

Evaluation Through a Gender Lens: Foregrounding Global South Reflections and Experiences

Sonal Zaveri
International Evaluation Academy

Sibongile Sithole
International Evaluation Academy

Silvia Salinas Mulder
Research and Evaluation Consultant

Amrita Gupta
Research and Evaluation Consultant

Anuradha Chatterji
Creating Resources for Empowerment in Action

Shyam Singh
Tribhuvan Sahkari University

Background: Gender-focused evaluation in the Global South is frequently shaped by colonial and Western epistemologies that can reinforce knowledge hierarchies and reproduce structural inequalities. Mainstream evaluation approaches, largely developed in the Global North, may inadequately account for forms of oppression rooted in colonial histories and locally specific structures of power. When applied in Global South contexts without sufficient critical adaptation, they can overlook the intersecting effects of structural violence, internal colonialism, institutionalized and relational patriarchy, racialized hierarchies, and other context-specific inequalities, thereby reproducing the very power imbalances that evaluation seeks to address.

Purpose: To propose a feminist Global South (GS) decolonized framework for gender-transformative evaluation that centers GS feminist epistemologies, and to illustrate how this framework will reshape evaluation design, evidence generation, interpretation, and use across diverse contexts.

Setting: Global South, with case studies drawn from Latin America, South Asia (India), and the Middle East.

Intervention: Four illustrative case studies: (1) interrogation of power blindness in Latin American evaluations using the power blindness wheel; (2) CREA's cocreation of feminist monitoring, evaluation, and learning systems with structurally excluded groups across South Asia and the Middle East;

(3) gender-transformative evaluation of masculinity norms with adolescent boys in urban India; and (4) social network analysis of gendered norm enforcement in rural India.

Research Design: Thematic analysis drawing on a secondary literature review of GS feminist scholarship, synthesis of Three Horizons foresight evidence, framework development, and cross-case analysis of the four evaluations.

Data Collection and Analysis: Not applicable. This is a conceptual and comparative article drawing on existing evaluations and scholarship rather than primary data collection.

Findings: The article proposes a framework organized around five interlocking domains (power dynamics, intersectionality, inclusion and representation, methodological approaches, and ethics) that function as a scaffold for context-specific methodological choices across the evaluation cycle. The case analysis reveals that gender-transformative evaluation requires naming and interrogating underlying power relations within evaluation itself; explicitly intersectional sampling and analysis; co-ownership of meaning-making with affected communities; context-specific feminist methodologies; and relational, iterative ethics embedded throughout the process. The analysis underscores the need for structural transformation of Global North-dominated evaluation mindsets and systems to create genuine space for Global South feminist and decolonial epistemologies and evaluators.

Keywords: *Global South; gender equality; gender-transformative evaluation; feminism; decolonization; intersectionality; power dynamics; evaluation ethics*

Social power is the capacity of different individuals or groups to determine who gets what, who does what, who decides what, and who sets the agenda.

—Srilatha Batliwala, *All about Power: Understanding Social Power and Power Structures*

Gender inequality persists despite five decades of global commitment. The commitment to gender equality was formally elevated in June 1975 at the first World Conference of the International Women's Year in Mexico and later reaffirmed through the Beijing Declaration and Platform for Action (United Nations, 1996). Yet neither a gender focus within the Millennium Development Goals (MDGs) nor Sustainable Development Goal (SDG) 5 “to achieve gender equality and empower all women and girls” has delivered significant and sustained progress (Commission on the Status of Women, 2014; Equal Measures 2030, 2024a; UN Women & DESA, 2025).

Between 2019 and 2022, progress stalled sharply, with nearly 40% of countries stagnating or regressing on gender equality and the lowest scores concentrated in the Global South (UN Women & DESA, 2025; Equal Measures 2030, 2024b). Current projections suggest gender equality in parliaments will not be achieved until 2063 (UN Women et al., 2024). Other estimates are even less optimistic: World Bank analysts have projected that, at current rates of progress, full legal gender equality may not be achieved for another 134 years (around 2158; World Bank, 2025). These trends warrant a critical examination not only of gender strategies but also of the evaluation frameworks and approaches used to define, measure, and validate “progress,” particularly using gender-transformative evaluations.

The Problem with Dominant Frameworks

Mainstream evaluation approaches have largely been developed and institutionalized within Global North (GN) contexts. Applied in the Global South (GS), these frameworks often struggle to capture intersecting forms of oppression shaped by colonial legacies, structural violence, internal colonialism, institutionalized and relational patriarchy, and racialized hierarchies. In practice, evaluations may reference “cultural context” or “local perspectives” while still imposing external analytical categories, indicators, and theories of change. The result is a

persistent paradox: Efforts intended to advance gender equality can inadvertently reproduce the power imbalances they seek to address. This underscores the need for perspectives for evaluation that better align with contextual understandings of gender dynamics and inequalities.

Clarifying Global South and Key Evaluation Distinctions

It is important to note that “Global South” is not a purely geographical designation. Rather, it describes structures of oppression, marginalization, and colonial legacies that shape societies and knowledge systems. The GS is grounded in historical and ongoing colonial relationships, unequal global economic structures, and epistemic hierarchies that privilege GN knowledge systems. This article differentiates feminist gender-transformative evaluation as being distinct from gender evaluation (which may be descriptive) and gender-responsive evaluation (which adapts to gender issues). Gender-transformative evaluation explicitly seeks to challenge and shift unequal power relations, norms, and structures that perpetuate gender inequality, working toward substantive justice and equity. We do so with an understanding of GS realities.

This article focuses on gender-transformative evaluation, building on feminist epistemologies and methodologies from the GS. This distinction is important because it clarifies that we are not offering a generic “gender lens” but rather a transformative, decolonial, justice-oriented approach to evaluation.

Purpose and Methodology

This article argues for centering Global South feminist epistemologies in gender evaluation especially for international development programs and projects, many situated in the Global South. We contend that meaningful progress toward gender equality requires rethinking how gender *inequality* is understood, analyzed, and evaluated across diverse contexts.

The article adopts a comparative, thematic approach that foregrounds epistemic, methodological, and political distinctions between GS and GN perspectives, while recognizing shared commitments to gender justice and divergences shaped by history, geography, contextual dynamics, and positionality. It examines areas where GS

scholarship has challenged, expanded, or reframed dominant GN frameworks, especially in relation to power, intersectionality, inclusivity and representation, methodology, and ethics, and translates these into practical insights for commissioners and evaluators seeking to decolonize gender in evaluation practice.

Methodology

To analyze how evaluations could benefit from a GS gender lens, the article draws on four linked components:

- Theoretical foundations: a secondary literature review of key contributions of GS feminist scholarship, distinguished from dominant GN approaches, to show how GS perspectives reshape evaluation theory and practice from a gender-transformative lens.
- Three Horizons (3H) evidence base: findings from a recent foresight approach, the Three Horizons (3H) study (Ofir et al., this issue), which examines dimensions of the transformational imperative to address the polycrisis and ensure social equality and planetary health. Subsequent analysis mapped the shifts in evaluation practices across horizons from H1 (patriarchal, colonial, and positivist systems embedding power inequalities) through H2 (the need for rights-based, context-specific, gender-responsive frameworks) to H3 (recentering power and equality through decolonized approaches).
- Framework development: development of a GS framework based on the GS literature review and cross-referencing insights from the 3H study.
- Case studies and synthesis: four case studies illustrating components of the GS approach for gender-related evaluations, followed by a comparative analysis to consolidate learning and propose a way forward. Throughout, we emphasize that this article does not argue for replacing one hegemonic framework with another; rather, it argues for creating space for multiple, context-grounded approaches to flourish.

Use of Artificial Intelligence (AI)

AI tools were integrated at various stages of developing the article. Claude and ChatGPT supported brainstorming of ideas and concepts, particularly those related to gender emerging from

the 3H Maps and assisted with formatting selected sections. AI was used as a complementary tool, supporting but not replacing the human creativity and critical reflection that underpin feminist and gender-transformative insights from the Global South.

Methodological Limitations

This article is conceptual and illustrative rather than empirical, and its findings should be read with several methodological limitations in mind. First, the literature review was purposive rather than systematic: It foregrounds GS feminist scholarship that the authors, working primarily in Latin America and South Asia, are most familiar with. Second, each case study is shaped by the positionality, training, and institutional location of the evaluators involved, including the authors. Third, the article relies on secondary materials, published evaluation reports, and the authors' reflective accounts. Fourth, the comparative discussion between the cases is thematic rather than strictly comparative in a methodological sense: Cases are read against the five domains of the framework. Finally, framing through a GS / GN binary, while analytically useful, risks flattening internal diversity within and across these geographies, including the experiences of Indigenous, queer, disabled, and migrant feminists whose epistemologies do not map neatly onto either category. We encourage readers to treat the framework as an invitation to dialogue and adaptation rather than as a finished tool, and to extend it through future work grounded within additional regions, sectors, and evaluator standpoints.

Theoretical Foundations: Global South Perspectives of Gendered Evaluations

Academic knowledge production has historically been concentrated in the GN, reproducing a pyramidal model (Connell, 2015) in which theory and legitimacy “trickle down” to the GS, limiting the availability and uptake of Southern feminist perspectives (Das, 2017; Batliwala, 2011). GS feminist scholarship is frequently positioned as peripheral, reinforcing epistemic injustice that mirrors colonial legacies of power and exclusion (Macheke & Michaeli, 2023). This unidirectional flow sustained by Northern institutional dominance shapes research priorities in ways that reflect Northern interests and marginalize GS

voices (Huq, 2025; Macheke & Michaeli, 2023; GWL Voices, 2023).

