Synthesis of feminist evaluation innovations in crisis contexts

Rai Sengupta

Funded by the Feminist Innovation in Monitoring and Evaluation (FIME) Small Award for Young and Emerging Evaluators

Acknowledgements

This product was developed by the author Rai Sengupta with the support of the Feminist Innovation in Monitoring and Evaluation Award Program, implemented by the Global Evaluation Initiative with the generous financial support of Global Affairs Canada and in collaboration with EvalGender+.

The views expressed in this publication are those of the author and do not necessarily reflect the views of the Global Evaluation Initiative (GEI) or Global Affairs Canada. For more information, visit the GEI website.

Table of Contents

Executive Summary 4
1. Introduction 8
1.1 Crises compound gendered risks and vulnerabilities 9
1.2 Conventional evaluation metrics mask inequities 9
1.3 Why feminist evaluation offers a critical alternative 10
1.4 About the Project 11
2. Objectives 12
2.1 Research Question 13
2.2 Research Goals and Objectives 13
3. Methodology 14
3.1 Search Strategy 15
3.1.1 Search terms 15
3.1.2 Data sources 16
3.2 Study Screening and Inclusion 16
3.3 Data Extraction 20
3.4 Data Analysis 21
4. Overview of Included Evaluations 22
4.1 Search Outputs 23
4.2 Characteristics of Included Studies 24
5. Findings 28
5.1 Design Innovations 29
5.2 Methods Innovations 31
5.3 Analytical Innovations 33
5.4 Ethics Innovations 34
6. Discussion 37
6.1 Cross-Cutting Lessons 38
6.2 Implementation Challenges 39
6.3 Evidence and Practice Limitations 40
7. Conclusion & Future Directions 41
7.1 Directions for Policy and Institutional Reform 42
7.2 Directions for Evaluators and Practitioners 42
7.3 Implications for the Feminist Evaluation Toolkit 43
Annexes 45
Annex 1: Bibliography 46
Annex 2: List of Included Evaluations 48
Annex 3: List of Themes and Codes 48

Abbreviations and acronyms

ALNAP	Active Learning Network for Accountability and Performance
FIME	Feminist Innovation in Monitoring and Evaluation
GBV	Gender-Based Violence
GEI	Global Evaluation Initiative
IEG	Independent Evaluation Group
INGO	International Non-Governmental Organisation
MSC	Most Significant Change
NGO	Non-Governmental Organisation
OH	Outcome Harvesting
OM	Outcome Mapping
UN	United Nations
WPHF	Women's Peace and Humanitarian Fund
WLO	Women-Led Organisation
WRO	Women's Rights Organisation

Executive Summary

Background

In humanitarian crises, conventional evaluations often prioritise short-term outputs - such as aid delivered or services provided - which can unintentionally sideline deeper questions of equity, voice, power, and inclusion. Many traditional frameworks assume a level of safety and stability that rarely exists amid displacement, insecurity, and complex power dynamics. As a result, women, girls, and other marginalised groups are often under-represented in data collection and knowledge production, with insights instead coming primarily from those who are more visible or accessible, often men or community representatives in safer positions.

Feminist evaluation, grounded in feminist theory, seeks to address these gaps by more intentionally examining power relations, elevating marginalised perspectives, and fostering justice through participatory and inclusive approaches. It broadens both what is considered credible evidence and whose experiences shape it, anchoring analysis in the lived realities of crisis-affected populations. In fragile settings, feminist evaluation moves beyond conventional metrics to place stronger emphasis on accountability, equity, and the potential for transformation within humanitarian response systems.

Despite clear theoretical and practical imperatives, feminist evaluation innovations in crisis settings remain under-documented and underutilized. Operational challenges and a lack of consolidated frameworks limit methodological innovation and scalability.

This qualitative synthesis of feminist evaluation innovations in crisis contexts is one of two outputs from the project *From Insights to Action: Advancing Feminist Evaluation Innovations in Crisis Contexts*, selected as one of six global recipients of the Feminist Innovation in Monitoring and Evaluation (FIME) Small Award for Young and Emerging Evaluators. The Award was implemented by the Global Evaluation Initiative in alliance with EvalGender+, with the financial support of Global Affairs Canada. This synthesis also informs a Feminist Evaluation in Crisis Toolkit (linked here), designed to support inclusive and context-sensitive evaluation in humanitarian emergencies.

Purpose and Objectives

The purpose of this qualitative synthesis is to explore how feminist evaluation innovations have been adapted and applied in crisis-affected contexts.

This specific objectives of the synthesis are to:

- Systematically identify and synthesise feminist tools, methods, and frameworks in crisis settings.
- Examine the operationalization of feminist principles such as power shifting, intersectionality, and voice across emergency contexts.
- Generate actionable insights to inform evaluation policy actors, institutions, evaluators, and practitioners, and to guide development of a practical toolkit for applying feminist evaluation in emergency and humanitarian settings.

Methodology

A qualitative evidence synthesis was undertaken, consolidating diverse, context-specific feminist evaluation practices from crisis settings. The search strategy, informed by the

SPIDER framework and the '7 S Approach,' included iterative term refinement to capture studies employing feminist evaluation methods in humanitarian and crisis contexts. Systematic searches spanned six evaluation-specific databases maintained by United Nations (UN) and humanitarian institutions, complemented by handsearching, citation tracking, and grey literature identification through generic search engines.

Screening adhered to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines with inclusion criteria focusing on qualitative or mixed-method evaluation reports issued between 2019 and 2024, written in English, and rooted in explicit or implicit feminist orientation.

Data extraction used a structured matrix capturing purpose, crisis type, feminist features, methods, analysis, and ethics. NVivo facilitated reflexive thematic coding, combining deductive and inductive codes to identify innovations in design, methods, analysis, and ethics, interpreted via feminist theory to reveal shifts in power, knowledge, and accountability.

Results

The review identified 17 evaluations meeting inclusion criteria, primarily commissioned by multilateral agencies and international non-governmental organisations (NGOs). Studies addressed a range of crises including refugee and displacement emergencies, public health outbreaks (Zika, cholera, COVID-19), climate-induced disasters, and economic crises, encompassing both multi-country and country-specific interventions. Geographically, evaluations spanned major crisis-affected regions. The timeframe extended from 2019 to 2024, showing sustained interest and application of feminist evaluation approaches in recent humanitarian programming.

Findings

Thematic analysis distilled 60 initial codes into 13 descriptive themes and 4 core analytical themes spanning design, methods, analytical, and ethics innovations.

- **Design innovations** reflected deliberate reconfigurations of evaluation architecture to incorporate feminist theory, rights-holder-driven sampling, participatory framing, and indigenous epistemologies. Examples included explicit feminist empowerment frameworks operationalizing multi-dimensional empowerment levels and transformational change models, alongside implicit feminist practices embedded within culturally sensitive programme evaluations.
- **Methods innovations** encompassed creative visual, arts-based, and narrative data collection tools (participatory video, child-centred recall methods), participatory exercises (such as Most Significant Change, Outcome Harvesting, and Outcome Mapping), and co-production strategies that empowered marginalized groups to self-represent and co-define evaluative change. These methods enhanced both data richness and participant agency in volatile and resource-constrained settings.
- **Analytical innovations** involved rigorous intersectional analysis incorporating gender, age, disability, displacement, and socio-economic status, transforming data disaggregation from a reporting formality into a critical feminist power analysis tool. Community-led validation and interpretation processes redistributed analytical authority to affected populations, strengthening credibility and ethical accountability. These approaches revealed nuanced patterns of inclusion, exclusion, and resilience otherwise obscured in conventional analyses.

- **Ethics innovations** advanced participatory safeguarding beyond procedural compliance, integrating 'Do No Harm' as a proactive ethical framework. Adaptive fieldwork protocols ensured participant safety, agency, and dignity amidst crisis risks, including remote data collection adaptations during COVID-19 and multi-layered ethical and data security oversight, embedding feminist ethics as a core methodological driver.

Lessons, conclusions and future directions

Several key lessons and conclusions emerged from this synthesis, revealing the strategic shifts needed at policy, institutional, and practice levels to realize the transformative potential of feminist approaches.

Lessons and conclusions:

- Evaluators and practitioners must integrate feminist framing holistically from design through implementation and analysis, prioritize participatory and arts-based methods that amplify marginalized voices, and use intersectional analysis strategically to shape programmatic and advocacy efforts.
- Institutionalizing participatory evaluation and co-design processes is key. Integrating mechanisms such as rights-holder-driven sampling, community co-validation, and community interpretation within organizational evaluation policies will reinforce feminist evaluation legitimacy, contextual relevance, accountability, and ethical rigor.
- Capacity-building investments focused on qualitative participatory methods, intersectional feminist analysis, community facilitation, and ethical risk management are vital for professionalizing feminist evaluation competencies in humanitarian and development sectors. Sustained institutional reforms are needed to embed these as core practice standards.
- Embedding thoughtful, power-sensitive ethics into evaluation, safeguarding, and risk management should go beyond basic compliance. It needs to promote flexible approaches that adapt to changing crisis conditions through collaboration across sectors.

Future directions and toolkit development:

Building on these findings, the Feminist Evaluation in Crisis Toolkit translates innovations in design, methods, analysis, and ethics into practical guidance for applying feminist principles in crisis contexts. These components strengthen rigor, feasibility, and transformative impact, advancing equity-centred evaluation as a humanitarian priority.

Photo credit: CARE International

Introduction

1. Introduction

In humanitarian crises, evaluations often focus on measuring short-term relief outputs - meals distributed, shelters built, cash transfers issued. These figures may provide a quick snapshot of activity, but they leave unasked the deeper questions of equity: who could access support, whose voices shape responses, whose knowledge counts, and whose needs remain invisible.

This blind spot is not accidental. Conventional evaluation frameworks often rest on assumptions of safety, stability, and participation that rarely hold in crisis environments. Under conditions of displacement, insecurity, and asymmetric power relations, data collection processes tend to privilege the perspectives of those who are most visible or accessible - typically men or community representatives in relatively secure positions - while systematically excluding the lived experiences of women, girls, and other marginalized groups.

Feminist evaluation, grounded in feminist theory, directly addresses these structural omissions by interrogating power relations, amplifying marginalized voices, and advancing social justice through participatory and inclusive inquiry (Podems, 2010). It redefines both what counts as evidence and who determines it, anchoring analysis in the realities of those most affected by crisis. In fragile and high-risk contexts, feminist evaluation thus moves beyond traditional performance measurement to become a mechanism of accountability and social justice - ensuring that humanitarian response systems are evaluated not only for their efficiency and effectiveness but also for their equity, inclusivity, and transformative potential.

