukan skala Likert.

“Bagaimana rasa kopi X menurut pengalaman konsumsi anda?”

 Sangat tidak enak Tidak enak Netral Enak Sangat enak

Penjelasan berikutnya didasarkan pada perspektif bahwa respon pada skala Likert adalah persetujuan responden atas pernyataan atau pernyataan. Contoh berikut adalah skala Likert tujuh tingkat.

1. I'm confident that I can do an excellent job on the assignments and tests during my study in this university

Disagree Very disagree Nutral Strongly disagree Very Agree Very Agree Strongly Agree
2. I expect to do well in this university

Strongly agree Very agree Agree Neutral Disagree Very Disagree Strongly Disagree
3. I'm certain I can master the skills being taught in this university

Strongly disagree Very disagree Disagree Neutral Agree Very Agree Strongly Agree
4. Considering the difficulty of materials offered in this university, the teacher, and my skills, I think I will do well in this university

Strongly agree Very agree Agree Neutral Disagree Very Disagree Strongly Disagree
5. I aim to completely master any material presented in this university

Neutral Strongly disagree Disagree Very Disagree Agree Very Agree Strongly Agree

Karena pilihan jawaban berjenjang, maka setiap pilihan jawaban bisa diberi skor. Skor 1 bisa diberi pada jenjang jawaban terendah, misalnya 'sangat tidak setuju', bisa pula pada jenjang jawaban tertinggi, misalnya 'sangat setuju', asal dilakukan secara konsisten.

2.2. Potensi Bias Urutan Pertanyaan

Bias adalah respon atau jawaban yang tidak sesuai dengan sikap responden sesungguhnya. Dalam statistic bias ini disebut juga error pengukuran (*measurement error*).

Dalam skala Likert, walaupun variabel pengamatan valid dan pertanyaan dibentuk dalam format yang tepat, bias disebabkan oleh urutan pertanyaan atau pernyataan dan posisi jawaban dalam skala. Kedua sumber bias ini perlu diperhatikan peneliti.

Urutan pertanyaan perlu diacak, agar setiap pertanyaan mendapat perhatian yang sama. Ada kecenderungan responden memberi perhatian lebih banyak pada pertanyaan-pertanyaan yang diposisikan di awal dari pada di akhir kuesioner. Pada contoh di atas, apabila urutan tidak diacak, pernyataan "ATM bank A banyak", akan memperoleh perhatian lebih tinggi (lebih serius) dibanding "Reputasi bank A baik". Kalau jumlah pertanyaan puluhan, misalnya 40 pertanyaan, pertanyaan terakhir bisa dijawab asal-asalan. Karena itu, urutan pertanyaan dasar (pertanyaan terkait variabel-variabel penelitian) perlu diacak.

2.3. Potensi Bias Posisi Jawaban

Hindari keseragaman pilihan jawaban dengan mengacak titik ekstrim positif dan negatif. Seperti pada contoh, letak sangat setuju, sebagian di sebelah kiri, sebagian lagi di sebelah kanan. Variasi posisi ini perlu untuk menghindari pemberian jawaban seragam untuk semua pilihan karena kemalasan responden untuk memeriksa pertanyaan satu per satu. Contoh berikut ini dapat menyebabkan bias posisi jawaban.

1. I'm confident that I can do an excellent job on the assignments and tests during my study in this university

Strongly disagree *Very disagree* *Disagree* *Neutral* *Agree* *Very Agree* *Strongly Agree*

2. I expect to do well in this university

Strongly disagree *Very disagree* *Disagree* *Neutral* *Agree* *Very Agree* *Strongly Agree*

3. I'm certain I can master the skills being taught in this university

Strongly disagree *Very disagree* *Disagree* *Neutral* *Agree* *Very Agree* *Strongly Agree*

4. Considering the difficulty of materials offered in this university, the teacher, and my skills, I think I will do well in this university

Strongly disagree *Very disagree* *Disagree* *Neutral* *Agree* *Very Agree* *Strongly Agree*

5. I aim to completely master any material presented in this university

Strongly disagree *Very disagree* *Disagree* *Neutral* *Agree* *Very Agree* *Strongly Agree*

Untuk menghindari bias posisi jawaban, posisi pilihan jawaban bisa diacak seperti contoh skala Likert di atas.