For gendered evaluations, these epistemic hierarchies are not abstract; they shape the framing of evaluation questions, the categories used to classify people and change, the indicators treated as credible evidence, and the theories of change used to explain progress. GS feminists, working within histories of colonialism and patriarchal social orders, have therefore critiqued dominant feminist frameworks shaped primarily in the GN that portray women in the GS as passive victims rather than active agents in their own narratives (Das, 2017; Ernstberger, 2020; Brewster et al., 2019; Mohanty et al., 1991). They have also advanced the “how” of centering voices historically overlooked by dominant paradigms of knowledge production (Ernstberger, 2020; Parekh & Wilcox, 2020), moving beyond the simplistic explanations in some Western feminist accounts reproducing paternalism or idealization of the GS or minimizing the maturity of GS feminists. GS feminists have worked on deeply foregrounding local practices within historical, economic, and social contexts shaped by intersecting oppressions across Asia, Latin America, and Africa (Mohanty, 1988; Narayan, 1998).

Additionally, GS feminists have emphasized the need to move beyond binary framings of gender and the cisgender focus of many dominant feminist approaches (Batliwala, 2012). Gender-transformative evaluation in the GS must also attend to queer, nonbinary, and transgender identities, recognizing that gender justice encompasses freedom from prescribed gender roles and norms for all people.

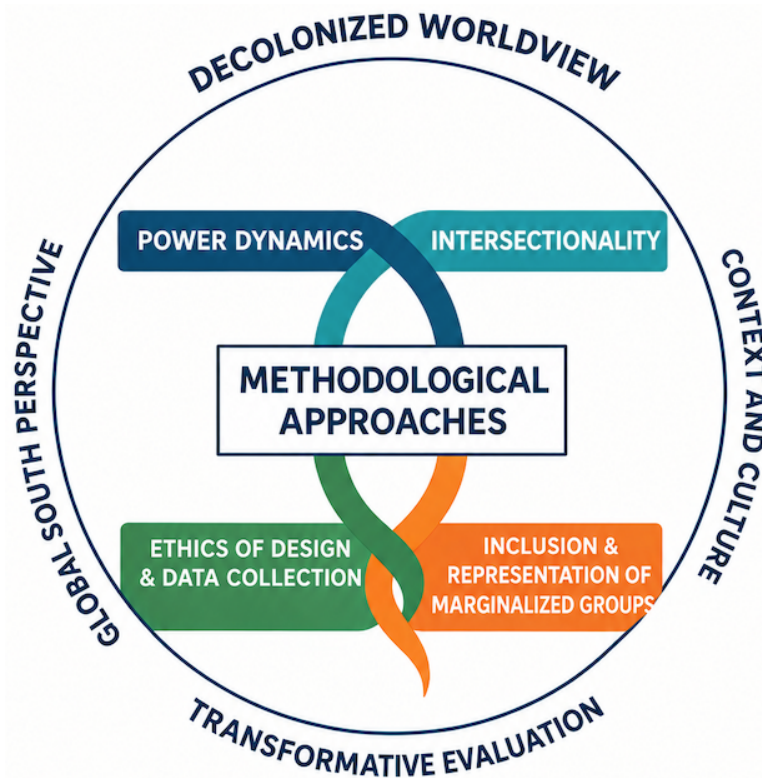
It is in this context that the GS decolonized framework introduced in the next section foregrounds five domains—power dynamics, intersectionality, inclusion and representation, methodological approaches, and ethics—as evaluation-relevant entry points for challenging epistemic dominance and strengthening gender-transformative evaluation practice.

A Global South Approach and Framework

GS feminist evaluators work within a dual terrain: local patriarchal structures and enduring colonial narratives that shape development practice, knowledge hierarchies, and what counts as “evidence.” A GS feminist approach is therefore not only conceptual; it is a practice framework for evaluation, guiding how evaluators frame problems, define criteria, generate evidence, interpret meaning, and negotiate use in ways that either reproduce or disrupt inequality.

This article proposes a GS decolonized framework for the praxis of gender-transformative evaluation organized around five epistemological domains: (1) power dynamics, (2) intersectionality, (3) inclusion and representation, (4) methodological approaches, and (5) ethics. These domains are not a fixed template; they function as a scaffold for context-specific methodological choices across the evaluation cycle: framing, design, data generation, analysis, validation, and use. Importantly, the five domains are deeply interrelated rather than siloed: Intersectionality, for example, is the foundation on which inclusion and representation rest. Co-ownership of meaning-making is meaningful only when those involved in cocreation are themselves selected through an intersectional lens that surfaces who is otherwise silenced by overlapping inequalities of gender, caste, race, ethnicity, class, sexuality, ability, age, language, and location. Likewise, power, methodology, and ethics cut across all the other domains: questions about who defines categories, who participates, whose knowledge counts, and whose safety is protected recur throughout sampling, analysis, governance, and use. The domains should therefore be read as overlapping entry points into the same underlying praxis, not as separate compartments (Figure 1).

Figure 1. Global South Decolonized Framework



Note. Created by authors for this article.

The sections that follow explain how each of these concepts is essential to transform entrenched inequalities and advance gender rights in evaluation theory and practice.

Power

Power is central to gender transformation and to evaluation itself. Foucault argues that power circulates through relationships, institutions, and everyday practices; it is productive in shaping norms and knowledge as much as it is repressive (Foucault, 1990). GS feminist scholarship emphasizes that feminist knowledge is a discursive and political practice inscribed with power relations (Mohanty, 2003b), and that dominant frameworks are often applied as if universally valid, with implications for whose realities are recognized and legitimized (Mama, 2007; Paredes & Silva Guimarães, 2025).

A recurring example is “empowerment.” Western feminist discourse has often conceptualized empowerment in terms of individual rights, autonomy, and choice (Cornwall & Rivas, 2015) and measured it through education,

labor participation, representation, or access to resources (Mohanty, 2003a). GS feminists argue that empowerment must be understood as the transformation of power relations, collective as well as individual, including organizing and challenging oppressive structures (Batliwala, 2010).

Evaluation Implications. GS feminist evaluators suggest power is both (a) an object of inquiry (what power relations are changing?) and (b) a feature of the evaluation (who controls design, implementation, interpretation, and dissemination?). This requires evaluation questions and criteria that examine who defines success, whose interests are served, how norms are enforced, and how institutional and relational power shifts over time. Beyond individual evaluator reflexivity, however, evaluation must also acknowledge that GN-dominated evaluation systems and structures themselves perpetuate power imbalances. Transforming gender evaluation therefore requires not only individual consciousness-raising but also structural change in how evaluation systems are organized, who

controls resources, and whose knowledge is valued (UN Women, 2026).

Intersectionality

Intersectionality is essential because “women” are not a homogeneous category. It explains how identity and structure interact to produce exclusion or privilege (UN Women, 2022). In much GN literature, intersectionality is frequently anchored in race, gender, and class (Crenshaw, 1989; Collins, 2000). In GS contexts, GS feminists emphasize intersecting inequalities such as gender, caste, class, and location that reinforce each other over time (Kabeer, 2010), and broaden the lens to include colonialism, imperialism, global capitalism, land dispossession, and cultural-religious systems that shape women’s lives (Lugones, 2007; Espinosa Miñoso et al., 2014; Mama, 2007; Mohanty, 2003a; Nnaemeka, 2004). GS scholarship also links intersectionality to transnational struggles, migration, displacement, and diasporic identities (Anthias, 2012; Piper, 2008) and highlights ecological injustice, including climate impacts mediated by unequal access to resources and decision-making spaces (Dankelman, 2010; Arora-Jonsson, 2011).

Evaluation Implications. Intersectionality must shape sampling and analysis: Evaluators should test how benefits and harms are distributed across intersecting groups; identify who is missing from evidence because of language, safety, gatekeeping, or social control; and interpret outcomes through structural conditions rather than individual change alone.

Inclusion and Representation

GS feminists challenge knowledge hierarchies that universalize categories and silence Indigenous histories and marginalized women’s realities (Mohanty, 2003b; Mama, 2007; Paredes & Silva Guimarães, 2025). This emphasis on situated knowledge and plural feminisms has direct evaluative relevance: Inclusion is not only participation; it is co-ownership of meaning-making.

Evaluation Implications. Gender-transformative feminist evaluation requires shared governance over evaluation purposes and questions, co-interpretation of findings, and validation processes where community sense-making is treated as evidence. The 3H synthesis for a transformative evaluative imperative in the context of the

polycrisis aligns with this perspective by calling for context-specific approaches that value Indigenous, GS, and diverse knowledge systems; co-ownership by affected communities; and amplifying the voices of vulnerable groups while positioning communities as decision-makers (Ofir et al., this issue).

Methodological Approaches

GS feminist scholars challenge metric-driven approaches and the hierarchical positioning of evaluator/researcher as neutral authority, which often fails to capture differential experiences of marginalized women and Indigenous perspectives of change (Kabeer, 1996; Murthy, 2001; Batliwala & Pittman, 2010). They critique tokenistic quantitative reporting and advocate participatory methodologies that enable listening, contextual exploration of gender and power, and reflexivity about evaluator bias and disciplinary grounding (Kabeer, 1996; Murthy, 2001; Batliwala & Pittman, 2010). While GN feminist qualitative traditions such as life histories and ethnography have long valued centering women’s voices (Reinharz, 1992; Collins, 2000), GS feminist scholarship also uses storytelling to surface collective histories, silenced knowledges, and structural violence linked to colonialism and global political economy (Mama, 2007; Mohanty, 2003a).

