

**A New Dimension of Evidence Synthesis: Developing and Piloting the Gender and Social  
Context Critical Appraisal Tool (GSC-CAT)**

**Carmen N. Benson, PhD**

Executive Director  
Cooking with Kids, Inc.  
[cbenson@cookingwithkids.org](mailto:cbenson@cookingwithkids.org)

**Theresa Pesi Murphrey, PhD**

Professor  
Texas A&M University  
[t-murphrey@tamu.edu](mailto:t-murphrey@tamu.edu)

## **A New Dimension of Evidence Synthesis: Developing and Piloting the Gender and Social Context Critical Appraisal Tool (GSC-CAT)**

### **Introduction**

Agricultural education initiatives contribute to livelihoods, workforce readiness, agricultural productivity, food security, and youth development in low- and middle-income countries (FAO, 2023; World Bank, 2024). Yet access, participation, and benefits are not necessarily experienced equally; they may differ by gender, socioeconomic status, age, disability, ethnicity, and geography (Huyer et al., 2024). Systematic reviews synthesize evidence to guide research, policy, and practice (Bangdiwala, 2024), but conventional critical appraisal tools focus primarily on methodological rigor, with limited attention to whether studies examine these differences. Consequently, methodologically rigorous evidence may still obscure who participates, who benefits, who is excluded, and for whom outcomes differ (Drucza & Abebe, 2018; Huyer et al., 2024). To address this gap, we developed the Gender and Social Context Critical Appraisal Tool (GSC-CAT) to systematically assess how primary studies account for gender and social context in study design, analysis, interpretation, and reporting.

### **Conceptual Framework**

The GSC-CAT is grounded in the Transformative Paradigm. This perspective centers attention on how social, cultural, economic and political contexts affect participation, experiences, and outcomes across groups (Mertens, 2020). In other words, high-quality evaluation depends not only on rigorous methodology, but also on whose lens is used to view an intervention, whose voices are heard throughout the evaluation process and whose power and access affect findings and conclusions. The Transformative Paradigm supports an emerging understanding within the education sector that social context affects participation, experiences, and outcomes and should be taken into account when thinking about programs' effectiveness (Estepp et al., 2022; Huyer et al., 2024). The framework also complements the EVAL Framework by extending evaluation competency beyond methodological rigor to include systematic assessment of gender equality and social inclusion within program evaluation and impact research (Strong et al., 2021).

### **Methodology**

We developed the tool through an iterative review of literature, evaluation frameworks, and practitioner guidance on gender equality and social inclusion in program evaluation (Drucza & Abebe, 2018; Patton, 2022; Tavenner et al., 2022; UN Working Group, 2017). Drawing on established approaches to instrument development (Boateng et al., 2018; DeVellis & Thorpe, 2021), we translated key constructs into nine appraisal criteria with definitions, decision rules, and scoring guidance. The GSC-CAT examines whether and how studies account for representation, participation, social disparities, and power across the research process and classifies studies from GSC-Unaware to GSC-Transformative. We used a systematic review of 79 youth agricultural education studies in low- and middle-income countries as an opportunity to pilot test the tool. Two reviewers independently appraised each study, recorded time required, and compared ratings across criteria to assess feasibility and initial inter-rater reliability. We then refined definitions and scoring guidance and identified areas for further validation.

## Results/Findings

We successfully applied the GSC-CAT to all 79 studies (N=79) included in the systematic review (Benson et al., 2024). When completed alongside traditional quality appraisal, the framework required approximately eight to 11 additional minutes per study, demonstrating its feasibility as a rapid critical appraisal tool. Two independent reviewers achieved 96% agreement across 711 paired criterion ratings (79 studies  $\times$  9 appraisal criteria), demonstrating strong preliminary inter-rater agreement. We used findings from the pilot to refine the operational definitions and scoring guidance and to identify priorities for future validation.