1.1 Crises compound gendered risks and vulnerabilities

Extensive evidence demonstrates that crises exacerbate pre-existing gender inequalities. Women and girls experience disproportionate exposure to gender-based violence, economic insecurity, and mobility constraints during humanitarian emergencies (UN Women, 2021; UNICEF, 2025). The intensification of unpaid care burdens, loss of livelihoods, and contraction of essential services particularly affect those in low-income or informal sectors (Agarwal, 2021; UNDP, 2021). Among refugee and displaced populations, women face compounded vulnerabilities - including limited legal protection, precarious employment, and heightened risk of exploitation (Akande, 2025).

These effects are both immediate and intergenerational, entrenching cycles of social and economic disadvantage. Crisis responses that fail to engage with these gendered dynamics risk reproducing structural inequities rather than disrupting them. Accordingly, evaluation must extend its focus beyond quantifying assistance delivered to systematically interrogating who benefits, who is excluded, and how entrenched gender hierarchies mediate access, agency, and resilience.

1.2 Conventional evaluation metrics mask inequities

Despite these realities, evaluation in crisis contexts often remains disproportionately focused on outputs rather than outcomes. Quantifying services rendered or goods distributed may demonstrate operational efficiency, yet it obscures deeper questions of equity, access, and justice. For example, a humanitarian capacity-building programme may meet numerical targets for training sessions delivered while failing to assess whether women could safely participate, whether their priorities informed the design, or whether participation translated into sustained empowerment or agency.

Structural and contextual barriers further constrain the inclusion of women's and girls' perspectives in crisis evaluation. Safety risks, displacement, and psychosocial trauma frequently render participation infeasible or unsafe. Refugee and displaced women may self-exclude from consultations due to fear of reprisal or exploitation; in active conflict zones, visibility itself may entail threats of violence or social sanction. Traditional evaluation frameworks - often conceived for stable, controlled environments - are ill-equipped to capture these hidden dynamics of exclusion. Consequently, data generated under such conditions tend to privilege the perspectives of those most accessible or visible - typically men or community representatives in safer positions - while systematically underrepresenting the most marginalised voices.

As Crupi and Godden (2024) observe, donor-driven evaluation agendas may exacerbate these limitations by privileging indicators aligned with funder accountability requirements over locally grounded knowledge systems. The result is an incomplete evidentiary base - one that reflects the mechanics of delivery rather than the lived realities of crisis-affected populations. Without deliberate feminist framing, evaluations risk perpetuating epistemic inequities and reinforcing, rather than challenging, structural patterns of exclusion in humanitarian response.

1.3 Why feminist evaluation offers a critical alternative

To tackle these gaps, feminist evaluation brings a fundamentally different lens. Feminist evaluations of humanitarian programming start from the recognition that men, women, girls, boys, and people of all genders experience crises very differently - and that these intersecting inequalities (linked to race, class, disability, or sexuality) must shape humanitarian action (Ward & Voss, 2024). For evaluations, this does not just mean better representation on paper but building processes where those the most affected are directly involved in shaping what success means and how it is measured.

Several guiding principles make feminist evaluation distinct and essential in fragile contexts:

1. **Valuing marginalized voices**. Feminist evaluation elevates the knowledge of groups most likely to be silenced, not only recording their experiences but creating structured spaces where they actively inform decisions and hold programs accountable (Podems, 2010; Sette, 2012). This is where feminist evaluation's advocacy roots show most clearly - it ensures that unheard voices do not just enter the report but shape the response. In crisis contexts, this matters because those most at risk -such as displaced women or survivors of violence - often hold the most urgent insights for shaping safe and effective interventions.
2. **Recognizing multiple ways of knowing**. Mainstream frameworks often rely on quantitative metrics, but feminist evaluation acknowledges evidence grounded in emotion, experience, and intuition (Minnich, 1990; Patton, 2002; Hughes, 2002). In humanitarian contexts, this is vital: trauma narratives, coping strategies, and grassroots insights from Women's Rights Organisations (WROs) add depth that numbers alone cannot capture.
3. **Challenging hierarchies**. Feminist evaluation refuses to treat evaluation as neutral, instead questioning who drives the agenda. Are donor priorities overriding local concerns? Are women led groups meaningfully included as decision-makers, or simply consulted as an afterthought (Ward & Voss, 2024)? In fragile contexts, such questions

determine whether evaluations reinforce or dismantle unequal power structures (McClean, 2000).

4. **Intersectionality**. Conventional gender evaluations may often treat women as a single group. Feminist evaluation, by contrast, insists on recognizing diversity - between rural and urban women, between single mothers and married women from different classes etc. (Mohanty, 1997; Connelly et al., 2000). By foregrounding these differences, feminist evaluation ensures responses are not "one size fits all," but truly tailored. In crisis contexts, this is particularly crucial because intersecting inequalities can magnify vulnerability - meaning that a generic "women-focused" response risks leaving the most marginalized behind.
5. **Evaluator as activist**. Feminist evaluators are not detached observers but engaged actors, consciously using evidence to challenge inequities and advocate for systemic change (Podems, 2010; Sette, 2012). In fragile settings, this role is crucial: evaluators become allies capable of amplifying women led initiatives and holding humanitarian systems accountable for justice.

1.4 Gaps in evidence base and uptake

Despite the clear need for feminist evaluation in crisis settings, there remains **limited consolidated evidence** on how feminist evaluation practices have been adapted, tested, and operationalized under emergency conditions. Their application in humanitarian and fragile contexts continues to be **underexplored**, constraining both methodological innovation and the ability to draw transferable lessons across diverse settings. While feminist evaluators have long advocated for inclusive, participatory, and context-sensitive approaches, these practices remain difficult to implement in crises without structured guidance, practical tools, and validated frameworks.

The absence of consolidated knowledge on feminist evaluation innovations across contexts represents a critical barrier to scaling. Without systematic documentation and synthesis, evaluators lack the evidence base needed to identify which innovations are effective, feasible, and adaptable under varying crisis conditions. This can make it more challenging to integrate feminist innovations into humanitarian evaluation and may reduce the likelihood that equity, agency, and context-specific inclusion are consistently prioritised in crisis response. Consolidated knowledge would enable evaluators and policymakers replicate successful innovations and strengthen the overall rigor and transformative potential of feminist evaluations in fragile settings.

1.5 About the project

The project - '*From insights to action: Advancing feminist evaluation innovations in crisis contexts*' is one of 6 global recipients of the Feminist Innovation in Monitoring and Evaluation (FIME) Small Award for Young and Emerging Evaluators. The Award was implemented by the Global Evaluation Initiative in alliance with EvalGender+, with the financial support of Global Affairs Canada. This project addresses these gaps by synthesizing existing feminist evaluation innovations in crisis settings (this document), providing consolidated evidence across geographies. Additionally, it develops a practical Feminist Evaluation in Crisis Toolkit (linked here) that offers clear, adaptable guidance and tools for implementing feminist evaluation methods in crisis settings.

Photo credit: Beyond Intractability

Synthesis Objectives

2. Synthesis Objectives

2.1 Research Question

This review aims to answer the question: "How have feminist evaluation innovations been applied and adapted in evaluations conducted in crisis-affected contexts?"

Reflecting the qualitative nature of this review, the research question is intentionally exploratory designed to broadly map existing knowledge while allowing space for emergent themes and alternative lines of inquiry (Flemming & Noyes, 2021). The question was structured using the SPIDER framework (Sample, Phenomenon of Interest, Design, Evaluation, and Research type) as shown in **Table 1** below, which supports the development of research questions suited to qualitative evidence synthesis (Cooke et al., 2012).

Table 1: Developing the research question using the SPIDER search tool

SPIDER Element	Description
S- Sample	Evaluation reports of projects, programmes, or policies conducted in crisis-affected contexts (natural or human-made) across diverse geographies and sectors.
PI – Phenomenon of Interest	Feminist evaluation innovations adapted and applied within crisis-affected settings.
D- Design	Qualitative or mixed-methods evaluations with implicit/explicit feminist framing.
E- Evaluation	Innovation, insights and lessons related to feminist evaluation in crisis contexts, including challenges, adaptations, enabling factors, and transformative potential.
R- Research Type	Qualitative and mixed-methods evaluation reports (not reviews)

2.2 Research Goals and Objectives

The overall goal of this qualitative synthesis is to advance understanding of how feminist evaluation approaches have been applied and adapted in crisis-affected settings, in order to inform more responsive and equitable evaluation practices for future humanitarian contexts.

The specific objectives of this synthesis are to:

1. **Systematically identify and synthesise feminist evaluation innovations** - including tools, methods, and frameworks – in crisis contexts.
2. **Examine the operationalisation of core feminist evaluation principles** - such as power shifting, intersectionality, and voice - in complex and rapidly evolving humanitarian settings.
3. **Generate actionable insights** to inform future evaluation policy and institutional actors, evaluators, and practitioners, and to guide the development of a practical toolkit for applying feminist evaluation in emergency and humanitarian settings.

Photo credit: UN Women

Methodology

3. Methodology

A **qualitative evidence synthesis** has been chosen for this project due to its ability to consolidate diverse, context-specific insights around feminist evaluation practices across crisis settings. The synthesis element is crucial because it allows for the aggregation and analysis of a wide range of qualitative data, enabling the identification of common themes, patterns, and innovative feminist evaluation approaches used in different crisis contexts.

3.1 Search Strategy

3.1.1 Search terms

Search terms were developed using the SPIDER framework (Cooke et al., 2012) and informed by the '7 S Approach' (Booth, 2016), as detailed in **Table 2**. The search process was iterative, with initial terms tested across evaluation databases and refined based on initial search outputs. Based on preliminary results, the search strategy was refined to improve sensitivity and inclusivity. For example, while the term *"feminist evaluation"* was central, it became evident that several evaluations embedded feminist principles and practices without explicitly using the term. Consequently, related terms such as *"gender-transformative," "gender-sensitive,"* and *"intersectional"* were incorporated into the search string to capture relevant reports that aligned with the review's conceptual scope but used alternate terminology. *Truncation* (*) and *wildcard* (?) symbols were used to enhance search flexibility, and *Boolean operators* 'OR' and 'AND' were applied to build comprehensive search strings.