For gender-transformative evaluations, GS feminists and evaluators promote combining approaches that can capture power dynamics and contextual complexity, including outcome mapping, outcome harvesting, most significant change, and gender-analytical tools such as the social relations framework, Gender@Work, and social network analysis (Batliwala & Pittman, 2010; Schaeffer & Zaveri, 2023; Wilson-Grau & Núñez, 2007).

Evaluation Implications. Method choice should be justified by transformative intent: Methods must be able to detect shifts in agency, relationships, norms, and institutional practice, not only outputs and individual outcomes, and avoid mechanized one-size-fits-all criteria that ignore intersectionality and ethics (Ofir et al., this issue). Relatedly, systems-oriented work in the GN that challenges reductionist efficiency logics, such as ecorelational system science theories, signals convergence with feminist and Indigenous emphases, though these approaches remain marginal within dominant evaluation commissioning and mental models (Parsons & Winters, 2023).

Ethics of Design and Data Collection

Standardized ethics rooted in GN assumptions about autonomy, risk, and individual consent can become unethical when they ignore local norms, colonial histories, and gendered power relations. Communities may consent under pressure and later resist through evasion (Salinas Mulder et al., 2000). Ethics should therefore be contextual, negotiated, and iterative rather than a one-off signature exercise. Salinas Mulder and colleagues (2000) show how consent, vulnerability, risk, and participation are interpreted differently across cultural frames, requiring ongoing attention to voluntariness and power, and ethics protocols that are revisited throughout the evaluation.

Additionally, trauma-informed evaluation practice is essential to gender-transformative work, particularly in contexts marked by violence, displacement, and systemic oppression. Trauma-informed approaches recognize that many individuals, communities, and organizations have experienced collective trauma rooted in violence, colonialism, and structural inequality. Evaluation processes must be designed with awareness of potential retraumatization, with built-in safety mechanisms, flexibility around participation, and trained facilitators who can hold space for healing while gathering evidence (Kallos, 2025).

Evaluation Implications. Gender-transformative evaluation ethics should include revisiting consent over time, culturally appropriate participation (including collective decision-making where relevant), explicit attention to gendered risk and backlash, and transparency about who controls dissemination and use, particularly where communities face political risk (Salinas Mulder et al., 2000; Ofir et al., this issue). Ethics frameworks should also incorporate trauma-informed principles, recognize the cumulative impacts of violence and oppression, and build in flexibility, choice, and control for participants throughout the evaluation process. More broadly, evaluation reports should foster “ethical transparency” and accountability by encouraging evaluators to openly document the ethical dilemmas and limitations they encountered and the decisions they took, without fear of being “punished” for doing so; the pursuit of an image of “ethical perfection” risks

reproducing a Global North tendency to pretend not to see what actually happens in the field (Rance & Salinas Mulder, 2001).

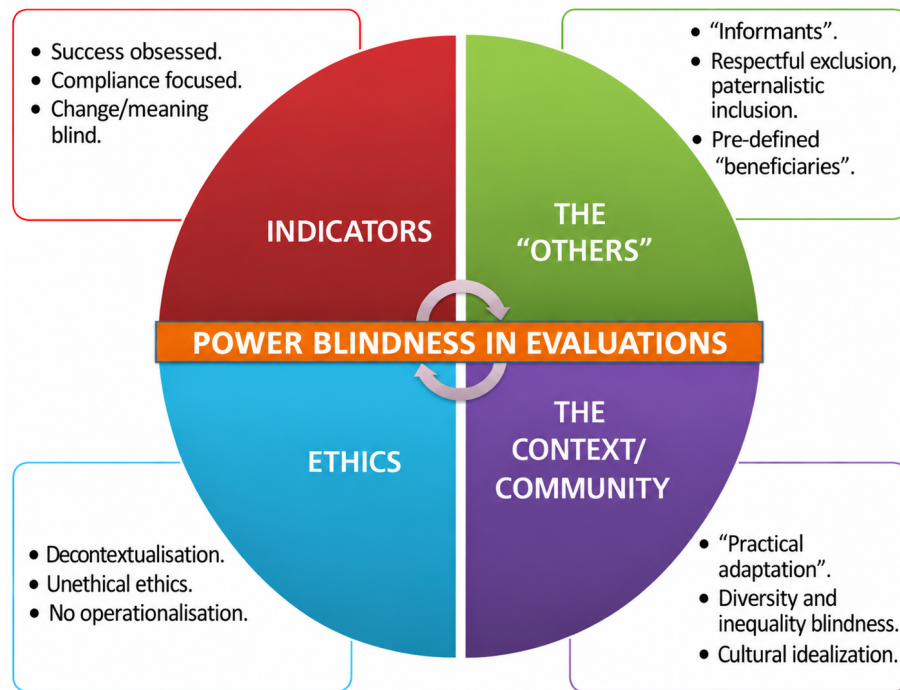
Case Studies

The four case studies further explore these five components that define and position the evaluation of gender in the Global South. The first case study explores evaluation from the perspective of power and the centrality of intersectional inequalities in Latin America impacting the transformation of gender inequalities. This case study is issue-based, drawing on multiple projects and distilling diverse evaluative experiences. The others are discrete case studies. The second case study by CREA discusses inclusion and representation, as well as the need to cocreate feminist methodologies with evaluators to ensure their use. The third case study describes the insights from a Global South–based evaluation in understanding how masculinity norms are shaped by power, inclusion, intersectionality, and the important role of ethics. The case study also discusses the need for use of participatory methodologies that interrogate deeply entrenched social and gender norms in different contexts. The fourth and last case study describes how social network–analysis methodology in an evaluation provides insights into complex interpersonal relationships in shaping gender norms. All case studies draw on evidence from the Global South.

Case Study 1: Latin America—Where Does Power “Hide” in Evaluations?

This case draws on Latin American realities, where plurinational and multilingual social fabrics coexist with entrenched, intersecting inequalities that disproportionately affect Indigenous and Afro-descendant peoples. It shows how internal colonialism and patriarchy, often subtle yet persistent, shape evaluative practice. Segato argues that internal colonialism intensifies and reorganizes pre-existing communal patriarchies into a “high-intensity” colonial/modern order in which racialized gender hierarchies, and gendered violence as a technique of governance, secure domination and territorial control (Segato, 2021).

Figure 2. Power Blindness Wheel



Note. From The Power Blindness Wheel [Conference presentation] by Salinas Mulder, November 2015, American Evaluation Association annual conference, Chicago, IL, United States.

Together, the four pillars in the wheel in Figure 2 operationalize this article's Global South framework by showing how power, intersectionality, inclusion/representation, methodological choice, and ethics are enacted through everyday evaluation routines.

In 2015, the question "Where does power hide in evaluations?" motivated author Salinas Mulder to present the power blindness wheel at the annual conference of the American Evaluation Association (AEA; Salinas Mulder, 2015). The wheel offers a feminist, decolonial approach to identify and address power blindness across four thematic areas—otherness, context, ethics, and success/indicators—each of which reveals how evaluation can reproduce inequality through apparently neutral or "good practice" routines. Power blindness can appear through the naturalization of social norms, paternalism, idealization, homogenization, generalization, and decontextualization. These expressions are not always overtly oppressive; they can be subtle and well-meant, yet still reinforce gendered and intersectional inequalities, exclusion, and discrimination in Global South contexts.

Otherness. Under time and resource constraints, evaluation practice can become technical and efficiency-driven, with evaluator-led, extractive engagement (Salinas Mulder & Amariles, 2016). Language is not neutral: People are framed as "informants," while those labeled "beneficiaries," often by default, before benefits are established, are positioned as passive recipients. Gender roles, rural–urban divides, language barriers, and Indigenous subordination shape who speaks and how. In rural Tarabuco, Bolivia, an evaluation on women's political participation showed how silence became women's "voice": As two white, urban, Spanish-speaking Bolivian women, the evaluators had to rely on a male intermediary to conduct and translate some interviews.

Context. Contextualization is frequently invoked as a buzzword, constrained by efficiency pressures rather than treated as critical, relational understanding. It can also appear as idealization of Indigenous and Afro-descendant realities framed as "cultural respect" (Salinas Mulder & Amariles, 2016). Yet culture cannot be "respected"

uncritically when the aim is to transform injustices rooted in social norms and beliefs. Power also works through simplified, uniform portrayals of rural, Indigenous, and Afro-descendant settings that ignore diversity and inequality. These assumptions shape who becomes a “key informant,” often reinforcing male- and adult-centered hierarchies. Broad categories such as “family” and “community,” sometimes defended as reflecting collective worldviews, can hide internal exclusions; they may also obscure contextual realities, such as the prominent child care role of grandmothers among vulnerable groups in Peru and Bolivia.

Ethics. Both universal and purely contextual ethics can hide power imbalances and reproduce harm. Universal templates can impose unrealistic requirements that increase risk. In Bolivia, illiterate women who could not read consent forms offered to “sign a blank sheet of paper” to document consent (Salinas Mulder et al., 2000). Conversely, a solely contextual frame can normalize place-based, gendered, and adult-centered bias. In Mexico, a team of women evaluators faced a dilemma: National law required parental consent for adolescent girls to join an evaluation on sexual and reproductive health and rights (SRHR) and gender-based violence (GBV). Despite legal compliance and in a setting with high sexual violence and adolescent pregnancy, where official data indicated most perpetrators were fathers, uncles, or other relatives, “parental consent” became ethically fraught. From a Global South feminist perspective, an ethically sound response would refuse to treat such dilemmas as a binary choice between “legal compliance” and “protection” of girls. It would instead reframe ethics as a collective, iterative, and power-aware practice. Concretely, this means: (a) negotiating an ethics protocol with local feminist organizations, women’s rights defenders, and adolescent girls themselves before fieldwork begins, so that risk, voluntariness, and confidentiality are interpreted from girls’ standpoint rather than from the perpetrator’s household; (b) seeking alternative consent pathways available under national child-protection or research ethics regulations (such as consent from a trusted guardian, school authority, or feminist NGO acting in loco parentis) where parental consent could itself constitute a risk; (c) embedding survivor-centered and trauma-informed safeguards including referral pathways, safe spaces, and the explicit right to withdraw without consequence; and (d) being transparent with commissioners and ethics committees about the dilemma, the reasoning behind the chosen path, and its

limitations, rather than performing “ethical perfection” (Salinas Mulder et al., 2000; Rance & Salinas Mulder, 2001). Ethical accountability, in this view, lies in surfacing and reasoning through dilemmas with affected communities, not in obeying procedural rules that may reproduce the very harms the evaluation seeks to address.

