

Final Report

External evaluation of the Netherlands WPS 2016-2019 and WPS 2020 programmes

Date: 15 November 2021

Authors: Henri Myrntinen, Rebecca Sibson, Lara Quarterman, Laura Barrios Sabogal, Lona Luduro and Jo Robinson

Submitted by Itad



Acknowledgements

We would like to thank everyone who supported us in this evaluation including all National Action Plan (NAP) III implementing partners and local partners who attended the theory of change workshop, critical reflection sessions or took part in interviews or focus group discussions: your thoughts and ideas are at the heart of this report. We would also like to thank The Taskforce Women's Rights and Gender Equality (TFVG) in the Ministry of Foreign Affairs, Netherlands for their engagement and support during the process.

Disclaimer

The views expressed in this report are those of the evaluators. They do not represent those of the Netherlands Ministry of Foreign Affairs or of any of the individuals and organisations referred to in the report.

'Itad' and the tri-colour triangles icon are a registered trademark of ITAD Limited.

Contents

Contents	ii
List of tables and figures	iii
List of acronyms	iv
Executive Summary	vi
1 Introduction	11
1.1 Evaluation purpose, scope, and target audience	11
1.2 Contextual background	12
1.3 Programme background	13
1.4 NAP theory of change	16
2 Methodology	18
2.1 Evaluation design	18
2.2 Modules 1 and 2: A theory-based approach	18
2.3 Module 3: Case studies	21
2.4 Sampling	22
2.5 Limitations and challenges	22
3. Findings	25
3.1. Inception and quick-scan	25
3.2. Theory of Change workshop	26
3.3. Document review	27
2.6 Critical reflection workshops	35
3.3. Case studies and survey	36
4. Analysis of findings	39
4.1. NAP programme contributions to impact and unintended outcomes	39
4.2. NAP programme contributions to the WPS framework three specific objectives	40
4.3. Alignment with other frameworks and contextual relevance	43
4.4. Sustainability	44
4.5. Efficiency and the consortium model	45
4.6. Lessons learned	47
5. Conclusions	49
6. Recommendations	50

Consortia and ways of working	50
Design, monitoring, evaluation and learning	50
Thematic focus areas	52
7. Annexes	53
Annex 1 – Case studies	53
Annex 2 – Data collection tools, case study approach and sampling	53
Annex 3 – Evaluation matrix	53
Annex 4 – Quick-scan results	53
Annex 5 – NAP III ToC, reconstructed ToC and ToC workshop interactive whiteboard capture	53
Annex 6 – Full list of consortium partners and local partners	53
Annex 7 – List of stakeholders consulted	53
Annex 8 - Bibliography	53
Annex 9 – Full Terms of Reference	53

List of tables and figures

Table 1: No-cost extensions and additional funding for 2020	16
Table 2: Programme evaluations assessed against the IOB criteria.....	21
Table 3: Thematic focus areas of the eight NAP programmes	28
Figure 1: NAP 2016-2019 Theory of change	17
Figure 2: Evaluation Modules.....	18
Figure 3: Contribution analysis steps	20
Figure 4: Theory of Change workshop feedback captured in an interactive whiteboard.....	220
Figure 5: Steps in the evaluation	25
Figure 6: Summary of strength of evidence	33
Figure 7: Sustainability.....	35
Figure 8: Budget variance.....	47

List of acronyms

AWEC	Afghan Women’s Educational Center
CBO	Community-Based Organisation
CCSD	Centre for Civil Society
CDJP	Diocesan Commission “Justice et Paix”
CEV	Comisión para el Esclarecimiento de la Verdad, la Convivencia y la No Repetición
CODACOP	Corporación de Apoyo a Comunidades Populares
CSO	Civil Society Organisation
DIOFAP	Le Directoire des Organisations Féminines pour les Actions de Paix
DRC	Democratic Republic of Congo
EQ	Evaluation Question
FARC-EP	Fuerzas Armadas Revolucionarias de Colombia—Ejército del Pueblo
FGD	Focus Group Discussion
GBV	Gender-Based Violence
HBORL	Hamida Barmaki Organization for the Rule of Law
INGO	International Non-Governmental Organisation
IOB	Policy and Operations Evaluation Department of the Dutch MFA
JEP	Jurisdicción Especial para la Paz
KII	Key Informant Interview
LGBTIQ+	Lesbian, Gay, Bisexual, Trans, Intersex, Queer and Other
LIMPAL	Liga Internacional de Mujeres por la Paz y la Libertad
MEL	Monitoring, Evaluation, and Learning
MFA	Ministry of Foreign Affairs
MTR	Mid-Term Review
NAP	National Action Plan
NGO	Non-Governmental Organisation
PARDE	Parlement des Enfants et des Jeunes
PFP	Psycho-social Focal Point
RMM	Resources Mapping and Mobilisation
SMM	Sautiya Mama Mukongomani
SPR	Synergie des Femmes pour la Paix et la Réconciliation des Peuples des Grands Lacs
SSR	Security Sector Reform
STAD	Support Trust for Africa Development
TFVG	Taskforce on Women’s Rights and Gender Equality
THA	The Hague Academy for Local Governance

ToC	Theory of Change
UN	United Nations
WG4CA	Women and Girls for Change Alliance
WPS	Women, Peace and Security
WPSO	Women and Peace Studies Organization
YLDF	Youth Leadership Development Foundation

Executive Summary

This external end-term evaluation of eight in-country programmes under the third Dutch Women, Peace, and Security (WPS) National Action Plan (NAP) 2016–2019 and its one-year extension, WPS NAP 2020, was commissioned by the Dutch Ministry of Foreign Affairs (MFA) and carried out by Itad, an independent, UK-based consulting company. The eight countries in which the programmes were implemented were Afghanistan, Colombia, Democratic Republic of Congo (DRC), Iraq, Libya, South Sudan, Syria and Yemen. In each of these countries, one consortium consisting of Dutch NGOs, Dutch knowledge institutions and local NGOs was funded by the MFA. It is important to note that the eight programmes constituted only part of the various efforts Dutch NAP signatories have undertaken to implement the third NAP of the Netherlands.