Applying the GSC-CAT identified substantial gaps in the evidence base. Although many evaluations demonstrated sound methodological rigor, few meaningfully considered how gender and social context shaped participation, experiences, and outcomes. Overall, 38 studies (48%) failed to meet the minimum expectation of reporting gender-disaggregated findings and were classified as GSC-Unaware. When broader dimensions of social inclusion, including socioeconomic and demographic characteristics, were considered, this proportion increased to 52 studies (67%). Only 22 studies (28%) articulated explicit GSC objectives, and just seven (9%) considered how power, agency, or decision-making influenced program outcomes. The GSC-CAT distinguished studies that appeared similarly rigorous using conventional appraisal tools but differed substantially in their consideration of representation, inclusion, and power.

## Conclusions

We developed and piloted the GSC-CAT, demonstrating that gender and social context *can be* systematically assessed when appraising primary studies within systematic reviews. The tool provides a practical approach for examining these dimensions alongside methodological quality, without requiring specialized expertise. Its application within a systematic review of youth agricultural education programs demonstrates why these dimensions *should be* assessed. Despite methodological rigor, most studies gave limited attention to differentiated outcomes, limiting what we learn about who benefited, who was excluded, and how social context shaped outcomes. These findings demonstrate that methodological rigor alone may not capture differences in experiences and outcomes across populations. Incorporating gender and social context into critical appraisal can provide a more complete understanding of the evidence.

## Impact on the Profession

Together, these findings suggest that assessing methodological quality alone is insufficient when appraising primary studies in systematic reviews. Integrating gender and social context considerations into critical appraisal can strengthen the equity, relevance, and policy value of evidence syntheses by helping researchers, practitioners, and decision-makers understand not only whether interventions work, but for whom, under what conditions, and why. This represents an emerging opportunity to broaden critical appraisal beyond methodological rigor and strengthen how evidence is interpreted and applied across diverse populations and contexts. Future research should refine and validate the GSC-CAT by examining inter-rater reliability, incorporating expert review, and testing its application across diverse evidence synthesis contexts.

## References

- Bangdiwala, S. I. (2024). *The importance of systematic reviews*. *Journal of Comparative Effectiveness Research*, 13(10), 713–715.  
<https://doi.org/10.1080/17457300.2024.2388484>
- Benson, C. N. (2024). *A state-of-the-evidence review: Impacts of youth agricultural education programs on global poverty, resilience, and food security outcomes* (Doctoral dissertation, Texas A&M University).
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, 6, Article 149.  
<https://doi.org/10.3389/fpubh.2018.00149>
- DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications* (5th ed.). Sage.
- Drucza, K.L., Abebe, L. (2018). *Gender equality and social inclusion in agriculture research for development guidelines*. CIMMYT
- Estepp, C. M., Cline, L. L., & Rodriguez, M. T. (2022). Members' perceptions regarding diversity and inclusion within the American Association for Agricultural Education. *Journal of Agricultural Education*, 63(2), 182–202.  
<https://doi.org/10.5032/jae.2022.02186>
- Food and Agriculture Organization of the United Nations. (2023). *The status of food and agriculture 2023*. FAO.
- Huyer, S., Loboguerrero, A. M., Chanana, N., & Spellman, O. (2024). From gender gaps to gender-transformative climate-smart agriculture. *Current Opinion in Environmental Sustainability*, 67, 101415. <https://doi.org/10.1016/j.cosust.2024.101415>
- Mertens, D. M. (2020). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods* (5th ed.). Sage.
- Patton, M. Q. (2022). *Utilization-focused evaluation* (5th ed.). Sage.
- Strong, R., Dooley, K. E., Murphrey, T. P., Strong, J., Elbert, C. D., & Baker, M. (2021). The EVAL framework: Developing impact evaluation scholars. *Advancements in Agricultural Development*, 2(3), 1–13. <https://doi.org/10.37433/aad.v2i3.139>
- Tavener, K., Crane, T. A., Bullock, R., & Galiè, A. (2022). Intersectionality in gender and agriculture: Toward an applied research design. *Gender, Technology and Development*, 26(3), 385–403. <https://doi.org/10.1080/09718524.2022.2140383>
- AI Statement:** Generative AI was used as an assistive tool for editing and proofreading. We, as authors, are responsible for all scholarship and content.