Likert Versus Cronbach's Psychometric Thresholds: Reducing Error and Maximizing Agricultural Education's Scholarship Impacts

Robert Strong
Texas A&M University

James R. Lindner
Auburn University

Bruce Herbert
Texas A&M University

Nicole Ray
Fresno State University

Clare Hancock
Auburn University

Karissa Palmer
Texas A&M University

Primary author:
Robert Strong
Texas A&M University
2116 TAMU
College Station, TX. 77843
979-845-1139
robert.strongjr@ag.tamu.edu

Introduction/Need for Research

Instrumentation is a critical function in measuring social and behavioral science impacts on stakeholders, teachers, and change agents (Thüm et al., 2014). Inquiries on instrument quality offer researchers' evidence of the best instrumentation tool to use (Dillman et al., 2014). Internal validity and reliability have long been considered social sciences' quality gatekeepers (Kodheli, et al., 2021). Field (2013) indicated internal validity is as simple as the instrument measures what it is designed to measure. Reliability is the internal consistency of the instrument results when utilized with a similar population (Cronbach, 1951). Warmbrod (2014) recommended that agricultural education researchers document the empirical evidence proving an instrument is valid and reliable. *Priority 2* of our *National Research Agenda* suggested examining research practices to advance agricultural teaching processes (Lindner et al., 2016).

Conceptual or theoretical Framework

The theory of planned behavior (Ajzen, 1991) framed this study by utilizing the constructs; subjective norms, attitudes, perceived behavioral control, and resulting behavior. Subjective norms are beliefs individuals will endorse and accomplish a specific behavior. Attitudes are developed paradigms of thinking that result in one's behavior. Perceived behavioral control is the discernment of the complexity of carrying out a preferred behavior (Ajzen, 1991). The purpose was to investigate data collection instrument development and reporting subjective norms and resulting behavior in agricultural education literature over the last five years. Research objectives were: 1) determine the number of instrumentation articles reporting construct reliability, and 2) determine the numerical range of items used to measure constructs.

Methodology

Wright et al. (2007) posited a systematic review uses a comprehensive search based on explicit protocols to review existing literature with a synthesis of data focusing on key questions. Systematic reviews are five steps; identify the critical question, formulate search parameters, systematically search databases, analyze data, and data summary interpretation (Lee et al., 2021). Using the five steps, authors systematically reviewed all articles from *Advancements in Agricultural Development* (*AAD*), *Journal of Agricultural Education* (*JAE*), *Journal of Extension* (*JOE*), and *The Journal of Agricultural Education and Extension* (*TJAEE*) from 2018 to 2022. The authors reviewed eight hundred ninety-six (N = 896) articles from the four publications.

Results/findings

JAE reported forty-seven (N = 47) articles in 2022 and seventeen (n = 17, 36.17%) reported data collection reliability coefficients. Of the seventeen articles, the numerical range of statements or questions utilized to measure constructs extended from 1 to 10. *JAE* reported in 2021 (N = 73) published articles and thirty-two (n = 32, 43.83%) utilized data collection reliability coefficients. Of those thirty-two, the numerical range of statements or questions was 1 to 19. In 2020, *JAE* reported eighty-three (N = 83) articles published and forty-one (n = 41, 49.39%) that tested reliability coefficients. Statements numerically ranged from 1 to 32. Thirty-two was the extreme outliner. In 2019, (N = 70) articles were published and (n = 36, 51.43%) implemented collection instruments. Statements ranged from 1 to 12. *JAE* reported (N = 80) articles and (n = 44, 55%) articles used data collection instruments in 2018. The range of statements was from 1 to 20.

JOE published thirty-four (*N* = 34) articles in 2022 and five (*n* = 5, 14.70%) reported testing reliability coefficients. The range of statements or questions was 1 to 10. *JOE* had eighty- two (*N* = 82) articles published in 2021 and six (*n* = 6, 7.31%) reported data collection reliability coefficients. One to twelve was the range of statements or questions. In 2020, *JOE* reported sixty-eight (*N* = 68) articles and nine (*n* = 9, 13.23%) had data reliability coefficients. The numerical range of statements or questions utilized to measure constructs ranged from 1 to 7. *JOE* published (*N* = 77) articles in 2019 and (*n* = 13, 68.88%) indicated the use of data collection instruments. Statements ranged from 1 to 12. In 2018, *JOE* reported (*N* = 78) articles and (*n* = 15, 21.43%) testing reliability coefficients. The range of statements was from 1 to 12.

