Basic Steps in Scale Development for Research

*by **Nguyen Hoang Mai Tram**,*
Ho Chi Minh City University of Foreign Languages - Information Technology, Vietnam
Email: tramnhm@huflit.edu.vn - kaylee.tramnguyen@gmail.com

Abstract:
This article outlines a practical four-phase approach to scale development, from construct definition to empirical validation. Phase 1 focuses on item generation; Phase 2 addresses content validity; Phase 3 details data collection procedures; and Phase 4 validates the scale through EFA and CFA. This structured process supports the creation of a reliable, theory-informed instrument for use in various research contexts.

Developing a scale requires a clear and systematic process. A well-designed scale helps ensure that measurement accurately reflects the construct and can be used confidently in both teaching and research contexts. Figure 1 illustrates the four-phase methodology, detailing the process from initial item pool creation through final scale validation.

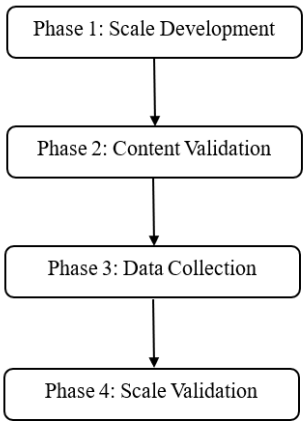

Figure 1. Scale development and validation process

Phase 1: Scale development

This phase begins with a precise definition of the construct, based on previous studies and relevant theories. Once the construct is clear, an initial pool of items is generated by examining how earlier research has operationalized similar concepts. For example, Demirci et al. (2025) developed a scale on motivation for GenAI-based language learning using self-determination theory, which includes autonomy, competence, and relatedness. With the construct and its dimensions mapped out, items may be adapted from validated scales or newly created to suit the specific research context. By the end of phase 1, a comprehensive and theoretically justified item pool should be available.

Phase 2: Content validity

It is important to involve TESOL experts who are not part of the author team, as well as a small group of EFL learners, to review the

initial item pool. These reviewers can provide critical feedback on item wording, construct coverage, clarity, and any potential gaps. Their input helps identify items that may be ambiguous, redundant, or misaligned with the intended construct. The feedback gathered from this trial stage is essential for refining the questionnaire, and any revisions made to the scale should be clearly documented to ensure transparency in the research process.

Phase 3: Data collection

The refined scale is then administered to participants. At this stage, researchers should specify the response format, the language of the questionnaire, and the sampling and recruitment methods. Before beginning the survey, some screening questions help confirm participant eligibility. Clear instructions, consistent administration, and ethical procedures ensure that the data collected is reliable for subsequent analyses.

Phase 4: Scale validation

In this phase, the full sample is divided into two sub-samples for exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) respectively. EFA is first conducted to uncover the underlying factor structure of the scale. Sampling adequacy is assessed using the KMO measure and Bartlett's test of sphericity (Bartlett, 1954). Factors are typically retained when their eigenvalues exceed 1, and total variance explained is above 50% (Kaiser, 1974). Items are generally expected to show factor loadings above 0.5. In practice, multiple criteria can be used to remove poorly performing items, based on both statistical standards and theoretical foundations.

CFA is then conducted on the second sub-sample to validate the factorial model obtained from EFA. Model fit is commonly evaluated using five indicators: χ^2/df (< 5), CFI (> .90), TLI (> .90), RMSEA (< .08), and SRMR (< .08) (Hu & Bentler, 1999; McDonald & Ho, 2002). In addition, reliability evidence in terms of internal consistency should be examined, typically using Cronbach's alpha, with values above .70 considered acceptable (Fornell & Larcker, 1981). Complementary indicators such as Composite Reliability (CR > .70) and Average Variance Extracted (AVE > .50) are also recommended, as they provide additional evidence of construct reliability and convergent validity.

Scale development is a step-by-step process that ensures instruments are both valid and reliable. By moving from clear construct definition to rigorous validation, researchers create tools that accurately capture learning outcomes. In TESOL, systematic scale development supports evidence-based teaching and guides future studies.

References

Bartlett, M. S. (1954). A note on the multiplying factors for various χ 2 approximations. *Journal of the Royal Statistical Society. Series B (Methodological), 16*(2), 296-298. https://doi.org/10.1111/j.2517-6161.1954.tb00174.x

Demirci, H., Kara, M., & Korkmaz, Ö. (2025). Measuring motivation for Generative AI-based language learning: Scale development and validation based on the self-determination theory. *International Journal of Applied Linguistics*. https://doi.org/10.1111/ijal.12868

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. https://doi.org/10.1177/002224378101800104

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*(1), 31–36. https://doi.org/10.1007/BF02291575

McDonald, R.P., & Ho, M.H.R. (2002). Principles and practice in reporting structural equation analyses. *Psychological Methods, 7*(1), 64–82. https://doi.org/10.1037/1082-989X.7.1.64

Nguyen Hoang Mai Tram is a lecturer at Ho Chi Minh City University of Foreign Languages - Information Technology, Vietnam. She is also a PhD Candidate at the Education University of Hong Kong, supported by the Hong Kong PhD Fellowship Scheme (HKPFS). Tram is a member of the management team for the TESOL Research Collaboration Network (TERECONET), a professional learning community supporting early-career researchers in Vietnam. Her research interests include educational technology, informal language learning, and AI in education. Her work has been published in System, Computers in the Schools, and Children and Youth Services Review.