Reducing Survey Bias Through Cultural and Response-Style Awareness

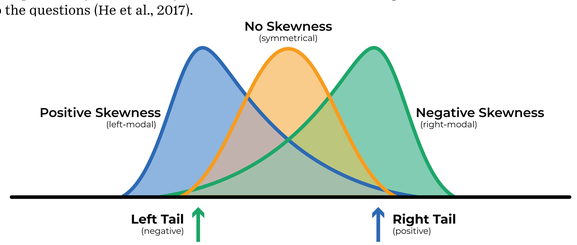
Prepared by Julie Shirah, Senior Data Analyst

World Education Services

Introduction

How does culture influence responses to surveys?

- ***Cultural bias*** can lead to response-style differences on questionnaires which can result in a set of systematic errors that undermines the validity of questionnaires administered in different cultures (He et al., 2017).
- One type of cultural bias is **method bias**: Individuals from different cultures have different response styles to Likert scale questions (He et al., 2017). This leads to differences in ratings that are not due to differences in actual ratings to the questions (He et al., 2017).



Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree

Note: Negative skew results in more responses to the left tail (negative) side of the distribution. Positive skew results in more responses to the right tail (positive) side of the distribution.

Moss and Vijayendra (2020) explored three different types of method bias:

1. **Acquiescent response style**: Participants from the Middle East and Africa tend to select "Agree" or "Strongly Agree" on a Likert scale.

Strongly Disagree	Disagree	Neither Agree nor Disagree	**Agree**	**Strongly Agree**

2. **Extreme response style**: Participants from Latin America tend to respond by selecting ratings at extreme ends of the scale (e.g., "Strongly Agree" or "Strongly Disagree").

Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	**Strongly Agree**

3. **Middle response style**: Participants from Asian countries tend to select ratings in the middle of the Likert scale (e.g., "Neither Agree nor Disagree," "Agree," "Disagree").

Strongly Disagree	**Disagree**	**Neither Agree nor Disagree**	**Agree**	Strongly Agree

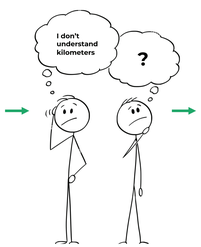
Methods to Reduce Bias

What techniques reduce bias in survey research?

1. **Use of Visuals:** Emojis have been widely used globally and may be a universal language that can bridge communication gaps between cultures because of their pictorial nature (Kejriwal et al., 2021; Kimura-Thollander & Kumar, 2019).

2. **Cognitive Pre-testing:** This is the process in which a few participants are asked questions orally after they complete a pilot survey to ascertain whether they understood the survey questions and thoughts they had while answering the survey (Willis & Miller, 2011*).

How many kilometers do you commute to campus every day round trip?
A. 1-5 kilometers
B. 6-10 kilometers
C. 11-15 kilometers
D. 16-20 kilometers
E. I use another mode of transportation
F. I am an online-only student
G. Other: _____



How many kilometers/miles do you commute to campus every day round trip?
A. 1-5 kilometers/0.6-3.1 miles
B. 6-10 kilometers/3.7-6.2 miles
C. 11-15 kilometers/6.8-9.3 miles
D. 16-20 kilometers/9.9-12.4 miles
E. I use another mode of transportation
F. I am an online-only student
G. Other: _____

*Example is an expanded idea from He and van de Vijver (2012)

3. **Standardization**: Scores can be standardized in two ways: (1) Individual participant scores are adjusted by subtracting the mean of all their responses from each of their individual responses (Fischer, 2004). (2) The mean score for a question across all participants in a culture can be subtracted from each individual response (Fischer, 2004).

Summary of All Techniques

Technique	Pros	Cons
1. Use of Visuals (Kejriwal et al., 2021; Kimura-Thollander & Kumar, 2019)	Emojis are used globally and may be a universal language that can bridge communication gaps between cultures.	Emojis' meaning or prevalence may be different cross-culturally.
2. Cognitive Pre-testing (Willis & Miller, 2011)	Ensures the survey is well-designed and understood by target audience.	Lengthens the survey design process, and pilot participants may not be representative of entire survey audience.
3. Standardization (Fischer, 2004)	A technique that can be used without altering survey design (e.g., visuals in place of Likert scale).	May remove variation that is due to communication style differences across cultures.

Application of Standardization Technique

Example Scenario:

A course is delivered to students from the Middle East, Latin America, and Asia. All students respond to a course evaluation that contains the following statements* on a 5-point Likert scale (Strongly Agree – Strongly Disagree):

1. "The instructor was well prepared for class."
2. "The instructor communicated clearly and was easy to understand."
3. "The instructor encouraged student participation in class."