TJAEE had (*N* = 22) articles published in 2022 and three (*n* = 3, 13.63%) articles had construct reliability coefficients. Statements ranged from 1 to 36. Thirty-six was an outlier given the small number of instrumentation studies in 2022. In 2021, thirty-three (*N* = 33) articles were published and seven (*n* = 7, 21.21%) reported construct reliability coefficients. The numerical range of statements was 1 to 10. There were twenty-five (*N* = 25) articles and three (*n* = 3; 12%) articles that tested construct reliability coefficients in 2020. Statements or questions ranged from 1 to 7. *TJAEE* reported (*N* = 25) articles published in 2019 and (*n* = 9, 36%) indicated the use of data collection instruments and statements ranged from 1 to 8. In 2018, (*N* = 25) articles were published and (*n* = 11, 44%) utilized data collection instruments. Statements ranged from 1 to 10.

AAD began in 2020, and therefore, articles from 2018 and 2019. *AAD* had (*N* = 14) articles published in 2022 and seven (*n* = 7, 50%) articles reported data collection reliability coefficients. The numerical range of statements ranged from 1 to 7. Twenty-seven (*N* = 27) articles were published in 2021 and twelve (*n* = 12, 44.44%) utilized data collection reliability coefficients. The range of statements was from 1 to 7. Twenty- three articles (*N* = 23) were published in 2020 and eight (*n* = 8, 34.78%) reported reliability coefficients. Statements ranged from 1 to 5.

Conclusions

Fewer items produced lower construct reliability coefficients and thus, produced higher levels of error (Likert, 1932). Much of our published scholarship has not utilized instruments to collect data over the last five years; when they have, smaller numbers of items measured constructs.

Implications/recommendations/impact on profession

Likert's (1932) convention in his quintessential work on measuring social variables suggested that for measurements to be reliable an alpha of .9 should be achieved. While Cronbach's (1951) convention postulates that construct reliability of .7 be achieved. Besides, what difference does .2 make anyway? With a threshold of .7, a potential variance of up to 30% exists; subsequently with a threshold of .9, a potential variance of only up to 10% exists (Field, 2013; Saris & Gallhofer, 2007). A difference of 20% variance can be a substantial difference in the power of analysis and interpretation of effect size. As a profession we should seek the highest level of reliability as possible, when possible. When developing an instrument, researchers should include a maximum number of statements and questions and eliminate those that do not contribute to reliability and add additional questions when acceptable levels of reliability are not achieved.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16,* 297–334. https://doi.org/10.1007/BF02310555

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: the tailored design method* (4th ed.). John Wiley & Sons.

Field, A. P. (2013). *Discovering statistics using IBM SPSS Statistics* (4th ed.). SAGE Publications.

Lee, C.-L., Strong, R., & Dooley, K. E. (2021). Analyzing precision agriculture adoption across the globe: A systematic review of scholarship from 1999–2020. *Sustainability, 13*(18), 10295. https://doi.org/10.1080/1389224X.2020.1844767

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 140,* 1–55.

Lindner, J. R., Rodriguez, M. T., Strong, R., Jones, D., & Layfield, D. (2016). Research priority area 2: New technologies, practices, and products adoption decisions. In Roberts, T. G., Harder, A., & Brashears, M. T. (Eds). *American Association for Agricultural Education national research agenda: 2016-2020*. Gainesville, FL: Department of Agricultural Education and Communication.

Kodheli, O., et al. (2021). Satellite communications in the new space era: A survey and future challenges. *IEEE Communications Surveys & Tutorials, 23*(1), 70–109. https://doi.org/10.1109/COMST.2020.3028247

Saris, W. E., & Gallhofer, I. (2007). Estimation of the effects of measurement characteristics on the quality of survey questions. *Survey Research Methods, 1*(1), 29–43. https://doi.org/10.18148/srm/2007.v1i1.49

Thüm, T., Apel, S., Kästner, C., Schaefer, I., & Saake, G. (2014). A classification and survey of analysis strategies for software product lines. *ACM Computing Surveys, 47*(1), 1–45. https://doi.org/10.1145/2580950

Warmbrod, J. R. (2014). Reporting and interpreting scores derived from Likert-type scales. *Journal of Agricultural Education, 55*(5), 30–47. https://doi.org/10.5032/jae.2014.05030

Wright, R. W., Brand, R., Dunn, W., & Spindler, K. (2007). How to write a systematic review. *Clinical Orthopedics and Related Research, 455*, 23–29. https://doi.org/10.1097/BLO.0b013e31802c9098