Table 2: Search terms developed using the SPIDER search tool

SPIDER Element	Search Terms
S- Sample	("evaluation*" OR "project evaluation*" OR "programme evaluation*" OR "policy evaluation*") AND ("crisis" OR "emergency" OR "disaster" OR "conflict" OR "epidemic" OR "displacement" OR "humanitarian" OR "climate")
PI – Phenomenon of Interest	("feminist evaluation" OR "gender-transformative evaluation" OR gender transformative evaluation" OR "gender-sensitive evaluation" OR "gender sensitive evaluation" OR "intersectional evaluation" OR "feminist method*" OR "feminist innovation*" OR "gender justice" OR "intersectional" OR "gender") AND ("crisis" OR "emergency" OR "humanitarian" OR "conflict" OR "disaster")
D- Design	("qualitative" OR "mixed-method*" OR "participatory" OR "transformative" OR "empowerment-based" OR "co-created" OR "community-led" OR "arts-based") AND ("evaluation")
E- Evaluation	("innovation*" OR "lesson*" OR "insight*" OR "good practice*" OR "challenge*" OR "adapt*" OR "enabling factor*" OR "barrier*" OR "transformative potential") AND ("feminist evaluation" OR "gender evaluation")
R- Research Type	("evaluation report*" OR "case study") NOT ("systematic review" OR "literature review")

3.1.2 Data sources

The review conducted *systematic searches across six evaluation-focused databases*: UNICEF Evaluation Database, UN Women Global Accountability and Tracking of Evaluation Use (GATE), UNDP Evaluation Resource Centre, World Bank Independent Evaluation Group (IEG) Repository, UNHCR Evaluation Database, and Active Learning Network for Accountability and Performance (ALNAP) Humanitarian Evaluation, Learning and Performance (HELP) Library. These databases were selected for their comprehensive coverage of evaluations conducted in crisis-affected contexts, their focus on development and humanitarian programming, and their inclusion of grey literature produced by UN agencies, multilateral institutions, and humanitarian networks.

To enhance the comprehensiveness of the search, *supplementary strategies* were employed. These included handsearching selected peer-reviewed journals, as well as conducting forward and backward citation tracking to identify additional relevant studies. Generic search engines such as Google and Google Scholar were also used to locate grey literature and practitioner-led evaluations that may not be indexed in formal repositories.

3.2 Study Screening and Inclusion

The screening of search results followed a structured, multi-stage approach aligned with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Page et al., 2021). All retrieved records were first imported into the reference management software Zotero (Zotero, 2022), where duplicates were identified and removed. Initial screening was then carried out based on titles and abstracts to eliminate clearly irrelevant studies. The remaining documents were transferred to the qualitative analysis software NVivo (QSR International, 2020), where full-text reviews were undertaken using the inclusion criteria outlined in **Table 3**.

- **Study type**: The review included evaluation reports that employed qualitative or mixed-methods approaches to data collection and analysis, were conducted during or after periods of natural (e.g., disasters, epidemics, climate change) or human-made crises (e.g., conflict, displacement, political instability). Only evaluations of specific policies, programmes, or interventions - whether implemented in a single country or across multiple regions - were considered. Previously conducted systematic reviews, meta-analyses, or theoretical papers were excluded. *This focus ensured that the synthesis was grounded in primary evaluation data from real-world crisis contexts, enabling deeper insights into methodological adaptations and innovations in feminist evaluation practice.*
- **Feminist orientation**: Evaluations were not required to be explicitly labelled as “feminist.” Instead, studies were assessed for feminist orientation based on the presence of relevant characteristics in their design, methods, ethics, or intent - such as community-led or co-created processes, the use of arts-based or narrative methodologies, intersectional and empowerment-based analysis, and a clear focus on transforming gendered power relations. *This approach allowed the inclusion of a broader and more nuanced range of feminist-informed evaluations, recognising that many practitioners may embed feminist principles without using the explicit terminology.*

- **Timeframe and language**: Only studies published between 2019 and 2024, and available in English, were included in the review. *This timeframe was chosen to ensure the relevance of findings to recent and ongoing crisis contexts - such as the COVID-19 pandemic and escalating climate and conflict-related emergencies - while the language restriction was applied for practical reasons related to capacity and consistency in analysis.*

Table 3: Inclusion and Exclusion Criteria

Category	Inclusion Criteria	Exclusion Criteria
S – Sample	Evaluation of a project, programme, or policy (including formative, summative, developmental, or real-time types). Evaluations conducted in crisis-affected contexts, including: • Natural crises (e.g., disasters, epidemics, climate change) • Human-made crises (e.g., conflict, displacement, political instability)	Documents that are not formal evaluations (e.g., standalone research studies, audits, or needs assessments). Evaluations with no clear reference to crisis contexts.
PI - Phenomenon of Interest	Evaluation is conducted in the context of a crisis and is feminist in orientation, evidenced by one or more of the following: • **Design**: Community-led or co-created evaluations that shift power to women, girls, and gender-diverse groups as knowledge holders. • **Methods**: Use of arts-based, narrative, or embodied data collection (e.g., photovoice, storytelling, body mapping) that surface lived, emotional, and embodied experiences. • **Analysis**: Application of intersectional, empowerment-based, or gender-transformative analytical frameworks. • **Ethics and Reflexivity**: Critical reflection on evaluator positionality, bias, and power relations, alongside use of ethics of care, trauma-informed, and safeguarding approaches. • **Transformative Intent**: Grounded in feminist or social justice goals, with evaluations designed to challenge gender norms, redistribute power, or influence equity-focused policy.	Evaluations that are: • Participatory without clear feminist framing, intent, or methods. • Limited to basic sex-disaggregated data without deeper gender/power analysis • Lacking any form of critical reflection, trauma-informed ethics, or equity lens • Not oriented toward transformative change or rights-based goals.

Category	Inclusion Criteria	Exclusion Criteria
D – Design	Qualitative or mixed-methods evaluations which employ feminist participatory, arts-based, narrative, or co-creative methods.	Quantitative-only evaluations; systematic reviews, meta-evaluations, or theoretical papers
E – Evaluation	Reports providing insights into methodological innovations, critical reflections, or learning from practice. Includes discussion on challenges, adaptations, power dynamics, and potential for transformative change.	Evaluations that report only on outputs or outcomes without discussion of methods, principles, or feminist praxis.
R – Research	Qualitative and mixed-methods evaluation reports.	Quantitative-only evaluations; previously conducted reviews and syntheses.
Other Considerations		
Time Frame	Evaluations published between 2019 and 2024.	Evaluations published before 2019 or after 2024.
Language	Evaluations published in English.	Evaluations published in languages other than English.

3.3 Data Extraction

Data extraction was conducted using a structured matrix designed to capture key features of each evaluation while allowing for flexibility to document context-specific insights. This matrix collected information on the evaluation's purpose, context, crisis type, feminist characteristics, methodological design, data collection tools, analytical frameworks, and ethical considerations. The matrix also included fields for identifying innovations, adaptations to crisis contexts, and reflections on power dynamics or positionality. This structured yet flexible approach facilitated consistent documentation while accommodating the complexity and diversity of feminist-aligned evaluation practices.

Each included study was reviewed in full, and relevant details were extracted manually into the matrix. Acknowledging that data extraction is not a strictly linear process but involves iterative engagement with the data and emerging themes (Flemming & Jones, 2020), the extraction form was piloted on three evaluations to test its relevance and comprehensiveness. Based on this pilot, some modifications were made to ensure the tool adequately captured the depth and diversity of feminist-aligned evaluation practices. Particular attention was paid to identifying both explicit and implicit markers of feminist orientation, such as participatory processes, use of narrative or arts-based tools, intersectional analysis, and trauma-informed or reflexive practices. Where available, notes on methodological challenges, lessons learned, and recommendations were also extracted. This iterative process allowed the data extraction framework to evolve in response to the characteristics of the included studies, laying the foundation for a nuanced and contextually grounded synthesis. A summary of the data extraction matrix is presented below in **Table 4**.

Table 4: Data Extraction Matrix

Domain	Types of data extracted
Basic information about the evaluation	Title, year, evaluation type (formative, summative, real-time, etc.), commissioning agency, country/region
Crisis context	Type of crisis (natural or human-made), scope and severity, duration
Feminist orientation and design	Evidence of feminist principles (explicit or implicit), including values, goals, or framing
Methods	Use of qualitative, mixed, arts-based, narrative, or embodied methods (e.g. photovoice, storytelling, body mapping)
Analytical frameworks	Use of intersectional, empowerment-based, or gender-transformative analysis
Ethics and reflexivity	Attention to evaluator positionality, trauma-informed approaches, ethics of care, safeguarding
Transformative intent	Efforts to shift power, challenge gender norms, promote social justice, or influence equity-focused policy

Domain	Types of data extracted
Key learnings/reflections	Insights into methodological innovations, challenges, adaptations, or power dynamics in evaluation practice

3.4 Data Analysis

All included evaluation reports were imported into NVivo (QSR International, 2020) for qualitative analysis. An initial round of close reading was undertaken to familiarise with the content, following principles of reflexive thematic analysis (Braun & Clarke, 2006). Line-by-line coding was used to label meaningful segments of text. A hybrid approach guided the development of the coding framework: deductive codes were drawn from the review's core questions - particularly around feminist evaluation features and innovations - while inductive codes allowed novel and unexpected insights to surface from the data (Cavanagh, 1997; Thomas & Harden, 2008). This approach ensured conceptual coherence while remaining responsive to the distinct ways feminist approaches were articulated in practice. The codebook was iteratively refined to ensure that emerging feminist methodological, ethical, and analytical strategies were systematically captured.

The synthesis followed a multi-stage thematic process designed to trace both descriptive patterns and deeper analytical meanings. Open coding first identified a broad range of themes, including feminist innovations in design, methodology, analysis, and ethics. These codes were then clustered into a hierarchical tree structure to explore thematic relationships and intersections. From this, two levels of synthesis were developed: (1) **descriptive themes**, which summarised recurring strategies, tools, and framings used in feminist evaluations in crisis contexts; and (2) **analytical themes**, which interpreted how these practices reflected shifts in power, knowledge production, and ethical accountability. This deeper layer drew on feminist theory and evaluation literature to surface conceptual innovations and practice-based hypotheses (Britten et al., 2002). Rather than merely aggregating findings, the synthesis critically examined how feminist evaluation approaches reimagined knowledge hierarchies, foregrounded lived experience, and advanced gender-transformative goals within volatile and high-stakes settings.