Ketiga, hindari memberi nilai pada setiap pilihan. Secara psikologi, manusia memiliki sifat menyukai nilai tinggi dibanding rendah (kalau nilainya 1, 2, 3, 4, 5) atau lebih suka nilai positif daripada negatif (kalau nilainya -2, -1, 0, 1, 2). Penggunaan nilai secara tidak sadar, dapat menggiring pilihan pada jawaban nilai lebih tinggi atau positif. Karena itu, pilihan jawaban lebih baik disertai kotak, lingkaran atau garis kosong, seperti pada contoh-contoh di atas. Jadi, pertanyaan seperti ini sebaiknya dihindari:

“Bank A memiliki banyak ATM”

- 1: Sangat tidak setuju
- 2: Sangat setuju
- 3. Netral
- 4: Setuju
- 5. Sangat setuju

“Reputasi Bank A buruk”

- | | | | | |
|---------------|--------|--------|--------------|--------|
| -2 | -1 | 0 | 1 | 2 |
| Sangat setuju | Setuju | Netral | Tidak setuju | Setuju |

2.4. Jenis Data

Apakah data skala Likert bersifat ordinal ataukah interval telah lama diperdebatkan para ahli (Sullivan dan Artino, 2013). Sebagian ahli (misalnya Jamieson, 2022; Zikmund, 2000) menganggap ordinal, sebagian lagi menganggap interval (misalnya Cooper dan Schlindler, 2018) mengatakan bahwa skala Likert dapat diperlakukan sebagai data interval.

Kelompok ketiga menyatakan bahwa data skala Likert adalah ordinal, namun dapat diperlakukan sebagai data interval dengan ketentuan, pertama, jumlah pilihan banyak, bukan tiga, lima atau tujuh. Hu dan Leong (2017) mengajukan 11 pilihan. Masalahnya, semakin banyak pilihan jawaban, semakin sulit mencari kata-kata yang dipahami secara umum.

Dengan menggunakan bahasa Inggris, umumnya para peneliti mengajukan pilihan tujuh tingkat berikut biasa dipakai: *extremely disagree*,

strongly disagree, disagree, neither agree nor disagree, agree, strongly agree, extremely agree. Dalam bahasa Indonesia padanannya adalah: amat sangat tidak setuju, sangat tidak setuju, tidak setuju, netral, setuju, sangat setuju, amat sangat setuju. Kalau pilihannya sebelas, kata-kata apa yang mewakili, sehingga pilihan di sisi kanan dan kiri 'netral' seimbang? Dalam bahasa Indonesia, usulan Hu dan Leong (2017) sulit diwujudkan.

Kedua, instrumen pengukuran melibatkan sejumlah item (*multi-item measurement*), bukan item tunggal (*single item measurement*) (Grace-Marthin, n.d.; Carificio; Perla, 2008; Sullivan & Artino, 2013). Jadi, apabila pengukuran memiliki banyak item (*multi-items measurement*), data skala Likert dapat diperlakukan sebagai data interval dan diolah dengan teknik-teknik parametrik.

Contoh berikut ini multi-items measurement dari Eliot dan Murayama (2008) untuk mengukur *mastery goals*.

"My aim is to completely master the material presented in this class"

Disagree Very disagree Neutral Strongly disagree Very Agree Very Agree Strongly Agree

"I am striving to understand the content of this course as thoroughly as possible"

Neutral Strongly disagree Disagree Very Disagree Agree Very Agree Strongly Agree

"My goal is to learn as much as possible"

Strongly agree Very agree Agree Neutral Disagree Very Disagree Strongly Disagree

Contoh *single-item measurement*:

"Anda ... pada sistem pemilihan tertutup proporsional"

Sangat tidak setuju Tidak setuju Netral Setuju Sangat setuju

Ketiga, boleh mengolah data skala Likert dengan teknik-teknik parametrik, namun jangan sekali-kali interpretasi hasil analisis pemusatan (*central tendency analysis*), kecuali kalau data berdistribusi normal (Grace-Marthin, n.d.; Sullivan & Artino, 2008). Dengan demikian, peneliti tidak dibolehkan melakukan uji apapun berdasarkan rata-rata kalau prinsip normalitas data tidak terpenuhi.