Indicators and Success. Many evaluations are implicitly oriented toward success and compliance, encouraging selective narratives. Indicators may be reported with limited space for reflection; critical findings can be vetted and stories shaped to demonstrate success even when reality is mixed. In Bolivia, a program expanding rural women’s access to finance received a very positive rating. While many women obtained credit, loan officers pressured women to take loans to meet targets, and some husbands forced women to sign so they could control the money, dynamics that remained outside the official story. Indicators and success narratives can therefore miss context-specific meanings and unintended effects that are essential for understanding the complexities of gender transformation (Salinas Mulder & Amariles, 2014, 2016).

Ultimately, power blindness is not a side issue in evaluation; it shapes who is heard, what counts as evidence, and how change is judged. Naming where power hides (in othering, context talk, ethics, and success narratives) enables evaluation designs that widen participation, surface contradictions, and treat learning as central. Only then can evaluations contribute to gender-transformative change rather than reproduce the inequalities they assess and aim to transform.

Case Study 2: Cocreating Feminist MEL Tools: CREA’s Grounded, Contextual, and Inclusive Approach

This case study examines how CREA reviewed, revisited, and redesigned conventional monitoring, evaluation, and learning (MEL) systems to capture diverse experiences of structurally excluded groups. It starts from a simple premise: If MEL systems cannot meaningfully capture gendered transformation, then tweaking indicators alone will not produce the social change sought.

Who Is CREA and Why Is Their Approach Specifically Global South? Creating Resources for Empowerment and Action (CREA) is a Global South-based organization grounded in South Asia and the Middle East, with over 15 years’ experience working with structurally excluded groups

including sex workers, LGBT groups (including LGBTQ+ and gender-diverse people), women with disabilities, and people marginalized by caste, class, race, and ethnicity. CREA emerged from and is accountable to social movements in the Global South, working within contexts shaped by colonial legacies, intersecting oppression, and shrinking civic space. Their approach is distinctly Global South because it is rooted in the lived realities, knowledge systems, and organizing traditions of marginalized communities in Asia and the Middle East. CREA positions MEL not as a donor compliance tool but as a site of liberation and accountability to movements for justice. Their work challenges the Northern-centered, standardized MEL systems that dominate development evaluation and insists on contextual, community-led alternatives.

CREA has sought to develop MEL approaches that are defined and owned by the people whose lives and struggles they document. This case study highlights how CREA aimed to reframe mainstream MEL, drawing on Global South feminist monitoring and evaluation practice.

Through CREA's Feminist Accountability Project (FAP), a collaboration with community MEL practitioners and partners across South Asia and the Middle East, MEL was reframed as liberatory, safe, relational, and accountable to communities. The process was intentionally cocreated and non-extractive: Interpreters were engaged early, safeguarding protocols were explained, and facilitators built in pauses and check-ins to support consent, reflection, and safety, ensuring horizontal dialogue.

FAP used three complementary approaches—participatory workshops, peer in-depth interviews, and culturally specific, context-grounded methods (including poetry and other art forms, alongside conflict resolution tools oriented to healing and justice)—to surface what “change” looks like in context. These approaches unpacked power, intersectionality, and culture and context, including well-being, to account for structural fatigue¹, emotional labor², and political risk. Such reflective practice supported diagnosis of structural barriers to accountability while enabling participatory and safer evaluative spaces. It also created a concrete basis for educating donors about

transformative MEL expectations, including what counts as evidence and whose interpretations matter.

Participants cocreated tools grounded in lived realities, using facilitated role-play skits and iterative design to develop surveys and processes guided by principles of equalizing power, collective care, flexibility, intersectionality, and community voice. For example, GBV-focused groups emphasized trauma-informed and survivor-centered approaches, including non-extractive data collection with marginalized women in the conflict-ridden and militarized Chittagong Hill Tracts,³ and trust-building with hijra (transgender⁴) communities where safety, language, and cultural familiarity are critical to ethical engagement. This included explicit attention to how gender identity, sexuality, and non-normative expressions are constrained by evaluation systems and how MEL can create space for authenticity and agency.

The MEL development process also enacted cultural transformation: It acknowledged emotional histories participants brought to work, created an affirming space for critical reflection, and modeled language justice, accessibility, and collective care. Participants visualized how power operates through social positions and privileges in MEL, from whose voices are heard, to who controls data, to how findings are interpreted. Tools were adapted to reflect multiple identities (gender, caste, language, age, sexuality) and constraints such as emotional labor and mobility restrictions, framing MEL not only as a technical practice but also as a site where intersectional oppression can be reproduced or resisted.

Using participatory workshops and in-depth interviews, participants cocreated the FAP tool kit, which addressed both political realities (criminalization of activism, shrinking civic spaces, intersectional exclusions) and practical constraints (inaccessible donor reporting systems, internal resource limits). Because organizations and community MEL practitioners were involved throughout, reported uptake was high, with participants committing to revise organizational tools to embed care, consent, and contextual relevance, and some organizations beginning to introduce feminist MEL to grassroots women and queer communities.

¹ The trauma and fatigue arising out of existing in and navigating structures that continually discriminate, silence voices, deny rights and freedoms, and perpetuate violence and exclusion.

² The effort and passion that goes into the affected people to engage in the work to amplify their voice and visibility.

³ <https://www.pcjss.org/cht-history-struggle/>

⁴ (in South Asia) a person who was registered as male at birth, but who identifies as female or as neither male nor female.

Through cocreation, intersectional design, and cultural grounding, this feminist MEL approach reframed MEL from a bureaucratic requirement into a site of feminist transformation: a political, relational, and ethical practice that must be continually cocreated. Rather than centering evaluators or donors, it aimed to redistribute power across all stages—co-designing tools with communities; building consent, accessibility, and safety into data collection; and shifting from surveillance to collective ownership of change. The case also demonstrates that feminist accountability is not incompatible with transparency or rigor in MEL systems; instead, it expands what rigor requires when evaluations seek to document gendered and intersectional transformation.

For commissioners and evaluators, this case shows that gender-transformative MEL requires changing not only indicators but also governance: who defines change, how evidence is generated, and who interprets findings. It demonstrates concrete methodological alternatives (peer interviews, arts-based reflection, covalidation) and ethical practices (iterative consent, safeguarding, language justice) that make evaluation both more credible and less extractive in contexts shaped by unequal power and political risk. Further, it highlights a shift from the role of “powerful” evaluator to that of facilitator, which has implications for the competencies of evaluators (see Chaplowe et al., this issue).

Case Study 3: ECF India: Evaluating the Transformation of Masculinity Norms in the Global South

Equal Community Foundation (ECF), based in Pune, India, has worked since 2009 with adolescent boys (14–17 years) in low-income urban communities to transform masculinity norms and prevent violence against women. This case study is particularly significant for Global South feminist evaluation because it demonstrates how gender-transformative approaches can be applied to work with boys and men, shifting the historical focus on women’s empowerment to include examination of how patriarchal masculinities are constructed and can be transformed. Its Action for Equality (AfE) program uses a structured curriculum delivered by mentors across foundation (15 weeks), action (15 weeks), and leadership (10 months) levels. Over

3,000 boys across 28 communities have participated.

An independent evaluation⁵ used a hybrid approach combining feminist evaluation⁶, utilization-focused evaluation (UFE),⁷ and developmental evaluation (DE)⁸, framed by a gender-transformative lens. This is significant because gender-transformative evaluation has rarely been applied to programs working with boys and men. The evaluation involved 106 stakeholders using mixed methods: attitude surveys; participatory tools (body maps, communication maps); focus group discussions (FGDs) with mothers, sisters and female relatives; and interviews. The evaluation was collaborative and iterative, designed to support ECF’s learning and organizational development while generating credible evidence.

Power. The evaluation revealed how power operates through constructions of masculinity. Boys’ concepts of hegemonic masculinity were being reshaped through the program. Deep-seated gender attitudes (dowry, son preference) were particularly resistant to change, illustrating how patriarchal power is embedded in social norms. Power also operated through household hierarchies—mothers and sisters validated (or resisted) boys’ changes, making household relationships crucial sites of transformation. The theory of change was collaboratively reconstructed using outcome-mapping concepts (expect to see, like to see, love to see), revealing that the program was primarily about constructing positive male identity during adolescence rather than directly preventing violence. This finding exemplifies how Global South feminist evaluation can challenge program assumptions and reframe what counts as success, moving beyond linear assumptions about violence prevention to recognizing identity formation and gender norm transformation as foundational.