Photo credit: Oxfam

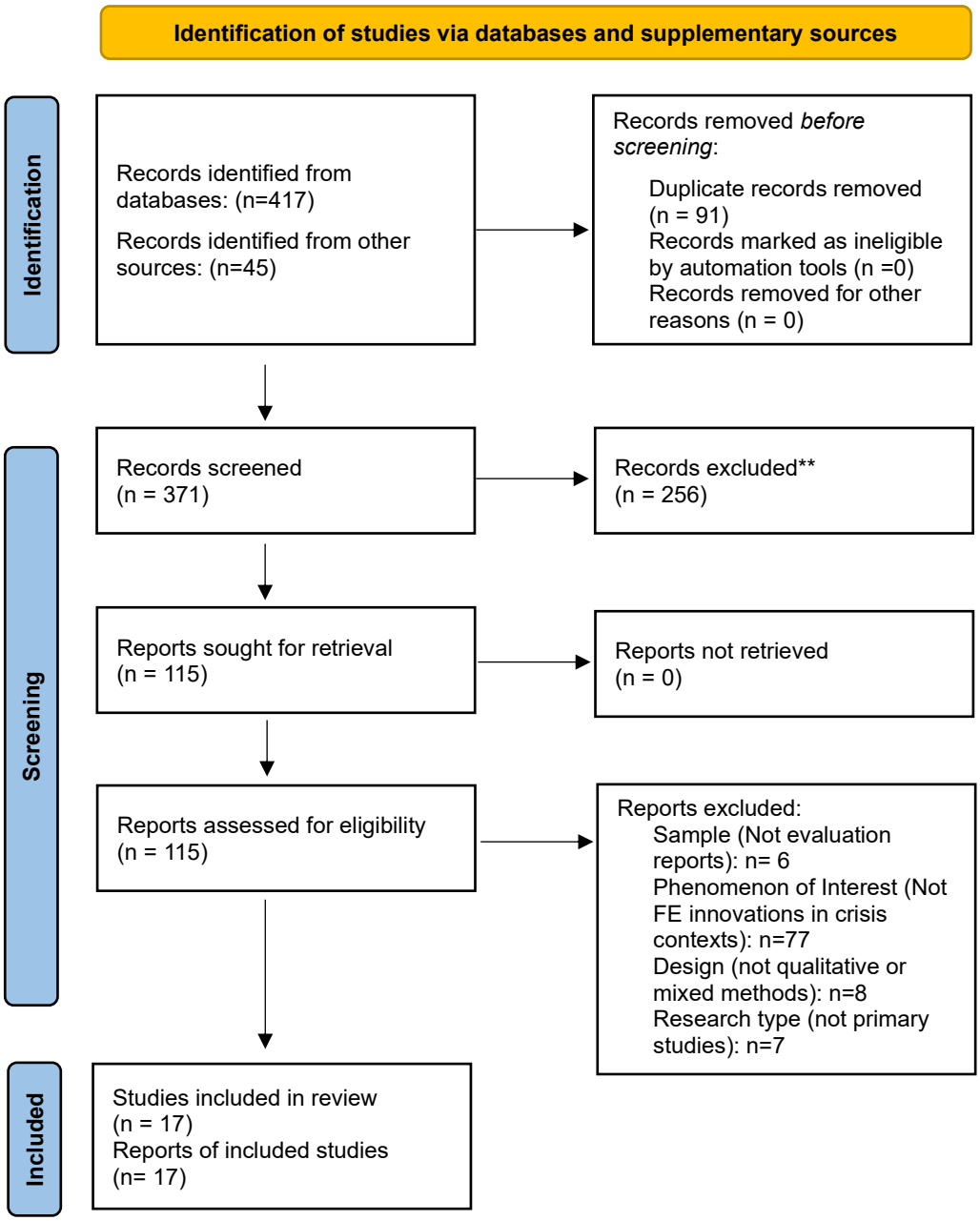
Overview of included evaluations

4. Overview of Included Evaluations

4.1 Search Outputs

The search process yielded a total of 417 records from the 6 evaluation databases. Further, supplemental strategies yielded 45 records: handsearching (19) and grey literature search (26). After the removal of duplicates (n=91 excluded) and the screening of titles and abstracts using the inclusion/exclusion criteria (n=256 excluded), the remaining 115 studies were screened for full texts. Eventually, 17 records were selected for inclusion in the present review, as depicted in the adapted PRISMA flow diagram (Page *et al.*, 2021) in **Figure 1**. No further records were found by handsearching the reference lists of the included studies.

Figure 1: PRISMA flow diagram



4.2 Characteristics of Included Studies

Study type: All 17 evaluation reports were commissioned or published by multilateral agencies or major International Non-Governmental Organisations (INGOs). The largest number were commissioned by UNICEF with 5 reports, followed by UN Women with 4 reports (including one joint evaluation with UNFPA). Other multilateral entities included UNHCR and IEG, while the INGOs included CARE International, Plan International, and the International Federation of Red Cross and Red Crescent Societies (IFRC). Of the total, 8 evaluations were multi-country, while the remaining 9 were country-specific evaluations of humanitarian or development interventions in crisis contexts.

Crisis type: In line with the synthesis objectives, the evaluations encompassed both natural and man-made crises. Refugee and forced displacement crises dominated the evaluation reports (about 7 studies), followed by public health emergencies (5 studies across Zika, cholera, and COVID-19). Climate change-related disasters and economic crises feature in a smaller number of cases.

Geographical spread and timeline: The 17 evaluation reports demonstrate wide geographical coverage across all major regions. There is a strong concentration in Turkey (three evaluations linked to the Syria refugee response), significant representation across Africa (including Nigeria, Uganda, Zimbabwe, Malawi, Mozambique, Burkina Faso, Niger, Mali, Chad, and the Democratic Republic of Congo). Additional coverage includes Central and South America and the Caribbean (Colombia, Honduras), West Asia (Lebanon, Iraq, Jordan), and the Pacific (Solomon Islands). The 17 studies span 2019–2024, with the largest cluster in 2019 (4 studies) and a largely steady spread thereafter - 2020 (1), 2021 (2), 2022 (3), 2023 (4), and 2024 (3). This shows a steady flow of crisis-related evaluations over six years.

The characteristics of the included studies are summarised in **Table 5**. Each evaluation has been assigned a unique code (e.g., E1, E2), which is used to reference the specific evaluation in the findings section. The full list of included studies is presented in **Annex 2**.

Table 5: Characteristics of Included Studies

S. No	Year	Code	Evaluation Report	Commissioning organisation	Country/ies	Crisis Type
1	2019	E1	Evaluation of the community-based child protection services in response to the Syria refugee crisis in Turkey.	UNICEF	Turkey	Syrian refugee crisis in Turkey
2	2019	E2	Evaluation of the Education Personnel Management Strategy as a key component of the UNICEF response to the Syria crisis in the education sector in Turkey.	UNICEF	Turkey	Syrian refugee crisis in Turkey
3	2019	E3	Corporate thematic evaluation: UN Women's contribution to humanitarian action.	UN Women	Multiple countries Case studies in Bangladesh, Cameroon, Colombia, Jordan and South Sudan	Various humanitarian emergencies
4	2019	E4	Community action on Zika project in Honduras and Colombia: Participatory video evaluation report.	IFRC	Honduras and Colombia	Zika virus public health emergency
5	2020	E5	Integrated GBV prevention and response to the emergency needs of newly displaced women, men, girls, and boys in Borno State, North-East Nigeria: Mid-term evaluation report.	CARE International	Nigeria	Refugee crisis in Nigeria
6	2021	E6	Strengthening the resilience of Syrian women and girls and host communities programme: EU Madad - Final evaluation.	UN Women	Iraq, Jordan and Turkey	Syrian refugee crisis in Iraq, Jordan and Turkey
7	2021	E7	A Lifesaving GBV, Women's Leadership, and SRMH Support for Refugees, in Uganda, Arua District, West Nile, Endline Evaluation - Final Report	CARE International	Uganda	Refugee crisis in Uganda
8	2022	E8	Final Evaluation of Gender Inequality of Risk and Promoting Community Resilience Project in Solomon Islands	UN Women	Solomon Islands	Climate change-induced natural

S. No	Year	Code	Evaluation Report	Commissioning organisation	Country/ies	Crisis Type
						disasters in the Solomon Islands
9	2022	E9	Midterm evaluation: Partners for justice - Final report.	Search for Common Ground	Lebanon	Economic contraction, civil unrest, COVID-19 pandemic
10	2022	E10	MEESA report: An evaluation of adolescent girls and young women's continued access to education during COVID-19 in the Middle East, East, and Southern Africa (March 2020–March 2021).	Plan International	Egypt, Uganda, Rwanda, Zimbabwe, Zambia	COVID-19 pandemic
11	2023	E11	Evaluation of the UNICEF emergency social cash transfer programme.	UNICEF	Zimbabwe	Economic crisis and COVID-19 pandemic
12	2023	E12	Evaluation of UNFPA/UN Women GBV two-year Central Emergency Response Fund (CERF) allocation.	UN Women and UNFPA	Multiple countries	Gender-based violence in various emergency contexts
13	2023	E13	Evaluation of UNHCR's response to multiple emergencies in the Central Sahel region: Burkina Faso, Niger, Mali.	UNHCR	Burkina Faso, Niger, Mali	Forced displacement and refugee crisis in Burkina Faso, Niger, and Mali
14	2023	E14	Addressing gender inequalities in countries affected by fragility, conflict, and violence: An evaluation of the World Bank Group's support.	IEG	Burkina Faso, Chad, the Democratic Republic of Congo, Lebanon, the Solomon Islands,	Multiple forms of fragility, conflict, and violence

S. No	Year	Code	Evaluation Report	Commissioning organisation	Country/ies	Crisis Type
					and the Republic of Yemen	
15	2024	E15	Final Evaluation of UNICEF's Response to the Level 2 Cholera and Cyclone Freddy Emergencies in Mozambique.	UNICEF	Mozambique	Cholera public health emergency and Cyclone Freddy
16	2024	E16	Final Evaluation of UNICEF's Response to the Level 2 Cholera and Cyclone Freddy Emergencies in Malawi.	UNICEF	Malawi	Cholera public health emergency and Cyclone Freddy
17	2024	E17	Final evaluation report - Women's Peace and Humanitarian Fund (WPHF) 2019–2023.	Women's Peace and Humanitarian Fund	Multiple countries	Multiple humanitarian contexts – including COVID-19 and peace and security settings

Photo credit: CARE

Findings

5. Findings

Based on the thematic analysis of the 17 studies, 60 codes emerged (see **Annex 3** for the complete list). After grouping these codes into categories, 13 descriptive themes and 4 analytical themes were identified, as shown in **Table 6**.