*All quotes are from University of Wisconsin-Madison. (n.d.).

Student (P)	Q1	Q2	Q3
P1 (Middle East)	5	4	4
P2 (Latin America)	5	5	1
P3 (Asia)	3	4	2

		Mean Centering (MC)		
Student (P)	**Mean***	**MC Q1**	**MC Q2**	**MC Q3**
P1 (Middle East)	4.33	5 – 4.33 = **0.66**	4 – 4.33 = **-0.33**	4 – 4.33 = **-0.33**
P2 (Latin America)	3.66	5 – 3.66 = **1.33**	5 – 3.66 = **1.33**	1 – 3.66 = **-2.66**
P3 (Asia)	3	3 – 3 = **0**	4 – 3 = **1**	2 – 3 = **-1**

*For each participant across their responses to all three questions.

		Z Score (Z)		
Student (P)	**Standard Deviation***	**Z Q1**	**Z Q2**	**Z Q3**
P1 (Middle East)	0.57	0.66 / 0.57 = **1.15**	-0.33 / 0.57 = **-0.58**	-0.33 / 0.57 = **-0.58**
P2 (Latin America)	2.31	1.33 / 2.31 = **0.58**	1.33 / 2.31 = **0.58**	-2.66 / 2.31 = **-1.15**
P3 (Asia)	1	0 / 1 = **0**	1 / 1 = **1**	-1 / 1 = **-1**

*For each participant across their responses to all three questions.

Interpretation: Z scores indicate a Mean of 0 and a Standard Deviation of 1. Therefore, Z scores can be interpreted as the number of standard deviations from the mean on a bell curve; the higher the Z score, the further from the mean.

Impact: For Q1 this indicates that a Middle Eastern student rated the professor as more prepared than a Latin American student did, and an Asian student rated the professor as average. For Q2 this indicates that a Middle Eastern student rated the professor as less easy to understand than a Latin American student rated the professor, and an Asian student rated the professor as easy to understand.

References

Fischer, R. (2004). Standardization to account for cross-cultural response bias: A classification of score adjustment procedures and review of research in JCCP. *Journal of Cross-Cultural Psychology*, *35*(3), 263-282. https://doi.org/10.1177/0022022104264122

He, J., & van de Vijver, F. (2012). Bias and equivalence in cross-cultural research. *Online Readings in Psychology and Culture*, *2*(2). https://doi.org/10.9707/2307-0919.1111

He, J., van de Vijver, F. J. R., Fetvadjiev, V. H., de Carmen Dominguez Espinosa, A., Adams, B., Alonso-Arbiol, I., Aydinli-Karakulak, A., Buzea, C., Dimitrova, R., Fortin, A., Hapunda, G., Ma, S., Sargautyte, R., Sim, S., Schachner, M. K., Suryani, A., Zeinoun, P., & Zhang, R. (2017). On enhancing the cross-cultural comparability of Likert-Scale personality and value measures: A comparison of common procedures. *European Journal of Personality*, *31*(6), 642-657. https://doi.org/10.1002/per.2132

Kejriwal, M., Wang, Q., Li, H., and Wang, L. (2021). An empirical study of emoji usage on Twitter in linguistic and national contexts. *Online Social Networks and Media*, 24. https://doi.org/10.1016/j.osnem.2021.100149

Kimura-Thollander, P., & Kumar, N. (2019). Examining the "Global" language of emojis: Designing for cultural representation. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19). *Association for Computing Machinery*, New York, NY, USA, Paper 495, 1–14. https://doi.org/10.1145/3290605.3300725

Moss, F., & Vijayendra, B. (2020). *When Difference doesn't mean Different: Understanding cultural bias in global CX programmes*. Ipsos. Retrieved from https://www.ipsos.com/sites/default/files/ct/publication/documents/2020-09/understandingculturalbias_2.pdf

Willis, G., & Miller, K. (2011). Cross-cultural cognitive interviewing. *Field Methods*, *23*(4), 331-341. https://doi.org/10.1093/poq/nfu092

University of Wisconsin-Madison. (n.d.). Best Practices and Sample Questions for Course Evaluation Surveys. https://assessment.wisc.edu/best-practices-and-sample-questions-for-course-evaluation-surveys/

Bell Curve Image: https://dataanalyze.wordpress.com/skewness/

Stick Figure Image: https://www.alamy.com/stock-photo/two-speech-bubble.html?sortBy=relevant

Face Rating Scale: https://stock.adobe.com/search?k=smiley+face+scale

Scan to learn more

wes.org