

A NEW DIMENSION OF EVIDENCE SYNTHESIS

DEVELOPING AND PILOTING THE GENDER AND SOCIAL CONTEXT CRITICAL APPRAISAL TOOL (GSC-CAT)

Carmen N. Benson, PhD¹ | Theresa Pesi Murphrey, PhD²
¹ Cooking with Kids, Inc. ² Texas A&M University

PURPOSE

 To develop and pilot the Gender and Social Context Critical Appraisal Tool (GSC-CAT) to systematically assess how primary studies account for gender and social context.

More inclusive evidence for stronger, more equitable solutions.

BACKGROUND

Agricultural education initiatives contribute to livelihoods, workforce readiness, agricultural productivity, food security, and youth development in low- and middle-income countries. Yet access, participation, and benefits are not experienced equally across gender, socioeconomic status, age, disability, ethnicity, and geography.

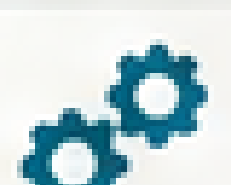
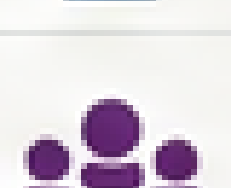
Conventional critical appraisal tools focus on methodological rigor, with limited attention to these differences—which can obscure who participates, who benefits, who is excluded, and for whom outcomes differ.

Methodologically rigorous evidence may still obscure who participates, who benefits, who is excluded, and for whom outcomes differ.

— Drucza & Abebe, 2018

THE GSC-CAT TOOL

Nine criteria examine how studies account for gender and social context across the research process. Studies are classified along a continuum from GSC-Unaware to GSC-Transformative.

#	CRITERION	KEY QUESTION
1	 Representation (Results)	Were results disaggregated by gender?
2	 Representation (Results)	Were results disaggregated by other socioeconomic characteristics?
3	 Sampling	Were sampling methods sensitive to representing underrepresented groups?
4	 Sampling	Were sampling methods sensitive to representing other marginalized populations?
5	 Objectives	Did the study objectives include an explicit GESI focus?
6	 Analysis	Did the evaluation examine intersectional dimensions of advantage and disadvantage?
7	 Analysis	Did the evaluation assess the effects of the intervention on gendered power relations?
8	 Analysis	Did the evaluation assess the effects of the intervention on other power relations affecting marginalized populations?
9	 Participation	Did the evaluation methods allow participants affected by the intervention to contribute to evaluation planning, interpretation, or decision-making?



METHODS

 **Tool Development**
Iterative review of literature, evaluation frameworks, and practitioner guidance (Drucza & Abebe, 2018; Patton, 2022; Tavenner et al., 2022; UN Working Group, 2017).

 **Instrument Design**
Nine criteria with definitions, decision rules, and scoring guidance (Boateng et al., 2018; DeVellis & Thorpe, 2021).

 **Pilot Testing**
Applied to 79 youth agricultural education studies in low- and middle-income countries (Benson et al., 2024).

 **Review Process**
Two independent reviewers appraised each study, recorded time required, compared ratings, and refined the tool.

KEY FINDINGS

**79**
studies appraised

**8–11**
additional minutes per study

**96%**
inter-rater agreement
Across 711 paired ratings (79 studies × 9 criteria).

KEY GAPS IN THE EVIDENCE BASE

48%
Did not meet the minimum expectation of reporting gender-disaggregated findings (n = 38).

32%
Were classified as GSC-Unaware.

31%
Lacked gender-disaggregated findings.

 Although many studies demonstrated sound methodological rigor, few meaningfully considered how gender and social context shaped participation, experiences, and outcomes.

IMPLICATIONS: A MORE COMPLETE AND EQUITABLE EVIDENCE BASE

**FOR RESEARCHERS**
Use the GSC-CAT to make social inclusion in evidence synthesis more visible.

**FOR PRACTITIONERS AND EVALUATORS**
Provides practical, rapid guidance for designing and interpreting more inclusive, equitable, and participatory evaluations.

**FOR THE FIELD**
Broader application and validation can strengthen the evidence base and support more inclusive policies, programs, and investments.

NEXT STEPS

- ✓ Conduct additional inter-rater reliability testing and validity analyses.
- ✓ Explore application across additional sectors and geographical contexts.
- ✓ Integrate into evidence synthesis protocols.
- ✓ Refine and scale the tool for broader use.