Kelompok keempat mengakui bahwa data skala Likert adalah ordinal. Pengujian dengan teknik parametric dan non metric memberikan hasil yang sama (Harpe, 2015; Normal, 2010). Apabila data skala Likert ditransformasi menjadi menjadi data interval, hasil analisis teknik parametric sebelum dan setelah data ditansformasi adalah sama (Mondiana, et al., 2018).

3. Alternatif Skala Likert

Uraian tentang jenis data skala Likert di atas menunjukkan masih adanya perdebatan tentang apakah tergolong sebagai data ordinal ataukah interval. Jalan keluar dalam artikel ini adalah menggunakan skala yang dipercaya menghasilkan data interval, yaitu *semantic differential scale* dan *continuous scale*.

3.1. Semantic Differential

Semantic differential juga dapat mengukur sikap dan persepsi. Skala ini berisikan dua kata sifat berlawanan pada kutubnya. Di antara kedua kutub disediakan ruang untuk memberi respon.

Berikan penilaian anda pada bank X

ATM bank sangat banyak	:_:_:_:_:_:_:_:_:	ATM bank sangat sedikit
Pelayanan bank buruk	:_:_:_:_:_:_:_:_:	Pelayanan bank baik
Lokasi bank strategis	:_:_:_:_:_:_:_:_:	Lokasi bank tidak strategis
Hadiah bank banyak	:_:_:_:_:_:_:_:_:	Hadiah bank sedikit
Reputasi bank baik	:_:_:_:_:_:_:_:_:	Reputasi bank buruk

Ada tiga dimensi dasar respon. Pertama, dimensi sifat, seperti: manis-pahit, baik-buruk, menyenangkan-menyusahkan. Kedua, dimensi potensi, seperti: penuh-kosong, kuat-lemah, tinggi-rendah; banyak-sedikit. Ketiga, dimensi aktivitas, seperti: tenang-ribut, cepat-lambat, tepat-menyimpang.

Yang perlu diperhatikan dalam penggunaan skala ini adalah:

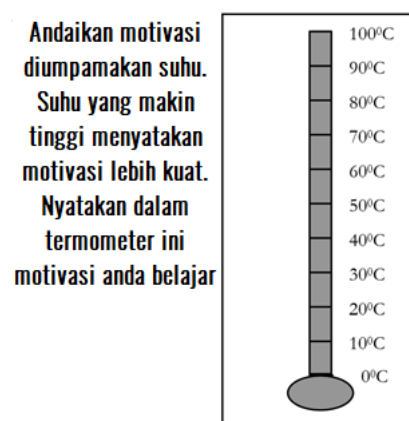
- Orientasi kutub kanan dan kutub kiri beragam, kadang positif, kadang negatif, seperti pada contoh di atas. Tujuannya adalah meminimalkan bias posisi jawaban. Kalau orientasi kutub seragam, ada kemungkinan responden memberikan jawaban condong ke kiri, tengah atau kanan.
- Jumlah ruang jawaban adalah ganjil, misalnya tiga, lima, tujuh, dan sembilan. Tidak ada ketentuan berapa jumlah skala terbaik. Namun, semakin banyak ruang, jawaban responden akan semakin terwakili, namun semakin sulit menilai ruang yang paling mewakili.

Keuntungan skala ini adalah data dianggap interval dan mudah digunakan. Kerugiannya adalah tidak ada penjas setiap ruang, terutama bila jumlahnya lima ke atas. Peneliti dapat memberi skor pada setiap ruang. Ruang di tengah dapat dianggap netral. Namun, pada ruang yang condong ke positif atau negatif tidak dapat dideskripsikan setiap angka mewakili apa.

3.2. Graphic Rating Scales

Dengan skala ini, responden diminta respon dengan menandai poin pada skala yang digunakan. Ada dua model yang populer, yaitu skala thermometer dan skala garis.

Graphic Rating Scales Model Termometer



Dengan skala garis, respon dinilai dengan cara mengukur panjang garis mulai bagian yang ditandai ke ujung kiri. Panjang garis dimaksud dibandingkan dengan panjang garis dan dinyatakan dalam angka desimal.