Intersectionality. Gender discrimination intersects with caste, class, ethnicity, and religion. The evaluation disaggregated data by these variables and deliberately sampled from Muslim and caste-marginalized communities. In low-income communities where poverty was the primary concern, gender equality was not a priority—showing how socioeconomic class and

⁵ <https://ecf.org.in/wp-content/uploads/2018/03/ECF-Evaluation-of-Action-for-Equality-programme-Summary-Report-Final.pdf>

⁶ <https://www.betterevaluation.org/methods-approaches/themes/feminist-evaluation>

⁷ <https://www.betterevaluation.org/methods-approaches/approaches/utilisation-focused-evaluation>

⁸ <https://www.betterevaluation.org/methods-approaches/approaches/developmental-evaluation>

gender intersect. The evaluation recognized that boys' capacity to take on more equitable gender roles depended not only on their individual awareness but also on their intersecting social positions and material constraints. Some boys remained resistant to participation, and the evaluation documented how caste status, class position, and religious identity shaped both access to the program and receptiveness to its messages.

Inclusion and Representation. Structurally excluded groups (related to caste, ethnicity, and religion, as well as class) were intentionally included. Mothers and sisters were consulted to validate boys' claimed changes—centering women's perspectives in an evaluation of a men's program. Boys participated through adolescent-friendly participatory learning and action (PLA) tools (body maps, communication maps). The UFE approach ensured ECF staff learned in real time and owned the evaluation. This inclusion strategy was crucial because it recognized that gender transformation cannot be evaluated through boys' self-reports alone; it requires validation from those most affected by changing masculinities.

Methodological Approaches. The hybrid approach (UFE + DE + feminist evaluation) was innovative for evaluating masculinity programs. Body maps enabled boys to express embodied changes in knowledge, attitude, and behavior. Communication maps tracked norm transmission within families and communities. Outcome-mapping concepts (spheres of influence/control) showed boys' influence was primarily at the household level. The collaborative theory of change process exposed the gap between stated goals and actual practice. An example is the variation between intent and practice: 50% of boys with 60%+ attendance showed positive attitude change; on specific issues like dowry, 90% demonstrated equitable attitudes. "Unintended outcomes" (boys' improved education, interfamily relationships) were significant changes emerging from critical thinking. Age and session attendance were important factors affecting change, with younger boys more likely to sustain participation and older adolescents sometimes more resistant to messages about gender equity. Further, the project employed male facilitators (influenced by similar patriarchal microcultures) and the feminist lens interrogated how changes in facilitators' awareness and the gender norm change were a factor in influencing those in participant boys and young men. Such a reflexive approach enabled discussion regarding

how the intervention stakeholders influenced (or not) norm change.

Ethics. Ethical consent from guardians was established at the outset; boys' assent was also included, and boys could withdraw at any time. The evaluation navigated sensitive topics around masculinity, violence, and family dynamics with care and cultural awareness. The collaborative UFE approach raised important questions about sustainability and whether deep transformation can be measured in the short term. The evaluator had to balance rigor with empathy and recognize iatrogenic risks—the possibility that engaging boys in gender awareness could create tension in households or expose them to ridicule from peers, making ongoing support and safety essential. Trauma-informed principles were embedded in how facilitators worked with boys who disclosed experiences of violence or abuse.

This evaluation demonstrates how GS feminist evaluation approaches—applied to programs working with boys and men—can surface hidden power dynamics, challenge program assumptions, and produce actionable evidence for transformative change. The reconstruction of the theory of change is a concrete example of how evaluation can move beyond measuring compliance to questioning the underlying logic of gender transformation. It also shows that gender-transformative evaluation is not limited to work with women and girls; it can illuminate how dominant masculine norms are constructed and offer pathways for reimagining masculinity aligned with gender justice.

Case Study 4: Social Network Analysis and Social Norms in India

Social norms shape and regulate gender roles through "reference networks," the social ties through which normative expectations are communicated, enforced, and validated. Individuals often conform to these expectations to secure acceptance and avoid exclusion, making norms resilient to change (De et al., 2021). This case study draws on social network analysis (SNA) to examine how sanitation-related social norms circulate in a village in Central Gujarat, India, including decisions to construct a household toilet (taboo in some contexts), use of toilet facilities, and women's participation in sanitation-related decision-making.

The network analysis suggests that gendered influence operates differently for men and women. Men often function as bridges in interfamily and intercommunity transmission of norms, sharing

expectations largely among male counterparts. Women, by contrast, tend to share norms locally, within family, close relatives, and neighbors. This local circulation is consequential: When women endorse restrictive norms within women's

networks, it can reduce the space for dissent and weaken possibilities for resistance, reinforcing norm stability at the household and neighborhood level.

Figure 3. Three Examples of Reference Networks of Sanitation-Related Social Norms

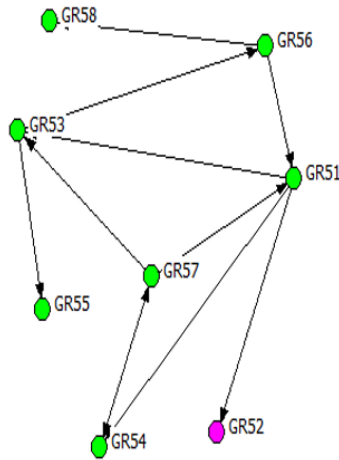


Figure 3.1. Reference Network of Sanitation-Related Social Norms

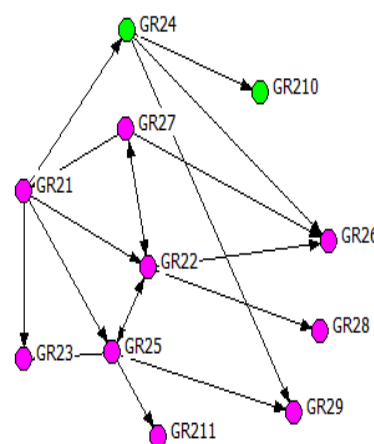


Figure 3.2. Reference Network of Sanitation-Related Social Norms

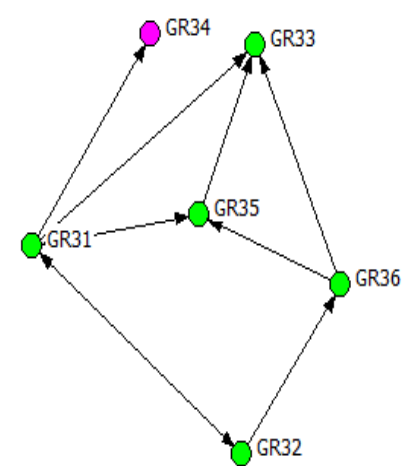


Figure 3.3. Reference Network of Sanitation-Related Social Norms

Note. Green indicates men and purple indicates women. The label of each node is a respondent ID. Arrows indicate who shares or practices social norms with whom.

A notable pattern concerns age and household hierarchy. Older women appeared more likely to transmit and enforce restrictive norms than younger women, who more often resisted them. This “role reversal” suggests that as women gain seniority within household hierarchies (through age and rank), they may also acquire incentives and authority to reproduce norms that they previously experienced as constraining. In this case, norms governing gendered roles inside the family appear more rigid and deeply rooted than norms shaping women's behavior outside the household.

Implications for Evaluation Lenses and Frameworks. These findings underscore the need for evaluation approaches that can analyze how power and norms circulate not only between men and women but also among women through household hierarchy and intra-gender enforcement. Intersectionality remains essential for understanding how multiple identities (e.g., caste/jati, class, religion, age, marital status) shape women's agency and constraint (Crenshaw, 1989). Likewise, empowerment frameworks that examine resources, agency, and achievements (Kabeer,

1999) and power relations (power within, power to, power with, power over; VeneKlasen & Miller, 2002) are strengthened when paired with methods that can empirically map who enforces norms, through which ties, and with what consequences. In this case, social network analysis makes visible a dimension that is easily missed in conventional evaluation: the relational mechanisms through which norms are reproduced within women's networks and household structures, including role reversal over the life course.

Contrary to expectations that reference networks in a cohesive society would be compact and highly interconnected, the village networks were relatively fragmented. Men's networks were more numerous but more disconnected, while women's networks were fewer but denser. This pattern suggests that women may play a limited role in broadcasting norms beyond close family groups, but a significant role in locally consolidating norms. It also cautions against assuming that macro-level interventions (e.g., women's economic participation) will automatically disrupt intra-household norm enforcement; evaluations may need a more

composite approach that examines both women's public roles and within-household dynamics.

Evaluation Takeaways. Mapping reference networks can strengthen gender-focused evaluation by improving contextualization of local power structures and hierarchies, clarifying who influences norm stability and change, and refining theories of change and indicators beyond individual behavior to relational mechanisms. Given the sensitivity of network data, ethical practice also requires attention to confidentiality and risk, particularly where mapping social ties could expose individuals to backlash or stigma, so that evidence generation does not intensify harm.

This case operationalizes the framework's methodological and power domains (SNA + power brokers), extends intersectionality through intra-

gender hierarchy, and flags ethics risks in evidence generation.

Cross-Case Analysis of Case Studies

The comparative analysis of the four cases (one issue-based and three discrete cases) highlights critical domains for advancing gender-transformative evaluation practice. Table 1 presents a comparative synthesis of the four case studies through the six domains of the Global South decolonized framework. It highlights how each case operationalizes gender-transformative evaluation practice through distinct mechanisms while reinforcing shared lessons across domains.