Table 6: Analytical and Descriptive Themes

Theme Type	Theme	Contributing evaluations
Analytical	**Design Innovations**	**14/17**
Descriptive	Embedding feminist framing in evaluation design	13/17
Descriptive	Embedding participatory and appreciative enquiry in design	14/17
Descriptive	Grounding design in local and indigenous epistemologies	1/17
Descriptive	Feminist-driven sampling approaches	8/17
Analytical	**Method Innovations**	**8/17**
Descriptive	Enabling self-representation through visual, arts-based, and narrative methods	4/17
Descriptive	Co-defining change via structured participatory tools	5/17
Analytical	**Analytical Innovations**	**8/17**
Descriptive	Embedding intersectional, identity-disaggregated analysis	5/17
Descriptive	Closing the loop through community validation and feedback	3/17
Analytical	**Ethics Innovations**	**7/17**
Descriptive	Embedding "Do No Harm"	5/17
Descriptive	Re-engineering fieldwork through risk-responsive protocols	3/17
Descriptive	Integrating multi-layered ethical and data security oversight	1/17

5.1 Design Innovations

Overview - what counts as a design innovation. Across the 17 reports, design innovations are understood as departures from conventional evaluation designs to better fit crisis conditions and feminist aims - e.g., explicitly enacting feminist theory in conceptual framing, implementing indigenous epistemologies, adopting crisis-responsive participatory and appreciative enquiry approaches, and implementing rights-holder driven sampling approaches.

Embedding feminist framing in evaluation design. A few reports provided **explicit feminist theoretical framing and architecture** (e.g., E3; E6; E17). These evaluations operationalised multi-level empowerment frameworks, deliberately designed tools to surface gendered power relations, and explicitly embedded feminist evaluation principles in the overall design. For instance, the 2021 UN Women evaluation which focussed on humanitarian programming for Syrian refugees (E6), used a feminist lens to frame its approach - focusing on systemic power imbalances in patriarchal societies, drawing on women's lived experiences, and **promoting the co-construction of knowledge with women and girls**. A three-dimensional model of women's empowerment guided the evaluation design and implementation: *micro level* –

individual confidence, autonomy, and access to financial, social, and cultural capital; *meso level* – relational empowerment within a woman's immediate environment; and *macro level* – societal structures and norms.

Likewise, the 2024 Women's Peace and Humanitarian Fund evaluation (E17) acknowledged Women-Led Organisations (WLOs) as credible, contextually grounded actors, framing their engagement in dialogue and action as both an **expression of respect and a means of empowerment**, thereby reflecting the feminist principles of participant agency, equality, and mutual learning. The evaluation design demonstrated awareness of power imbalances, engaging participants with humility and recognising them as experts of their own lived realities. This sensitivity extended through data collection and analysis, ensuring that findings were interpreted with fairness, inclusivity, and respect for participants' perspectives. These examples represent innovations by explicitly embedding feminist theory into the conceptual architecture and methodology of the evaluations.

Several other evaluations had implicitly feminist designs- i.e., they implemented practices aligned with feminist values (women-centred sampling, co-construction of knowledge, emphasis on intersectional protection concerns, focus on transformational change) without explicitly using feminist terminology. For instance, the 2023 joint UN Women and UNFPA evaluation (E12) adopted a culturally and gender-sensitive approach to facilitate the meaningful participation of critical stakeholders – including local women's organisations – in evaluation processes. Further, the 2021 CARE International evaluation focussed on humanitarian programming in Uganda (E7), took explicit steps to ensure the safe and meaningful participation of women and girls in the data collection and analysis process, thereby advancing the feminist values of voice, care, and agency. Similarly, the 2023 IEG evaluation of World Bank support to addressing gender inequalities in Fragility, Conflict, and Violence (FCV) affected countries (E14) assessed contributions to "transformational change". This model of transformational change encompassed five elements: relevance, inclusive ownership, depth of change, sustainability, and scale of change. The 'depth of change' element examined whether change was gender transformative and effective in advancing women's and girls' economic empowerment and in preventing and responding to gender-based violence (GBV) - thus implicitly signalling feminist values through its focus on shifting power dynamics without explicitly naming feminist evaluation.

Across these evaluations, feminist design elements were more often embedded in practice than explicitly named. By **normalizing the integration of feminist values in practical ways**, these evaluations demonstrate that transformational, equity-focused design is possible across diverse humanitarian and development evaluations, regardless of whether the approach is overtly named as feminist.

Embedding participatory and appreciative enquiry in evaluation design. Several evaluations deliberately **replaced standard extractive methodologies** with participatory techniques and appreciative enquiry to surface contextually relevant learning under crisis constraints, reflecting a shift in epistemic authority and valuing lived experience. For example, UNICEF's 2024 evaluations of humanitarian programming in the context of Cholera and Cyclone Freddy in Mozambique and Malawi explicitly used mixed methods framed by **appreciative, "light-footprint" approaches** to triangulate community perspectives rapidly while minimising harm (E15; E16).

The 2021 UN Women evaluation focussing on Syrian refugees (E6) similarly deployed a mixed-methods design with heavily participatory components to co-produce findings with

beneficiaries, shifting epistemic authority toward affected women and girls. These design choices functioned as feminist innovations by enabling intentional, rapid, context-sensitive learning while preserving stakeholder voice in volatile settings. Many reports also foregrounded methodological trust in qualitative data - **treating detailed, subjective accounts as valid evidence** for understanding impact and mechanisms in crisis contexts. The 2023 Plan International evaluation (E10) during the COVID-19 pandemic in Middle East, East, and Southern Africa (MEESA) exemplifies this stance, giving **weight to multiple stakeholder narratives** and cross-checking with quantitative data where available.

Grounding design in indigenous and local epistemologies. A 2022 UN Women evaluation – focussed on gender inequality and community resilience in the Solomon Islands (E8) - applied indigenous or locally situated methodologies as a primary design choice. Here, **Tok Stori**, a Melanesian approach framed interviewer–participant relations, question design, and ethical obligations. This shows a feminist innovation in **transferring epistemic authority to local knowledge-holders** and tailoring methods to cultural norms. Such indigenous grounding was not widespread in the dataset but - where present - served both methodological fit and feminist goals, including power-shifting and respect for local agency.

Reaching the right voices: feminist driven sampling approaches. Across several evaluations, sampling strategies explicitly prioritised the inclusion of women's and girls' voices, with purposive adjustments made to ensure gender balance in interview and FGD participants. This often-meant oversampling groups facing heightened exclusion - such as adolescent girls, women with disabilities, or female-headed households - and adapting recruitment protocols mid-process to correct imbalances.

In insecure and resource-constrained environments, a few evaluations introduced **adaptive, rights-holder-driven sampling strategies** to improve representativeness and reach marginalised voices. The 2020 CARE evaluation in Nigeria (E5) implemented a structured snowball protocol, where field enumerators conducted daily KIIs and household surveys guided by project beneficiaries. Gender alternation requirements reduced bias, while randomised beneficiary lists prevented over-representation of specific networks. In the Central Sahel, the 2023 UNHCR evaluation (E13) applied **Respondent-Driven Sampling (RDS**), starting with carefully selected "seeds" able to recruit diverse peers within their networks. Through chain-referral recruitment, combined with attention to network diversity, the method accessed hard-to-reach groups while mitigating bias linked to the initial seed selection. These approaches illustrate how participatory rights-holder-driven sampling can **expand access to marginalised populations** and enhance data credibility - an intentional feminist response where conventional frames are infeasible in volatile contexts.

5.2 Methods Innovations

Overview - what counts as a methods innovation. Across the 17 evaluation reports, a methods innovation is understood as a crisis-responsive use or creation of new methods - as well as the adaptation or redesign of standard evaluation practice - that reconfigures *how* evidence is generated, *who* shapes it, and *whose* knowledge counts. These innovations go beyond procedural flexibility to embed feminist principles that amplify marginalized voices, shift epistemic authority, and safeguard participant agency. They include the introduction of creative visual, arts-based, and narrative methods (such as participatory video, story cards, and child-centred simulated recall), alongside participatory exercises and tools (such as Most Significant Change, Outcome Harvesting, and Outcome Mapping). By combining or replacing

conventional tools with these participatory and rights holder–driven methods, evaluations in crisis contexts surfaced perspectives that are often excluded and generated more contextually grounded, equity-centred evidence.

Enabling self-representation: Visual, arts-based and narrative methods. A subset of evaluations introduced visual, arts-based, and narrative methods of data collection as intentional methodological departures from standard interview/survey formats, giving marginalised groups greater ownership and control over how their experiences were documented, interpreted, and valued in the evidence base. For example, the 2019 Community Action on Zika Project in Honduras and Colombia evaluation (E4) employed **participatory video** via the "Our Story" mobile app, complemented by timelines, photo walks, and story cards, explicitly designed to assess evaluation criteria through the lens of lived experience rather than solely external evaluation judgement.

Child-centred simulated recall was applied in the 2024 UNICEF evaluations of humanitarian programming in the context of Cholera and Cyclone Freddy in Mozambique and Malawi (E15; E16) enabling children to act as primary narrators. Here, children created drawings and guided evaluators through their recollections. This approach enabled exploration of experiences as they unfolded in children's own contexts, free from prescriptive evaluator framing – thereby operationalising feminist commitments to inclusivity, agency, and empowerment. Further, the 2019 UNICEF evaluation in Turkey (E1) applied **H-assessments** - a simple, participatory M&E tool used with adolescents and caregivers to identify a programme's strengths, challenges, and ideas for improvement through group discussions structured around a visual "H" diagram.

In these cases, feminist adaptations were evident, with data collection techniques intentionally designed to capture the perspectives of all gender identities. Particular attention was given to creating safe and enabling spaces for women and girls to speak openly, a choice that not only enhanced the richness and authenticity of the data but also advanced **feminist principles of voice and agency**. By replacing or augmenting conventional methods with creative, participant-led visual and narrative tools, these evaluations overcame common crisis-context limitations and surfaced marginalised perspectives that standard methods often miss.

Co-defining change: Structured participatory exercises and tools. Close to a third of the evaluations used structured participatory tools and exercises, often as a means to capture complex, non-linear change that standard quantitative indicators could not fully reflect. Among these, **Most Significant Change (MSC)** emerged as a particularly important innovation, applied across multiple crisis contexts as a rights-holder-driven way of defining and evidencing impact.

In the 2023 Plan International evaluation in the Middle East, East and Southern Africa (E10), deep dives into "most significant stories" illuminated adaptive education strategies for adolescent girls during COVID-19, while the 2021 CARE International evaluation in Uganda (E7) documented behavioural shifts through "success stories" from women-led initiatives in refugee settings, emphasising women's agency and voice. Similarly, the 2023 UNICEF evaluation of an emergency social cash transfer programme in Zimbabwe (E11) used rights-holder-driven case studies to capture not only the nature of change but also its underlying drivers, validating quantitative findings and providing a richer explanatory layer.