Contoh *Graphic Rating Scales*

Evaluasilah atribut Bank A dengan memberikan tanda 'X' pada garis yang posisinya mencerminkan respon anda

	Sangat tidak setuju	Sangat setuju
ATM sangat banyak	_____	
Pelayanan memuaskan	_____	
Lokasi strategis	_____	
Hadiahnya banyak	_____	
Reputasi baik	_____	

Skala thermometer dianggap menghasilkan data interval, sedangkan skala garis menghasilkan data rasio. Data dari keduanya dapat diolah dengan statistik parametrik. Perlu diperhatikan agar perbedaan kata di kedua ujung garis jangan terlalu ekstrim, agar tidak terjadi bias jalan tengah, yaitu kecenderungan memberi jawaban yang tidak terlalu ekstrim.

4. Penutup

Skala Likert adalah skala yang sangat populer bagi peneliti. Kepopuleran tersebut menyebabkan salah pengertian bahwa skala Likert adalah nama bagi semua skala yang pilihan jawabannya diitemisasi (*itemized scales*). Skala Likert hanya salah satu skala yang dapat digunakan dalam pengukuran. Untuk menghindari kontroversi terkait jenis data yang dihasilkan ordinal ataukah interval, peneliti dapat memiliki skala lain yang dipercaya menghasilkan data interval dan rasio.

Referensi

Aaker, D., Kumar, V. and Day, G. (2004). *Marketing Research*. John Wiley & Sons Inc., New York.

Churchill, G.A., & Iacobucci, D. (2005). *Marketing Research Methodological Foundations*. Thomson South-Western.

Cooper, D.R., & Emory, C.W. (1995). *Business Research Methods*. 5th Edition. Irwin.

Cooper, D. R., & Schindler, P.S. (2018). *Business Research Method*. 13th Edition. McGraw-Hill, Inc.

Corificio, J., Perla, R. (2008). Resolving the 50-year debate around using and misusing Likert scales. *Medical Education*, 40, 1150–1152. doi: 10.1111/j.1365-2923.2008.03172.x

Elliot, A. J., & Murayama, K. (2008). On the measurement of achievement goals: Critique, illustration, and application. *Journal of Educational Psychology*, 100(3), 613–628. <https://doi.org/10.1037/0022-0663.100.3.613>.

Grace-Marthin, K. (n.d.). Can likert scale data ever be continuous? The Analysis Factor [Statistic Blog]. Retrieved June 12, 2023, from <https://www.theanalysisfactor.com/can-likert-scale-data-ever-be-continuous/>

Hair, Bush dan Ortinau. 2003. *Marketing Research Within a Changing Information Environment*. Second Edition. McGraw-Hill/Irwin

Harpe, S. E. (2015). How to analyze Likert and other rating scale data. *Currents in Pharmacy Teaching and Learning*, 7(6), 836–850. <https://doi.org/10.1016/j.cptl.2015.08.001>

Olsson, U. (2022). Power Properties of Ordinal Regression Models for Likert Type Data. *Practical Assessment, Research, and Evaluation*, 27, Article 6. DOI: <https://doi.org/10.7275/er28-5992> Available at: <https://scholarworks.umass.edu/pare/vol27/iss1/6>

Jamieson, S. (2022, September 26). Likert scale. *Encyclopedia Britannica*. <https://www.britannica.com/topic/Likert-Scale>

Malhotra, Naresh K. (2010). *Marketing Research: An Applied Orientation*. 6th Edition. Pearson Education.

Mondiana, Y.Q., Pramoedyo, H., Sumarminingsih, E. (2018). Structural Equation Modeling on Likert Scale Data With Transformation by Successive Interval Method and With No Transformation. *International Journal of Scientific and Research Publications*, 8(5). <http://dx.doi.org/10.29322/IJSRP.8.5.2018.p7751>

Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education*, 15(5), 625–632. <https://doi.org/10.1007/s10459-010-9222-y>

Sekaran, U. (2003). *Research Methods for Business: A Skill Building Approach*. 4th Edition. John Wiley.

Sullivan, G. M., & Artino, A. R., Jr (2013). Analyzing and interpreting data from likert-type scales. *Journal of graduate medical education*, 5(4), 541–542. <https://doi.org/10.4300/JGME-5-4-18>

WHO Collaborating Center for Mental Health. (1998). WHO (Five) Well-Being Index. World Health Organization.

Wu, H., & Leong, S-O. (2017). Can Likert Scales be Treated as Interval Scales?—A Simulation Study. *Journal of Social Service Research*, 43(4), 527-532. <https://doi.org/10.1080/01488376.2017.1329775>

Zikmund, William. G. (2000). *Exploring Marketing Research*. Seventh Edition. The Dryden Press.