Table 1. Comparative Analysis of Cases and Their Implications for Gender-Transformative Evaluation

Analytical frames (Global South framework)	Case study 1: Latin America—Where does power “hide” in evaluations?	Case study 2: CREA—Feminist MEL	Case study 3: ECF India—Evaluating masculinity norms	Case study 4: India—Social network analysis and social norms
Power dynamics	Makes “power blindness” visible inside evaluation routines—othering language, thin contextualization, procedural ethics, and success-driven narratives—showing how evaluation can reproduce inequality unless power is explicitly analyzed.	Redistributes power by treating communities as definers of change and cocreators of MEL; positions MEL as relational and political rather than donor-driven compliance.	Reveals how power operates through constructions of hegemonic masculinity and household hierarchies; mothers and sisters both validated and resisted boys' changes, making intra-family relationships crucial sites of transformation. The collaborative reconstruction of the theory of change—using outcome-mapping concepts—challenges the program's own assumptions, reframing success from violence prevention to identity formation and norm transformation.	Shows how power circulates through reference networks and household hierarchy, including women's role in local enforcement and “role reversal” across age and rank.
Intersectionality	Situates evaluation within racialized gender hierarchies affecting Indigenous and	Centers structurally excluded groups (e.g., sex workers, LGBT groups,	Disaggregates data by caste, class, ethnicity, and religion and deliberately samples from Muslim and caste-marginalized	Highlights differentiated experiences shaped by gender, age, caste, and

Analytical frames (Global South framework)	Case study 1: Latin America—Where does power “hide” in evaluations?	Case study 2: CREA—Feminist MEL	Case study 3: ECF India—Evaluating masculinity norms	Case study 4: India—Social network analysis and social norms
	Afro-descendant peoples, demonstrating how intersecting exclusions shape voice, evidence, and interpretation.	women with disabilities; caste / class / race / ethnicity), treating change as nonlinear and shaped by political risk and lived realities.	communities. Documents how in low-income settings poverty overshadows gender equality as a priority, showing that boys’ capacity for equitable gender roles depends on intersecting social positions and material constraints and not individual awareness alone.	social status; uses network structure to show uneven constraint and agency across groups.
Inclusion and representation	Shows how “key informant” practices, translation dynamics, and mechanized engagement can silence marginalized voices and reproduce exclusion, even when framed as efficiency or cultural respect.	Builds inclusion through cocreation, language justice, accessibility, safeguarding, and non-extractive facilitation so that marginalized voices shape tools and interpretations.	Centers women’s perspectives in an evaluation of a men’s program by consulting mothers and sisters to validate boys’ claimed behavioral changes. Uses adolescent-friendly participatory tools (body maps, communication maps) to enable boys’ meaningful participation. The utilization-focused evaluation approach ensures ECF staff learn in real time and own the evaluation process, recognizing that gender transformation cannot be assessed through boys’ self-reports alone.	Demonstrates that representation depends on whose ties and voices are captured in network mapping; warns that local norm reinforcement can be missed if women’s intra-community dynamics are not surfaced.
Methodological approaches	Uses the power blindness wheel as a diagnostic to redesign methods that surface silences, contradictions, and unintended effects beyond standard Q&A and indicator reporting.	Employs participatory workshops, peer interviews, storytelling, arts-based methods, and trauma-informed approaches to cocreate MEL systems that capture complex change pathways.	Employs an innovative hybrid of utilization-focused evaluation, developmental evaluation, and feminist evaluation with a gender-transformative lens, rarely applied to programs working with boys and men. Body maps capture embodied changes in knowledge, attitude, and behavior; communication maps track norm	Uses social network analysis to explain how norms circulate, who enforces them, and where resistance is constrained, adding explanatory power to evaluations of social-norms interventions.

Analytical frames (Global South framework)	Case study 1: Latin America—Where does power “hide” in evaluations?	Case study 2: CREA—Feminist MEL	Case study 3: ECF India—Evaluating masculinity norms	Case study 4: India—Social network analysis and social norms
			transmission within families and communities. Outcome-mapping concepts (spheres of influence/control) reveal that boys’ influence operates primarily at the household level. The collaborative theory of change process exposes gaps between stated program goals and actual practice, while a reflexive feminist lens interrogates how male facilitators’ own patriarchal socialization shaped the intervention.	
Ethics	Shows that both universal and purely contextual ethics can reproduce harm—through unrealistic procedural demands or by normalizing gendered/adult-centered bias—and calls for iterative, power-aware ethics.	Embeds ethics as collective care, consent, safety, and non-extractiveness; foregrounds political risk, emotional labor, and safeguarding in evaluation design and facilitation.	Secures ethical consent from guardians while allowing boys to withdraw at any time. Navigates sensitive topics around masculinity, violence, and family dynamics with cultural awareness and care. Recognizes iatrogenic risks—engaging boys in gender awareness could create household tension or expose them to peer ridicule—making ongoing support and safety essential. Embeds trauma-informed principles in facilitator practice with boys who disclose experiences of violence. The collaborative approach raises important questions about whether deep gender transformation can be adequately measured within standard short-term evaluation timeframes.	Frames ethics around dignity, fairness, and women’s agency; points to the need for attention to risk and confidentiality when mapping sensitive social ties and norm enforcement.

Analytical frames (Global South framework)	Case study 1: Latin America—Where does power “hide” in evaluations?	Case study 2: CREA—Feminist MEL	Case study 3: ECF India—Evaluating masculinity norms	Case study 4: India—Social network analysis and social norms
Evaluator competencies for GS gender-transformative practice	Requires capacity to name and interrogate power blindness and redesign evaluation processes to widen participation, validate dissent, and resist “success-only” reporting.	Requires facilitation, reflexivity, and skills in cocreation, language justice, accessibility, and trauma-informed practice to support community ownership of MEL.	Requires capacity to apply hybrid evaluation approaches (UFE, DE, feminist evaluation) to masculinity programs; skill in facilitating participatory and adolescent-friendly methods; reflexivity to interrogate facilitator positionality and patriarchal socialization; and the ability to balance feminist accountability with empathy, cultural sensitivity, and trauma-informed practice when working with boys and men.	Requires systems and network literacy to interpret norm transmission and power relations, supporting deeper contextualization of local hierarchies and change mechanisms.

Lessons from the Global South—Advancing Gender-Transformative Practice

Power Dynamics

Across all four cases, addressing structural inequities is a defining feature of gender-transformative evaluation. Case Study 1 shows how “power blindness” is produced inside evaluation routines: othering language, thin contextualization, procedural ethics, and success-driven indicator narratives. Case Study 2 shows how cocreation, collective care, and peer learning redistribute power by enabling structurally excluded groups and feminist civil society to define what counts as change and credible evidence. Case Study 3 shows how power operates through constructions of hegemonic masculinity and household hierarchies; mothers and sisters both validated and resisted boys’ changes, making intra-family relationships crucial sites of transformation. The collaborative reconstruction of the theory of change challenged the program’s own assumptions, reframing success from violence prevention to identity formation and norm transformation. Case Study 4 extends the analysis by showing how power also operates through intra-household hierarchy, where older women may enforce restrictive norms on younger

women, making age, alongside gender, a site of privilege. Together, the cases suggest that evaluation must move beyond documenting intersectional inequalities to naming, interrogating, and redesigning the power relations that shape participation, evidence, interpretation, and use. However, it is crucial to recognize that individual evaluator reflexivity, while necessary, is insufficient. The cases point to the need for structural transformation of Global North-dominated evaluation systems themselves, including funding mechanisms, evaluation standards, and gatekeeping of what counts as “credible evidence” to create genuine space for Global South epistemologies and evaluators.

Intersectionality

Intersectionality functions as a foundational lens across cases, demonstrating that gender transformation unfolds through layered hierarchies rather than a single identity axis. Case Study 1 situates evaluation within ethnic, cultural, social, and gendered hierarchies that shape who is heard and what counts as evidence. Case Study 2 highlights that change is nonlinear and must capture experiences often hidden within intersecting identities and political realities. Case Study 3 demonstrates how gender discrimination intersects with caste, class, ethnicity, and religion by disaggregating data across these variables and

deliberately sampling from Muslim and caste-marginalized communities; in low-income settings, poverty overshadowed gender equality as a priority, revealing that boys' capacity for equitable gender roles depended on their intersecting social positions and material constraints. Case Study 4 illustrates how gender intersects with age, caste, and social status, producing differentiated experiences of constraint, agency, and norm enforcement. All the cases reinforce that evaluation questions, sampling, and analysis must be explicitly intersectional to avoid reproducing homogenized narratives of "women's (or men's) change."

Inclusion and Representation

A cross-cutting lesson is that inclusion is not an "add-on" but a mechanism of transformation that determines whose knowledge becomes evidence. Case Study 1 shows how "key informant" selection, translation dynamics, and mechanized engagement can marginalize those most affected. Case Study 2 demonstrates inclusion through cocreated processes that foreground language justice, accessibility, safeguarding, and non-extractive facilitation. Case Study 3 illustrates how centering women's perspectives by consulting mothers and sisters to validate boys' claimed behavioral changes strengthens evaluation credibility, while adolescent-friendly participatory tools (body maps, communication maps) and a utilization-focused evaluation approach ensure meaningful participation from both the boys and ECF staff. Case Study 4 suggests that representation also depends on whose social ties are visible in network mapping, underscoring that inclusion must be designed deliberately into evidence generation and sense-making.

Methodological Approaches

The cases point to the need for context-specific feminist evaluation capable of capturing norm enforcement, contradiction, and emergent change. Case Study 1 emphasizes adapting methods to surface silences, exclusions, and hidden power dynamics that conventional approaches may miss. Case Study 2 shows how participatory workshops, peer interviews, storytelling, and culturally grounded methods can be used not only to collect data but to cocreate MEL systems, including trauma-informed and survivor-centered practice. Case Study 3 demonstrates how a hybrid approach combining utilization-focused, developmental, and feminist evaluation with a gender-transformative lens can surface hidden dynamics in masculinity

programs; body maps captured embodied changes, communication maps tracked norm transmission through families, and a reflexive feminist lens interrogated how male facilitators' own patriarchal socialization shaped the intervention. Case Study 4 shows how social network analysis can reveal how norms circulate, who enforces them, and where resistance is constrained, adding explanatory power to gender evaluations of social-norms interventions.