In the case of some protracted crisis contexts, evaluators **combined multiple qualitative and participatory methods** to ensure triangulation, equity, and contextual depth. For example, the 2019 UNICEF evaluation in Turkey (E1) integrated MSC and case stories, segregated by

age and gender, with drawing exercises to elicit change narratives from children. This not only respected children's agency in defining "what change mattered most" but also created intentionally safe spaces where girls and boys could express experiences free from adult or evaluator framing. These stories reflected changes chosen by the child - whether personal, familial, communal, or societal - which were then discussed collectively to identify those best illustrating programme value and impact. Likewise, the 2020 CARE evaluation in Nigeria (E5) employed a multi-tier approach using **Outcome Harvesting (OH), Outcome Mapping (OM)**, and MSC. OH's backward-mapping of observed change to intervention contributions, coupled with iterative stakeholder engagement, allowed women, girls, and other marginalised groups to influence which outcomes were prioritised and validated, demonstrating feminist principles of inclusion, power-sharing, and accountability.

5.3 Analytical Innovations

Overview - what counts as an analytical innovation. Across the 17 evaluations in crisis contexts, analytical innovations are understood as the use or creation of new analytical approaches or feminist and crisis-responsive departures from standard practice - that re-engineer *how* evidence is interpreted, coded, and synthesised to reveal patterns of power, exclusion, and agency that might otherwise remain invisible. These innovations move beyond routine disaggregation to embed intersectional and identity-sensitive analysis at the core of interpretation and integrate community-generated narratives and validation processes directly into the analytical cycle. By involving affected communities - especially women and girls - in interpreting and verifying findings, these analytical frameworks and approaches transform disaggregation into a tool of power analysis, deepening understanding of how gender, age, disability, displacement, and other social identities shape differentiated experiences, outcomes, and pathways of change.

Embedding intersectional, identity-disaggregated analysis. Close to a third of the evaluations moved beyond routine disaggregation to pioneer systematic, crisis-responsive analytical frameworks that embedded feminist and intersectional principles into the structure of interpretation. Data were disaggregated by gender, age, disability, displacement status, and other relevant social markers, then actively applied as analytical lenses to surface previously obscured patterns of vulnerability, resilience, and agency.

In the Community Action on Zika Project in Honduras and Colombia (E4), disaggregation extended across multiple layers - men, women, youth, and community -enabling the extraction of population-specific themes such as community surveillance and behaviour change. This was not disaggregation for reporting's sake, but a **deliberate intersectional mapping** of how health-related challenges and responses differ across identities, reshaping the explanatory logic of the findings while demonstrating feminist commitments to power analysis and nuanced interpretation.

The 2020 CARE evaluation in Nigeria (E5) and the 2019 UN Women Corporate Humanitarian Aid evaluation (E3) exemplified this analytical re-engineering, separating the experiences of men, women, girls, and boys to **pinpoint differentiated protection risks and barriers**. Similarly, the 2023 UNICEF evaluation of an emergency social cash transfer programme in Zimbabwe (E11) used gender-disaggregated analysis to surface disparities that might otherwise be masked, highlighting contrasts in experiences and outcomes and embedding these into the decision-making evidence base. These studies demonstrate

how disaggregation, when harnessed as a **power-analysis tool**, can challenge prevailing assumptions and reframe programme priorities.

Some of the most advanced applications explicitly addressed intersecting vulnerabilities by **combining gender with markers such as displacement, age, and disability**. The 2023 joint UNFPA and UN Women evaluation (E12) rigorously analysed the compounded risks borne by adolescent girls, displaced women, and persons with disabilities, ensuring these realities shaped policy-relevant conclusions.

Closing the loop: Community-led interpretation and validation of findings. A smaller subset of evaluations introduced analytical innovations that embedded community validation and participatory interpretation directly into the analysis process, operationalising the feminist principles of agency and redistributive power. These approaches **intentionally shifted interpretive authority back to affected communities** - particularly women and girls - ensuring that evidence was not only collected *about* them but also analysed *with* them. Mechanisms for feedback and validation were designed to make sense of findings collectively, strengthen credibility, and ensure that insights were returned, discussed, and acted upon locally.

For example, the Community Action on Zika Project in Honduras and Colombia (E4) used participatory video via the "Our Story" mobile app, enabling communities to produce and screen their own stories, and retain copies - an explicit feminist approach to **returning knowledge** and affirming ownership. Similarly, the 2022 Search for Common Ground evaluation in Lebanon (E9) used **Outcome Harvesting validation workshops** held separately with staff and beneficiaries to reduce power imbalances, while the 2023 Plan International evaluation across the Middle East, East and Southern Africa (E10) captured feedback flows from national to household level to reflect adolescent girls' and young women's perspectives in the interpretation of results, thereby advancing the feminist principles of voice, empowerment, and inclusive decision-making.

Collectively, these analytical innovations operationalised feminist principles by redistributing interpretive power, embedding community voice in analysis, and ensuring that marginalised perspectives **directly informed evaluative conclusions** and their policy influence.

5.4 Ethics Innovations

Overview - what counts as an ethics innovation. Across the 17 evaluation reports, an ethics innovation is understood as an intentional, crisis-responsive departure from standard safeguarding practice that actively re-shapes the evaluation's design, methods, or governance to advance feminist aims. It goes beyond procedural compliance to embed participant safety, agency, and dignity as core methodological drivers - for example, making *Do No Harm* the overarching design principle, integrating bespoke referral systems for vulnerable groups, or instituting multi-layered ethical and data-security oversight tailored to conflict or disaster conditions. Such innovations anticipate and mitigate risks in ways that conventional ethics protocols would not, shifting power toward participants and ensuring ethical commitments are operationalised throughout the evaluation process.

Embedding 'Do No Harm'. The "Do No Harm" principle was a central feature in many evaluations, often explicitly framed through a feminist lens, centring participant safety, dignity, and agency. Notably, in the UNICEF evaluations of humanitarian programming in Mozambique and Malawi (E15; E16), Do No Harm was elevated from principle to overarching ethical

framework, guiding both design and data collection. Tools were deliberately crafted to be culturally, contextually, and gender-appropriate, with a focus on minimising risks for women and girls. Similarly, the IEG evaluation of World Bank support to addressing gender inequalities in FCV settings (E14) also reflected this ethic. By ensuring that country case studies were led by gender experts with deep contextual grounding - through lived experience, language fluency, or prior gender-focused work in the country - the **evaluation reduced risks of misinterpretation, inappropriate probing, and cultural harm**. This structural commitment positioned relational sensitivity, safety, and agency at the heart of the evaluation rather than treating them as procedural checks.

A smaller set of evaluations coupled Do No Harm with **active safeguarding mechanisms**. The 2023 Plan International evaluation (E10) integrated on-site referral systems for adolescent girls during data collection, ensuring immediate access to support if child protection concerns arose. This practice directly operationalised feminist ethics, protecting participants while **reaffirming their right to care and support**. Some evaluations chose to exclude high-risk interviews with GBV survivors - such as the 2021 UN Women evaluation (E6) - when safe conditions could not be guaranteed. This prioritisation of participant well-being over comprehensiveness reflects the feminist evaluation ethic of **care as evidence practice** -where moral responsibility supersedes extractive ambition.

Collectively, these examples illustrate how feminist evaluations can push ethics beyond compliance, transforming Do No Harm into a proactive, context-responsive methodology that anticipates harm, re-engineers tools and engagement processes, and embeds participant dignity and agency at the heart of crisis-context evaluation practice.

Re-engineering fieldwork through risk-responsive protocols. In some crisis-context evaluations, risk management was **elevated from a procedural requirement to the core ethical driver** of fieldwork design and execution, resulting in adaptations that embedded feminist principles of safety, equity, and voice directly into practice. In the 2021 UN Women evaluation focussed on the Syrian refugee crisis (E6), a comprehensive COVID-19 risk assessment led to a complete methodological pivot: all interviews and focus group discussions were conducted via secure audio- or video-based platforms. This not only removed physical contact to safeguard participants and researchers but also expanded agency for women and girls - enabling them to engage from safe, familiar environments, protect their privacy, and control the timing and conditions of their participation, thereby reducing exposure to surveillance, coercion, or stigma.

Similarly, in the 2022 UN Women Solomon Islands evaluation (E8), a second wave of COVID-19 triggered a **shift to remote data collection**, followed by targeted in-person fieldwork in Malaita once travel resumed. This adaptive approach helped maintain engagement, while continuously recalibrating to participant safety and access. In the 2020 CARE evaluation in Nigeria (E5), a **formally embedded Risk Mitigation Plan** was operationalised through daily, context-responsive field decisions grounded in feminist ethics. Risk rating tools informed each methodological step, supporting safe inclusion of marginalised groups while minimising exposure to harm and ensuring data collection processes upheld dignity, agency, and participant safety. These crisis-responsive, context-sensitive safeguards illustrate how feminist evaluations can transform risk protocols into mechanisms for voice and agency.

Integrating multi-layered ethical and data security oversight. Formal ethics approvals were a consistent feature across all evaluations, reinforcing procedural accountability and

participant protection, and in one case evolving into an ethics innovation. Notably, the 2023 UNHCR evaluation in Burkina Faso, Mali, and Niger (E13) underwent review by both an institutional ethics committee and an information security office. This **multi-layered ethical and data-security review process** extended oversight past participant welfare to systematically address privacy, information protection, and security risks in fragile, conflict-affected settings. Such integration of ethics and information governance signals a deeper commitment to feminist evaluation principles – by embedding multi-level, accountability to safeguard both people and information in complex humanitarian contexts.

Photo credit: Women's Refugee Commission

Discussion

6. Discussion

6.1 Cross-Cutting Lessons

Across the synthesis of 17 feminist evaluations in crisis contexts, several cross-cutting insights emerge with strategic relevance for both evaluation practice and policy influence. Taken together, the typology of design, methods, analytical, and ethics innovations embodies a coherent shift toward re-centring evaluation on agency, equity, co-production of knowledge, and contextual responsiveness, highlighting how evaluations in crisis settings can catalyse transformational change.

