Reducing Survey Bias Through Cultural and Response-Style Awareness

Overview of Likert Scale Design and Statistical Interpretation

Survey research often contains Likert scale items (1 Strongly Agree – 5 Strongly Disagree), which easily quantify a participant's perception of research questions. Likert scales vary in number of anchors (4-point versus 5-point scale), numerical representation (starting at 1 or 5), and placement of text anchors (endpoints only versus at all points; South et al., 2022). It is a matter of debate whether Likert scales should be used as interval or ordinal variables in data analysis (South et al., 2022). Most researchers treat Likert scale variables as an interval level of measurement; for instance, researchers could report means and standard deviations and use continuous diagram representations (South et al., 2022). Fewer researchers interpret Likert scales as an ordinal level of measurement; for instance, researchers could report median or quartiles and use discrete diagram representations (South et al., 2022).

Overview of Cultural Response-Style Bias

Cultural bias indicates that the culture a person was brought up in can influence how they interact with the world. It 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). It implies that differences in ratings might not reflect genuine differences in the constructs of interest (He et al., 2017). For instance, if researchers were interested in investigating international student interest in a new university service, differences in ratings might not reflect differences in interest, but differences in the ways different cultures respond to the Likert scale. According to He et al. (2017), there are a couple of ways a survey can be culturally biased: method and construct. *Method bias* occurs when individuals from different cultures have different response styles to Likert scale questions (He et al., 2017); this leads to differences in scale ratings that are not due to differences in the construct being measured, which reduces validity (He et al., 2017). *Construct bias* takes place when individuals from different cultures interpret the meaning of the construct differently (He et al., 2017); for instance, some cultures measure distance in kilometers whereas others measure distance in miles (He & van de Vijver, 2012).

Moss and Vijayendra (2020) outlined three types of method bias: acquiescence response style, extreme response style, and middle response style. Individuals from the Middle East and Africa tend to respond in an *acquiescence response style* by selecting "Agree" or "Strongly Agree" on a

Likert scale, no matter the question. Individuals from Latin America tend to respond in an *extreme response style* by selecting anchors at extreme ends of the scale (e.g., "Strongly Agree" or "Strongly Disagree"). Individuals from Asian countries tend to respond in the *middle response style* by selecting anchors in the middle of the Likert scale (e.g., "Neither Agree nor Disagree," "Agree," "Disagree").

Techniques for Reducing Bias

Tracking historical trends. Participants can be grouped by the country they are a citizen of, and scores for individual countries can be compared over time. For example, if you asked, "How likely are you to recommend University/Institution to a friend or colleague?" on a scale of 0 – 10, then you could aggregate scores of participants for individual countries (e.g., China, India, Nigeria, etc.) across time (e.g., yearly or quarterly). This method is beneficial because it allows you to track trends within each country over time and avoid cultural bias (Moss & Vijayendra, 2020). A limitation, however, is not having an aggregate view of the data (Moss & Vijayendra, 2020).

Use of visuals. Both facial emojis (Rintoul, 2019) and star ratings (South et al., 2022) can be used instead of traditional Likert scale anchors. Emojis are prevalent globally; because they are visuals, they function as a universal language that can bridge communication gaps between cultures (Kejriwal et al., 2021; Kimura-Thollander & Kumar, 2019). A caution regarding the use of emojis in surveys: their meaning or prevalence may be different cross-culturally (Kejriwal et al., 2021). The number and popularity of emojis individuals use differ cross-culturally, and some emojis are culture-specific (Kejriwal et al., 2021).

Cognitive pre-testing. Cognitive interviews can be used to pilot a survey before distributing it widely. 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). Participants can also be asked what they think specific words or questions mean, which can help researchers evaluate whether a question is difficult to understand (Sha, 2020). After the cognitive pre-testing is complete, researchers can then modify the survey based on trouble areas or misunderstandings that were brought up during the interviews. For instance, a survey might ask, "How many kilometers is your commute to campus every day?" Participants might state during a cognitive interview that they do not understand kilometers but do know miles. Researchers might modify the question to have answer choices that include both kilometers and miles.

References

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/en/when-difference-doesnt-mean-different-understanding-cultural-bias

Rintoul, D. (2019). Visual and animated response formats in web surveys: Do they produce better data, or is it all just fun and games? *University of Wollongong*. Retrieved from https://ro.uow.edu.au/theses1/714/

Sha, M. (with Aizpurua, E.). (2020). Pretesting methods in cross-cultural research (Gabel, T., Ed.), *The essential role of language in survey research* (pp. 132). RTI Press.

South, L., Saffo, D., Vitek, O., Dunne, C. and Borkin, M.A. (2022). Effective use of Likert scales in visualization evaluations: A systematic review. *Computer Graphics Forum*, *41*(3). https://doi.org/10.1111/cgf.14521

Willis, G., & Miller, K. (2011). Cross-cultural cognitive interviewing: Seeking comparability and enhancing understanding. *Field Methods*, *23*(4), 331-341. https://doi.org/10.1177/1525822X11416092