Ethics

Across cases, ethics extends beyond compliance to a relational feminist practice anchored in care, reciprocity, and agency. Case Study 1 shows how both universal and purely contextual ethics can reproduce harm either through unrealistic procedural requirements or by normalizing gendered and adult-centered biases. Case Studies 2 and 3 demonstrate ethics embedded in evaluation design and methods, including safeguarding, iterative consent practices, and trauma-informed approaches. Case Study 3 specifically addresses the iatrogenic risks of engaging boys in gender awareness: the possibility that this could create household tension or expose boys to peer ridicule, making ongoing support and safety essential. The evaluation also raises important questions about whether deep gender transformation can be adequately captured within standard short-term evaluation timeframes, and it embedded trauma-informed principles when facilitators worked with boys who disclosed experiences of violence or abuse. Taken together, these cases suggest that gender-transformative evaluation ethics must be sustained throughout design, data generation, interpretation, and dissemination, not treated as a one-off procedural step and ensure ethical accountability.

Challenges of Capturing Changes in Power

A critical limitation evident across the case studies is the challenge of measuring shifts in power relations and structural change. While individual attitude change (Case Study 3) and network-level norm transmission (Case Study 4) can be documented, deeper transformations in institutional power, decision-making structures, and whose interests are served are more difficult to measure within standard evaluation timeframes. Case Study 2 addresses this by reframing evaluation itself as a site of power redistribution rather than solely as a measurement tool. This suggests that evaluations seeking to document

power shifts must employ longitudinal designs, include structural indicators alongside individual measures, and be explicit about the limits of what can be known and measured within conventional evaluation cycles.

Global South Specificity. The case studies also underscore what makes these evaluations specifically Global South–rooted: They emerge from and are accountable to social movements and community knowledge in Asia, Latin America, and Africa. They address power relations shaped by colonialism, caste, racism, and global economic hierarchies. They integrate creative, context-relevant, informal, and non-Western ways of knowing alongside formal research methods. They center the leadership, voice, and analysis of Global South feminists, evaluators, and grassroots organizations. This contrasts with “gender evaluation in the Global South” conducted by external evaluators or organizations, which may inadvertently reproduce Northern frameworks. True Global South approaches require that Global South women/men/others, and particularly those most marginalized, lead both the evaluation work and the interpretation of findings as well as review recommendations and their implications.

It is important to acknowledge the limitations of the case studies presented. While they offer rich illustrations of Global South feminist evaluation practice, they are not exhaustive. They focus on certain regions (Latin America, South Asia, and references to the Global South more broadly) and certain types of programs (gender equality, social norms, feminist MEL, masculinities). They also reflect evaluations conducted by evaluators with positionalities and training. Future work should expand the case base to include evaluations from diverse geographic regions, sectors, and evaluator backgrounds to strengthen the framework further.

3H and the Case Studies: A Forward Look

The 3H study (Ofir et al., this issue) reinforces the cross-case lessons by calling for epistemological transformation in evaluation. Drawing on 24 prescriptive framework statements from Global South–oriented scholars (including gender), it highlights shifts that align closely with the five domains of the Global South decolonized framework, such as power dynamics, intersectionality, inclusion and representation, methodological approaches, and ethics. In essence, it calls for moving from standardized, externally driven evaluation to context-adapted, community-

owned, and justice-oriented practice: Methods and criteria should be tailored to local realities and diverse ways of knowing; participation should be co-ownership in framing questions, defining success, interpreting evidence, and validating findings; gender should be integrated and be intersectional with explicit attention to whose knowledge is prioritized; locally grounded analytical methods and participatory approaches should guide evidence generation and sense-making; and ethics should be relational and continuous, embedding power analysis, community safety, and accountability throughout the evaluation cycle so communities are decision-makers, not only subjects of evaluation.

Conclusion

We have argued that gender-transformative evaluation in the Global South requires centering GS feminist epistemologies and moving beyond tokenistic efforts to “add gender” or “respect culture” to evaluation methodology. The GS decolonized framework organized around power, intersectionality, inclusion and representation, methodological approaches, and ethics provides practical scaffolding for commissioners and evaluators seeking to reshape how gender is understood, evaluated, and acted upon in diverse contexts.

The four case studies demonstrate that this is not theoretical abstraction but lived evaluative practice grounded in struggle, learning, and commitment to gender justice—from Latin America’s interrogation of power blindness to CREA’s cocreation of feminist MEL with structurally excluded groups, to ECF’s gender-transformative work with boys and men in India, to social network analysis revealing intra-gender enforcement of norms. These cases show that alternative approaches are possible, rigorous, and more likely to catalyze transformative change than standardized, Northern-dominated frameworks.

Engaging explicitly with the colonial roots of contemporary inequality is what binds these cases together as GS feminist evaluation, rather than as “evaluation conducted in the Global South.” Case Study 1 traces how internal colonialism and high-intensity colonial/modern patriarchy in Latin America continue to organize gender hierarchies, Indigenous subordination, and language and translation dynamics that shape who can speak in an evaluation and whose knowledge is treated as evidence. Case Study 2 situates CREA’s feminist MEL practice within histories of colonial knowledge extraction and donor-driven

accountability, showing how language justice, cocreation, and survivor-centered design are explicit responses to those legacies, particularly for sex workers, LGBTQ+ and gender-diverse people, women with disabilities, and communities marginalized by caste, class, and ethnicity. Case Study 3 reads adolescent masculinities in low-income urban India through the long shadow of colonial constructions of caste, class, and respectability, recognizing that hegemonic masculinity is not only a patriarchal but also a colonial inheritance shaped by class and caste. Case Study 4 frames intra-household norm enforcement and the “role reversal” of senior women within colonial and post-colonial reordering of caste, kinship, and rural authority, in which women’s relative power within households is itself shaped by broader histories of dispossession and structural violence. Reading the cases through this colonial lens makes clear that gender-transformative evaluation in the Global South cannot be neutral about history: It must name colonial legacies as constitutive of present inequality, and design evidence and interpretation processes that allow communities to contest, rather than reproduce, those legacies.

The final insight concerns evaluator competencies and systemic change. These competencies—the capacity to name power blindness, design for intersectionality, redistribute participation, employ context-fit methods, and practice relational ethics—are not peripheral. They are the mechanism through which the five domains become everyday evaluative practices rather than being aspirational. However, evaluator competencies alone are insufficient. Systemic transformation of evaluation structures is also required. Evaluation systems must be reformed to value GS knowledge, support GS evaluators, and enable diverse approaches to flourish, moving toward greater equity in the global evaluation enterprise.

In conclusion, GS worldviews provide both analytic and practical foundations for transforming entrenched inequalities and advancing gender-just evaluation. Realizing this potential requires not only individual reflexivity but also collective commitment to decolonizing the evaluation systems and structures that perpetuate Northern epistemic dominance. The framework presented here offers entry points; the real work lies in implementation, adaptation, and community-led deepening of these approaches across diverse contexts and struggles for gender justice and liberation.

References

- Anthias, F. (2012). *Transnational mobilities, migration research and intersectionality: Towards a translocational frame*. Routledge.
- Arora-Jonsson, S. (2011). Virtue and vulnerability: Discourses on women, gender and climate change. *Global Environmental Change*, 21(2), 744–751. <https://doi.org/10.1016/j.gloenvcha.2011.01.005>
- Batliwala, S. (2010). Taking the power out of empowerment—An experiential account. In A. Cornwall & D. Eade (Eds.), *Deconstructing development discourse: Buzzwords and fuzzwords* (pp. 111–121). Practical Action Publishing.
- Batliwala, S. (2011). *Feminist leadership for social transformation: Clearing the conceptual cloud*. CREA. <https://creaworld.org/wp-content/uploads/2020/11/feminist-leadership-clearing-conceptual-cloud-srilatha-batliwala.pdf>
- Batliwala, S. (2012). *Changing their world, concepts and practices of the women’s movement*. Association for Women’s Rights in Development (AWID).
- Batliwala, S., & Pittman, A. (2010). *Capturing change in women’s realities: A critical overview of current monitoring & evaluation frameworks and approaches*. Association for Women’s Rights in Development (AWID).
- Brewster, A., MacDonald, A., & Kossew, S. (2019). Introduction: Gender and violence in cultural texts of the Global South. *Australian Humanities Review*, 64, 87–101. <https://australianhumanitiesreview.org/2019/05/27/introduction-gender-and-violence-in-cultural-texts-of-the-global-south/>
- Chaplowe, S., Morkel, C., de la Garza, T., Buckton, S., Ferretti, S., Nandi, R., & Zaveri, S. (2026). Rethinking Evaluation Competencies for the Transformational Imperative. *Journal of MultiDisciplinary Evaluation*, 22(53), 41–69.
- Collins, P. H. (2000). *Black feminist thought: Knowledge, consciousness, and the politics of empowerment* (2nd ed.). Routledge.
- Commission on the Status of Women. (2014). *Challenges and achievements in the implementation of the Millennium Development Goals for women and girls: Agreed conclusions (CSW58)*. UN Women. https://www.unwomen.org/sites/default/files/Headquarters/Attachments/Sections/CSW/58/CSW58_Agreed_Conclusions.pdf