- **The transformative power of feminist evaluation lies in integration**. In many evaluations, design, methodological, analytical, and ethical innovations were not used as standalone features but were integrated - feminist framing shaped design choices, participatory methods were paired with intersectional analysis, and ethical safeguards informed methodological decisions. This layering and interconnected use of feminist principles enhanced their overall methodological rigor and influence, suggesting that applying these elements collectively - rather than in isolation - can deepen the transformative potential of feminist evaluation practice.
- **Shifting epistemic authority is the common denominator.** Across all innovation types, a consistent thread was the deliberate relocation of evaluative power - from external expert judgement toward affected communities, especially women and girls. Whether through rights-holder-driven sampling, narrative self-representation, or community-led validation of findings, these shifts are not just methodological refinements - they reconfigure who shapes the evidence and, by extension, whose realities shape policy and programming priorities.
- **Feminist principles can permeate without explicit labelling**. Many evaluations demonstrated feminist practice without adopting explicit feminist terminology - due to institutional sensitivities, donor requirements, contextual constraints, or other factors influencing the choice of language and framing by evaluators and organizations. The findings show that embedded feminist values (voice, safety, agency, equity) can still transform evaluation processes and outcomes, which broadens the scope for mainstreaming feminist evaluation approaches across agencies that may hesitate to label them as such.
- **Crisis responsiveness is a catalyst for methodological legitimacy**. Innovating under urgent, resource-constrained conditions pushed evaluators toward pragmatic, context-tailored choices that simultaneously aligned with feminist principles. Such adaptations demonstrate that feminist innovations are not in conflict with operational efficiency - they can directly strengthen relevance, validity, and operational feasibility in crises.
- **Adequate time and preparation are critical enablers of participatory methodologies**. Multiple evaluations highlighted that meaningful implementation of participatory, arts-based, or community-led approaches requires extended engagement periods to build trust, align expectations, and ensure ethical practice. Structured preparation, such as orientation sessions or phased activity timelines, emerged as essential to fostering genuine participation and reliable data generation.

- **Community capacity-building strengthens both methodological rigor and ethical quality**. Investing in targeted training and orientation for community members or local enumerators was shown to enhance accurate application of evaluation tools while promoting ethical engagement. These capacity-building efforts improved the richness and validity of evidence and supported feminist commitments to participant agency and respectful data collection.
- **Intersectional power analysis transforms data into action**. Disaggregation alone has long been a compliance requirement, but in some of the evaluations reviewed, it evolved into a lens for interrogating power relations and structural exclusion. Layering gender alongside age, disability, displacement, and other identities - and making these intersecting vulnerabilities central to interpretation - lifted analysis from descriptive reporting to evidence capable of reframing programme priorities and policy dialogue.
- **Ethics as methodology, not compliance**. Elevating *Do No Harm* from a procedural safeguard to the core design driver reshaped how fieldwork was conducted and who could safely participate. Crisis-responsive ethical protocols worked in tandem with participatory approaches, reinforcing that feminist ethics are not a separate governance function but integral to the methods and analytical framing themselves.

6.2 Implementation Challenges

Implementing feminist evaluation innovations in crisis contexts entails navigating a range of complex operational, methodological, and ethical challenges. These challenges stem primarily from the difficult and volatile environments in which evaluations take place and have direct implications for the feasibility and quality of feminist-aligned evaluation approaches.

- **Security, access, and logistical constraints**. Crisis settings often impose severe limitations on evaluators' ability to access communities safely and consistently. Across many evaluations reviewed, security threats, conflict dynamics, and movement restrictions frequently forced early termination, relocation, or rescheduling of field activities. These factors collectively complicate the implementation of participatory, community-led, or time-intensive methodologies central to feminist evaluation. *This underscores the need for adaptive design frameworks, flexible budgeting, and contingency planning mechanisms that prioritize safety and continuity without compromising inclusivity.*
- **Participant engagement challenges**. In crisis contexts, survival priorities often compete with evaluation participation. In some of the evaluations reviewed, conflict, displacement, or lack of essential services sometimes led to non-response or disengagement, requiring significant mediation by local leaders or adaptation of engagement strategies. Ensuring privacy and ethical participation in these contexts further amplified the complexity of data collection. *This calls for realistic planning of participation schedules, strong collaboration with local organisations, the design of ethically appropriate, low-burden engagement strategies, and the operationalisation of safe spaces for data collection.*
- **Resource, time, and capacity limitations**. Feminist evaluation innovations - particularly participatory, arts-based, or community-led methodologies - require extended preparation, orientation, and capacity-building for both evaluators and

community participants. In a few cases, insufficient time or resources for these preparatory steps constrained methodological rigor and the quality of insights generated. Adaptive strategies to manage field-level disruptions, while necessary, placed additional demands on evaluator skill sets and institutional support structure. *This highlights the need to budget explicitly for preparation, capacity development, and iterative adaptation, recognizing these as essential investments for maintaining feminist quality standards in crisis evaluation.*

6.3 Evidence and Practice Limitations

Alongside these operational challenges, the evidence base of evaluation reports reviewed also revealed significant limitations in how feminist evaluation principles and innovations were applied, documented, and institutionalized. Addressing these limitations is essential to strengthening the consistency, transparency, and transformative potential of feminist evaluations in crisis contexts.

- **Community and stakeholder validation gaps.** Direct community feedback mechanisms were included in only a small subset of evaluations and have yet to become widespread practice, with many still relying predominantly on donor/institution-only validation processes. This approach risks excluding the voices of those most affected, potentially diminishing feminist commitments to co-ownership and participatory authority. *Evaluation efforts should prioritize scaling inclusive validation mechanisms that explicitly integrate the perspectives and decision-making power of affected communities.*
- **Uneven application of intersectional analysis**. While gender disaggregation was commonly applied, few evaluations extended analytical frameworks to incorporate other intersecting identities such as disability, age, or socioeconomic status. Culturally embedded analytical approaches were rare but demonstrated distinct value by producing more contextually resonant insights. *Strengthening intersectional and participatory analytical approaches is essential for enhancing both interpretive depth and the transformative potential of crisis evaluations.*
- **Ethics: Procedural versus reflexive applications**. Across many evaluations, procedural ethics - such as formal clearance processes, adherence to ethical codes, and data security - were widely implemented. However, locally embedded, reflexive ethics that dynamically respond to evolving crisis contexts, which are core to feminist innovation, appeared less consistently. *This highlights the need to deepen reflexive ethical practices and embed power-aware, adaptive ethics as a normative standard throughout the evaluation lifecycle.*
- **Limitations in the evidence base**. The documentation varies in depth: few evaluation reports systematically explicate how design choices were made (i.e., the logic linking feminist theory → method → analytic claims). Some innovations appear in practice but lack operational detail. This heterogeneity constrains claims about prevalence and effectiveness; thus, the synthesis treats reported innovations as indicative rather than universally representative.

Photo credit: UNICEF

Conclusions and future directions

7. Conclusion & Future Directions

7.1 Directions for Policy and Institutional Reform

To realize the full potential of feminist evaluation in crisis settings, foundational changes are required at the policy and institutional levels. Strengthening documentation, participatory governance, ethical frameworks, and capacity development creates the enabling environment that sustains feminist innovations and ensures their legitimacy, accountability, and transformative impact. These reforms establish the structural basis for practice-level advances in feminist evaluation methodologies and ethics.

Strengthen documentation and reporting standards: Develop detailed guidance and frameworks for comprehensive articulation of feminist conceptual frameworks through to methodological and analytical decisions. This will promote transparency in operationalizing feminist theory into practice, support replicability, and enhance systematic learning across diverse crisis contexts.

Institutionalize participatory governance and authority-shifting: Formally embed rights-holder-driven sampling, community-led interpretation, and co-validation of findings as evaluation standards. Codification within organizational policies is critical to sustaining feminist evaluation legitimacy, improving contextual fit, and enhancing policy influence through participatory ownership and ethical accountability.

Mainstream intersectional and power-aware ethics: Ensure reflexive, adaptive ethical standards that dynamically respond to crisis changing conditions are integrated across evaluation, safeguarding, and operational risk management systems. Cross-sector collaboration is essential to embedding feminist ethics as continuous, power-sensitive methodologies rather than procedural compliance checklists.

Invest in sustained capacity development and systemic reforms: Promote comprehensive professional development and resourcing for evaluators and institutions across qualitative participatory methods, intersectional feminist analysis, community facilitation, and ethical risk management. Institutional reforms should embed feminist evaluation competencies as core requirements, expanding their reach across humanitarian and development sectors.

7.2 Directions for Evaluators and Practitioners

Feminist evaluation requires intentional actions that embed gender-responsive and power-aware principles throughout all phases of evaluation design, implementation, and analysis. Evaluation practitioners should prioritize integrated feminist framing, intersectional analysis, dynamic ethical approaches, and adaptive, participatory methodologies to ensure evaluations remain contextually relevant, inclusive, and justice-oriented in complex crisis environments.

Embed feminist framing from inception through all evaluation phases: Emphasize integrating feminist principles early in conceptual design, guiding method selection, and analytical framework development, ensuring that principles shape processes holistically rather than retroactively.

Expand the use of participatory qualitative and arts-based methods: Foster inclusive approaches that enable broad community engagement, particularly women, girls, and marginalized groups, through creative, flexible tools that accommodate various constraints in volatile crisis environments.

Treat intersectional analysis as a strategic decision-making tool: Shift from viewing demographic disaggregation as a reporting formality to an active driver of programme design, adaptive advocacy, and policy recalibration. Analytical frameworks should incorporate multiple intersecting identities and structural exclusions, producing contextually rich insights that transform priorities.

Advance community-led, dynamic ethics: Operationalize community-driven ethical governance approaches that empower participants, protect dignity, and balance risk-benefit dynamically over the evaluation lifecycle. Feminist ethics should inform safety protocols, participation modalities, and decision-making criteria consistently.

Promote flexible, adaptive evaluation design: Encourage methodologies that respond pragmatically to logistical, security, and contextual constraints in crisis settings, balancing rigor with feasibility without sacrificing feminist commitments to inclusivity and agency.

7.3 Implications for the Feminist Evaluation Toolkit

Building on these findings, the Feminist Evaluation Toolkit places the four-part innovation typology at the centre of its guidance. The key lessons and insights emerging from the synthesis are translated into practical, useable tools within the toolkit, which operationalises these insights as follows:

- **Feminist framing from the outset.** Tools and templates guide evaluators to embed feminist principles - such as agency, equity, voice, and co-production - at the earliest planning and design stages. This includes support for developing feminist-oriented Terms of Reference, shaping evaluation objectives and questions, and selecting diverse, gender-sensitive teams.
- **Institutionalizing participatory and authority-shifting practices.** The toolkit provides standards, checklists, and case examples for embedding participatory approaches across the evaluation cycle. This includes rights-holder-driven sampling, arts-based and participatory data-collection methods, and tools for creating safe, accessible spaces for engagement in humanitarian contexts. Community validation processes, co-interpretation workshops, and participatory feedback loops operationalise feminist co-ownership, while adaptable facilitation tools support meaningful co-design with affected communities despite crisis constraints.
- **Reframing intersectionality as a strategic and analytical driver.** Moving beyond gender-disaggregated reporting, the toolkit offers practical guidance for applying intersectional power analysis to sampling, data collection, coding, triangulation, and community-driven interpretation. This positions intersectionality as a core decision-making tool - informing programme adaptations, equity analysis, and advocacy efforts - rather than a compliance requirement or an add-on.

- **Integrating ethics into operational risk management**: Ethical practice is reframed as an active methodological component and a core part of organisational risk-management systems. The toolkit outlines procedures for developing adaptive ethics protocols that merge Do No Harm principles with real-time crisis monitoring, safe referral mechanisms, privacy safeguards, and protective measures for both participants and evaluation teams. Guidance also supports evaluators in designing inclusive and trauma-sensitive spaces, enabling ethical decision-making to inform data collection, analysis, and dissemination in volatile environments.

Photo credit: UNICEF

Annexes

Annexes

Annex 1: Bibliography

- Agarwal, B. (2021). Livelihoods in COVID times: Gendered perils and new pathways in India. *World Development*, *139*, 105312. https://doi.org/10.1016/j.worlddev.2020.105312
- Akande, B. (2025). *Gender-sensitive approaches to building resilience in refugee and asylum-seeking women: Strategies, challenges, and policy implications*.
- Booth, A. (2016). Searching for qualitative research for inclusion in systematic reviews: A structured methodological review. *Systematic Reviews*, *5*(1), 1–23.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*.
- Britten, N., Campbell, R., Pope, C., Donovan, J., Morgan, M., & Pill, R. (2002). Using meta ethnography to synthesise qualitative research: A worked example. *Journal of Health Services Research & Policy*, *7*(4), 209–215.
- Cavanagh, S. (1997). Content analysis: Concepts, methods and applications. *Nurse Researcher*.
- Connelly, M., Li, T., MacDonald, M., & Parpart, J. L. (2000). Feminism and development: Theoretical perspectives. In J. Parpart, M. P. Connelly, & V. E. Barriteau (Eds.), *Theoretical perspectives on gender and development* (pp. 51–160). International Development Research Centre.
- Cooke, A., et al. (2012). Beyond PICO: The SPIDER tool for qualitative evidence synthesis. *SAGE Journals*.
- Crupi, K., & Godden, N. J. (2024). Feminist evaluation using feminist participatory action research: Guiding principles and practices. *American Journal of Evaluation*, *45*(1), 51–67. https://doi.org/10.1177/10982140221148433
- Flemming, K., & Jones, L. V. (2020). Using evidence from systematic reviews. In J. V. Craig & D. Dowding (Eds.), *Evidence based practice in nursing* (4th ed., pp. 109–125). Elsevier.
- Flemming, K., & Noyes, J. (2021). Qualitative evidence synthesis: Where are we at? *International Journal of Qualitative Methods*, *20*, 1609406921993276.
- Hughes, C. (2002). *Key concepts in feminist theory and research*. Sage.
- McClean, M. (2000). Alternative approaches to women and development. In J. Parpart, M. P. Connelly, & V. E. Barriteau (Eds.), *Theoretical perspectives on gender and development* (pp. 179–190). International Development Research Centre.
- Minnich, E. (1990). *Transforming knowledge*. Temple University Press.
- Mohanty, C. T. (1997). Under Western eyes: Feminist scholarship and colonial discourse. In N. Visvanathan, L. Duggan, L. Nisonoff, & N. Wiegersma (Eds.), *The women, gender and development reader* (pp. 79–86). David Philip.

- NVivo 12. (2020). QSR International Pty Ltd. Retrieved August 14, 2025, from https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ (Clinical Research Ed.)*, *372*, n71. https://doi.org/10.1136/bmj.n71
- Patton, M. Q. (2002). *Qualitative research and evaluation methods*. Sage.
- Podems, D. (2010). Feminist evaluation and gender approaches: There's a difference? *Journal of Multidisciplinary Evaluation*, *6*(14). http://survey.ate.wmich.edu/jmde/index.php/jmde_1/article/view/199/291
- Sette, C. (2012). Feminist evaluation approach is not just about women. *BetterEvaluation Blog*. https://www.betterevaluation.org/blog/feminist-evaluation-approach-not-just-about-women
- Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, *8*(1), 1–10.
- UN Women. (2021, July 1). *Measuring the shadow pandemic: Violence against women during COVID-19* [Rapid Gender Assessment summary]. UN Women Data Hub. https://data.unwomen.org/publications/vaw-rga
- UNICEF. (2025). *Gender in crisis: A regional compendium—Gender in emergencies programming*. UNICEF. https://knowledge.unicef.org/resource/gender-crisis-regional-compendium-gender-emergencies-programming
- United Nations Development Programme (UNDP). (2021). *Protecting women's livelihoods in times of pandemic: Temporary basic income and the road to gender equality* (Development Futures Series No. 5). https://doi.org/10.18356/30053307-5
- Ward, J., & Voss, J. (2024). *Guidance note on applying feminist approaches to humanitarian action*. Social Development Direct. https://sddirect.org.uk/sites/default/files/2024-08/Guidance%20Note_Applying%20Feminist%20Approaches%20to%20Humanitarian%20Action.pdf
- Zotero. (2020). Roy Rosenzweig Center for History and New Media.

Annex 2: List of Included Evaluations

- AAN Associates. (2023). *Evaluation of the UNICEF emergency social cash transfer programme*. UNICEF Zimbabwe.
- Arbulú, A., et al. (2019). *Corporate thematic evaluation: UN Women's contribution to humanitarian action*. UN Women Independent Evaluation Service.
- Asfour, H., et al. (2021). *Strengthening the resilience of Syrian women and girls and host communities programme: EU Madad - Final evaluation*. UN Women.
- Chan-Tung, A., & Hortle, R. (2022). *Final evaluation of Gender Inequality of Risk and Promoting Community Resilience Project in Solomon Islands: Final report, report to UN Women Fiji Multi Country Office*. Sustineo.
- Durston, S., et al. (2019). *Evaluation of the education personnel management strategy as a key component of the UNICEF response to the Syria crisis in the education sector in Turkey*. UNICEF.
- Emenogu, A., et al. (2020). *Integrated GBV prevention and response to the emergency needs of newly displaced women, men, girls, and boys in Borno State, North-East Nigeria: Mid-term evaluation report*. CARE International.
- Harvard Humanitarian Initiative & Brigham and Women's Physicians Organization. (2023). *Evaluation of UNHCR's response to multiple emergencies in the Central Sahel region: Burkina Faso, Niger, Mali*. UNHCR.
- International Federation of Red Cross and Red Crescent Societies. (2019). *Community action on Zika project in Honduras and Colombia: Participatory video evaluation report*.
- Lwanga, M. M., et al. (2021). *A lifesaving GBV, women's leadership, and SRMH support for refugees in Uganda, Arua District, West Nile: Endline evaluation – final report*. CARE International.
- Plan International. (2023). *MEESA report: An evaluation of adolescent girls and young women's continued access to education during COVID-19 in the Middle East, East, and Southern Africa (March 2020–March 2021)*.
- Samuel Hall. (2023). *Evaluation of UNFPA/UN Women GBV two-year Central Emergency Response Fund (CERF) allocation*. UN Women and UNFPA.
- Tong, K., et al. (2024). *Final evaluation report - Women's Peace and Humanitarian Fund (WPHF) 2019–2023*. Women's Peace and Humanitarian Fund.
- UNICEF. (2019). *Evaluation of the community-based child protection services in response to the Syria refugee crisis in Turkey*. UNICEF Turkey.
- UNICEF. (2024). *Final evaluation of UNICEF's response to the Level 2 Cholera and Cyclone Freddy emergencies in Malawi*. UNICEF Evaluation Office, New York.
- UNICEF. (2024). *Final evaluation of UNICEF's response to the Level 2 Cholera and Cyclone Freddy emergencies in Mozambique*. UNICEF Evaluation Office, New York.
- Voluntas. (2022). *Midterm evaluation: Partners for justice - Final report*. Search for Common Ground.

- World Bank. (2025). *Addressing gender inequalities in countries affected by fragility, conflict, and violence: An evaluation of the World Bank Group's support*. Independent Evaluation Group. **https://ieg.worldbankgroup.org/evaluations/addressing-gender-inequalities-countries-affected-fragility-conflict-and-violence**

Annex 3: List of Themes and Codes

Analytical Theme	Descriptive Theme	Code
Design Innovations	Embedding feminist framing in evaluation design	Explicit feminist theoretical framing
		Multi-level empowerment framework (micro/meso/macro)
		Addressing systemic power imbalances
		Drawing on the lived experience of women and girls
		Co-construction of knowledge with women and girls
		Recognising women-led organisations as credible actors
		Humility and respect for participants' expertise
	Embedding participatory and appreciative enquiry in design	Participatory techniques replacing extractive methods
		Appreciative enquiry approaches
		"Light footprint" approaches
		Triangulation of community perspectives
		Preserving stakeholder voice
		Valuing qualitative accounts as valid evidence
	Grounding design in local and indigenous epistemologies	Use of Tok Stori (Melanesian approach)
		Epistemic authority to local knowledge-holders
		Cultural tailoring of tools and process
	Feminist-driven sampling approaches	Purposive inclusion of women's and girls' voices
		Oversampling marginalised groups (e.g., adolescent girls, women with disabilities)
		Rights-holder-driven adaptive sampling
		Structured snowball protocols
		Gender alternation in respondent selection
		Respondent-Driven Sampling (RDS)
Methods Innovations	Enabling self-representation through visual, arts-based, and narrative methods	Participatory video methods
		Timelines, photo walks, story cards
		Child-centred simulated recall
		Drawings as elicitation tools
		H-assessments with adolescents/caregivers
		Safe and enabling spaces for women and girls
	Co-defining change via structured participatory tools	Most Significant Change (MSC)
		Case stories segregated by gender/age
		Outcome Harvesting (OH)

Analytical Theme	Descriptive Theme	Code
		Outcome Mapping (OM)
		Collective story selection
		Participant-defined change indicators
Analytical Innovations	Embedding intersectional, identity-disaggregated analysis	Disaggregation by gender, age, disability, displacement status
		Intersectional mapping of vulnerabilities
		Identifying differentiated protection risks
	Community-led interpretation and validation of findings	Local screenings of participatory video
		Retaining knowledge in communities
		In-country validation visits
		Separate validation workshops for staff and beneficiaries
		Layered feedback analysis (national to household level)
Ethics Innovations	Embedding "Do No Harm"	Elevating Do No Harm as the overarching ethical framework
		Designing culturally/contextually/gender-appropriate tools
		Minimising risks to women from WROs
		On-site referral systems for safeguarding
	Re-engineering fieldwork through risk-responsive protocols	COVID-19 risk assessments
		Expanding participation options for privacy/safety
		Adaptive data collection responding to evolving risk
	Integrating multi-layered ethical and data security oversight	Dual ethics and information security approvals
		Extended oversight for privacy and information protection
		Data security in fragile/conflict contexts

For further information, please contact:

Rai Sengupta (sengupta.rai01@gmail.com)