- Connell, R. (2015). Meeting at the edge of fear: Theory on a world scale. *Feminist Theory*, 16(1), 49–66. <https://doi.org/10.1177/1464700114562531>
- Cornwall, A., & Rivas, A.-M. (2015). From 'gender equality and women's empowerment' to global justice: Reclaiming a transformative agenda for gender and development. *Third World Quarterly*, 36(2), 396–415. <https://doi.org/10.1080/01436597.2015.1013341>
- Crenshaw, K. (1989). Demarginalizing the intersection of race and sex. *University of Chicago Legal Forum*, 1989(1), 139–167.
- Dankelman, I. (2010). *Gender and climate change: An introduction*. Earthscan.
- Das, D. (2017). What's in a term: Can feminism look beyond the Global North/Global South geopolitical paradigm? *M/C Journal*, 20(6). <https://doi.org/10.5204/mcj.1283>
- De, I., Singh, S., & Panda, S. S. (2021). *Social norms, gender and collective behaviour: Development paradigms in India*. Palgrave Macmillan.
- Equal Measures 2030. (2024a). *2024 SDG Gender Index: A gender-equal future in crisis*. https://equalmeasures2030.org/wp-content/uploads/2024/09/EM2030_2024_SDG_Gender_Index_EN_digital.pdf
- Equal Measures 2030. (2024b). *No country achieved gender equality: The SDG Gender Index findings* [Press release]. <https://equalmeasures2030.org/press/no-country-achieved-gender-equality-the-sdg-gender-index-findings/>
- Ernstberger, A. L. (2020). The enduring struggle: Findings from the history of women's and gender studies in the Global South. In J. Shayne & N. Carter (Eds.), *Persistence is resistance: Celebrating 50 years of Gender, Women & Sexuality Studies* (pp. 49–62). University of Washington Libraries. <https://uw.pressbooks.pub/happy50thws/chapter/the-enduring-struggle-findings-from-the-history-of-womens-and-gender-studies-in-the-global-south/>
- Espinosa Miñoso, Y., Gómez Correal, D., & Ochoa Muñoz, K. (Eds.). (2014). *Tejiendo de otro modo: Feminismo, epistemología y apuestas descoloniales en Abya Yala*. Editorial Universidad del Cauca.
- Foucault, M. (1990). *The history of sexuality: An introduction* (Vol. 1, R. Hurley, Trans.). Vintage Books. (Original work published 1978)
- GWL Voices. (2023). *Feminist framework for the Secretary-General's report: "Our Common Agenda"* [Policy briefs collection]. Coalition for the UN We Need; Global Governance Innovation Network. <https://ggin.stimson.org/wp-content/uploads/2023/05/Feminist-policy-briefs-on-OCA-1.pdf>
- Huq, C. (2025, July 2). Feminist interventions from the Global South in international law: Insights from Bangladesh. *Georgetown Journal of International Affairs*. <https://gjia.georgetown.edu/2025/07/02/feminist-interventions-from-the-global-south-in-international-law-insights-from-bangladesh/>
- Kabeer, N. (1996). Agency, well-being and inequality: Reflections on the gender dimensions of poverty. *IDS Bulletin*, 27(1), 11–21. <https://doi.org/10.1111/j.1759-5436.1996.mp27001002.x>
- Kabeer, N. (1999). Resources, agency, achievements: Reflections on the measurement of women's empowerment. *Development and Change*, 30(3), 435–464.
- Kabeer, N. (2010). *Can the MDGs provide a pathway to social justice? The challenge of intersecting inequalities*. IDS/MDG Achievement Fund.
- Kallos, A. (2025). *Applying trauma-informed evaluation principles in gender-based violence evaluations*. Eval Academy. <https://www.evalacademy.com/articles/applying-trauma-informed-evaluation-principles-in-gender-based-violence-evaluations>
- Lugones, M. (2007). Heterosexuality and the colonial/modern gender system. *Hypatia*, 22(1), 186–209. <https://doi.org/10.1111/j.1527-2001.2007.tb01156.x>
- Macheke, F., & Michaeli, I. (2023). *Global South civil society lens: Towards Germany's feminist international cooperation* [Policy paper]. Center for Feminist Foreign Policy; Association for Women's Rights in Development (AWID).
- Mama, A. (2007). Is it ethical to study Africa? Preliminary thoughts on scholarship and freedom. *African Studies Review*, 50(1), 1–26. <https://doi.org/10.1353/arw.2005.0122>
- Mohanty, C. T. (1988). Under Western eyes: Feminist scholarship and colonial discourses. *Feminist Review*, 30, 61–88.
- Mohanty, C. T., Russo, A., & Torres, L. (eds.). (1991). *Third World women and the politics of feminism*. Indiana University Press.
- Mohanty, C. T. (2003a). *Feminism without borders: Decolonizing theory, practicing solidarity*. Duke University Press.
- Mohanty, C. T. (2003b). Under Western eyes revisited: Feminist solidarity through anticapitalist struggles. *Signs: Journal of*

- Women in Culture and Society*, 28(2), 499–535.
- Murthy, R. K. (Ed.). (2001). *Building women's capacities: Interventions in gender transformation*. Sage Publications.
- Narayan, U. (1998). Essence of culture and a sense of history: A feminist critique of cultural essentialism. *Hypatia*, 13(2), 86–106.
- Nnaemeka, O. (2004). Nego-feminism: Theorizing, practicing, and pruning Africa's way. *Signs: Journal of Women in Culture and Society*, 29(2), 357–385. <https://doi.org/10.1086/378553>
- Ofir, Z., Kendrick, I., Goldman, I., Ferretti, S., Norins, J., Sithole, S., Belzer, A., Chaplowe, S., Diwakar, Y., Kalugampitiya, A., Lareza, A. E., Verveld-Nguyen, L., Zaveri, S. (2026). Where Next for Evaluation? Navigating the Polycrisis Through Three Horizons. *Journal of MultiDisciplinary Evaluation*, 22(53), 12–40.
- Paredes, J., & Silva Guimarães, D. (2025). Dialogue with Indigenous women's historical memory requires epistemic openings. *Review of General Psychology*, 29(1), 123–138. <https://doi.org/10.1177/10892680241289350>
- Parekh, S., & Wilcox, S. (2020). Feminist perspectives on globalization. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2020 ed.). <https://plato.stanford.edu/archives/spr2020/entries/feminism-globalization/>
- Parsons, B. A., & Winters, K. (2023). Paradigm-based evaluation for eco-just systems transformation. *Journal of MultiDisciplinary Evaluation*, 19(44).
- Piper, N. (2008). Feminisation of migration and the social dimensions of development: The Asian case. *Third World Quarterly*, 29(7), 1287–1303. <https://doi.org/10.1080/01436590802386427>
- Rance, S., & Salinas Mulder, S. (2001). *Investigando con ética: Aportes para la reflexión-acción*. CIEEP & Population Council.
- Reinharz, S. (1992). *Feminist methods in social research*. Oxford University Press.
- Salinas Mulder, S., & Amariles, F. (2014). Latin American feminist perspectives on gender power issues in evaluation. In S. Brisolara, D. Seigart, & S. SenGupta (Eds.), *Feminist evaluation and research: Theory and practice* (pp. 224–254). Guilford Press.
- Salinas Mulder, S., & Amariles, F. (2016, September). Cultural competence and power dynamics. *Evaluation Connections*. European Evaluation Society.
- Salinas Mulder, S. (2015). *The power blindness wheel* [Conference presentation]. American Evaluation Association annual conference, Chicago, IL, United States.
- Salinas Mulder, S., Rance, S., Serrate Suárez, M., & Castro Condori, M. (2000). Unethical ethics? Reflections on intercultural research practices. *Reproductive Health Matters*, 8(15), 104–112. [https://doi.org/10.1016/S0968-8080\(00\)90011-9](https://doi.org/10.1016/S0968-8080(00)90011-9)
- Schaeffer, H., & Zaveri, S. (2023). *Outcome Mapping + Equity, Gender, and Social Justice*. <https://www.outcomemapping.org/resources/outcome-mapping-equity-gender-and-social-justice>
- Segato, R. L. (2021). Gender and coloniality: From low-intensity communal patriarchy to high-intensity colonial/modern patriarchy. *Hypatia*, 36(3), 479–498.
- UN Women. (2022). *Intersectionality resource guide and toolkit*. UN Women.
- UN Women. (2026). *Ideologies, institutions and power: Addressing discriminatory social norms*. UN Women.
- UN Women & DESA. (2025). *Progress on the Sustainable Development Goals: The gender snapshot 2025*. UN Women; United Nations Department of Economic and Social Affairs, Statistics Division.
- UN Women, Women Count, & DESA. (2024). *The progress on the Sustainable Development Goals: The gender snapshot 2024*. <https://unstats.un.org/sdgs/gender-snapshot/2024/GenderSnapshot2024.pdf>
- United Nations. (1996). *Beijing declaration and platform for action: Fourth World Conference on Women, Beijing, China, 4-15 September 1995*. Department of Public Information. <https://digitallibrary.un.org/record/1648092>
- VeneKlasen, L., & Miller, V. (2002). Power and empowerment. In *A new weave of power, people & politics: The action guide for advocacy and citizen participation* (Chapter 3). World Neighbors.
- Wilson-Grau, R., & Núñez, M. (2007). Evaluating international social-change networks: A conceptual framework for a participatory approach. *Development in Practice*, 17(2), 258–271. <https://doi.org/10.1080/09614520701197226>
- World Bank. (2025). *134 years to go: The data behind achieving gender equality*. World Bank Blogs. <https://blogs.worldbank.org/en/opendata/134-years-to-go--the-data-behind-achieving-gender-equality-this->