This evaluation consisted of:

- A **meta-evaluation** of the country-programme evaluations that have been externally commissioned by the NAP consortia.
- **Additional desk research, literature review and data collection** to elaborate on and substantiate the conclusions from the country-programme evaluations that were externally commissioned by NAP consortia.
- Using **contribution analysis** to identify where programme outcomes support or challenge the theory of change (ToC) of the NAP.
- A **ToC workshop** with the implementing consortia on the ToC of the NAP.
- Two **critical reflection workshops** on the findings of the meta-evaluation and reflecting on underlying assumptions and causal pathways of the programmes.
- Three **case studies** (two country case studies and one thematic case study – see Annex 1) on programming conducted under the NAP, conducted remotely and in-country with national consultants. The two country case studies are Colombia and South Sudan, and the thematic case study focuses on the specific objective of the decrease of harmful gender norms.

The third Dutch WPS NAP, or NAP III, was guided by an overarching Theory of Change (ToC), upon which also the individual programme ToCs were modelled. The overall objective of the ToC was:

“Together we contribute to an enabling environment for women’s participation and empowerment in conflict and post-conflict environments, so they can meaningfully participate in conflict prevention, resolution, peacebuilding, protection, relief and recovery.”

This objective was to be reached by work on the following specific objectives:

1. **Enhanced protection** – *Better protect women and girls in conflict and post-conflict situations from violence and violations of their rights.*
2. **Decrease of harmful gender norms** – *Subvert harmful underlying gender norms, which are obstacles to sustainable peace.*
3. **Equal leverage in conflict prevention, resolution, peacebuilding, relief and recovery** – *Ensure that women have equal leverage in conflict prevention and resolution, peacebuilding, relief and recovery at all levels, and that their efforts are acknowledged and supported.*

The evaluation was met with data quality limitations in the country programme evaluations and the programme reporting as well as a limited number of independent sources. The evaluation is therefore unable to draw robust conclusions about the extent of outcome level change generated by the

programme, however, was able to develop important findings and learning on issues of relevance, efficiency, coherence, safeguarding and ethics which are relevant to NAP IV and broader WPS programming going forward.

Key findings

The eight in-country programmes sought to address these objectives through a variety of approaches, utilising different entry points and seeking to enact change at the individual, community, sub-national (e.g. municipal, district, provincial or similar), national and international level. As outlined in the activity reports of the consortium partners and local partners, the activities of the NAP programmes reached thousands of women, girls, men and boys in these eight conflict-affected countries. The evaluation of the impact of these broad range of activities proved difficult, however, as the reporting on the programmes mostly focused on activities and numbers of beneficiaries reached, rather than impacts.

Key findings of the evaluation thus were:

Key finding 1: Whilst the country evaluations were rich in information on activities, and sometimes anecdotal outputs and outcomes, evidence was generally weak (not validated or triangulated), and focused on activities and outputs rather than outcomes and impact. The methodologies used were not rigorous, making it difficult to provide results, well-founded conclusions and draw lessons from them. Furthermore, activity reports and evaluations of the country programmes did not always capture the full range of programme activities implemented by partners.

Key finding 2: Implementing partners considered the ToC broad and flexible enough to cover a variety of different approaches taken by the different country programmes; however, it was thought to be overly ambitious in its objectives and the logical pathways to change were found not well expounded.

Key finding 3: Country programme reporting captures activity and some output-level results to demonstrate what the programme has achieved. However, evidence of outcome-level change is sparse, anecdotal and poorly evidenced. Within reporting, there is a 'missing middle' in evidencing intermediary changes which are critical on the pathway towards more ambitious outcome-level change.

Key finding 4: Country programming has been agile and adaptive to difficult and dynamic contexts which has enabled implementing partners to be opportunistic in generating change and continue work in adversity. However, longer-term thinking for programme sustainability was generally not well planned for across programmes.

Evaluation questions

Based on the Terms of Reference of the Dutch MFA, the evaluation team formulated the following six evaluation questions, followed by summary findings.

Evaluation Question 1: What are the NAP programme contributions to the overall objective of WPS framework 2016–2019 and 2020? What are the programme contributions to other positive and negative outcomes (intended and unintended and sensitive to the 'do no harm' principle) and were there any significant differences in this respect between different consortium members/implementing partners?

The NAP programmes aligned well with the overall objective of the Dutch NAP III, and, based on the reporting from the programmes and the evaluation, were able to affect positive changes at the individual, community, sub-national, national and, to an extent, international level. Apart from the intended

outcomes, at least half of the programmes reported intended and unintended positive impacts, as beneficiaries were able to capitalise on skills gained and opportunities created by the programmes. All eight programmes reported some degree of resistance to the promotion of WPS issues, be it from family members of intended beneficiaries, community members, authorities, armed actors – or else decision-makers unwilling or reluctant to engage on these issues. While some of these resistances had been expected and risk mitigation measures taken, others had not been foreseen, or not been foreseen to the degree with which they materialised. The Colombia programme sought to pro-actively counteract the increased risks against women activists more broadly by making addressing these risks an integral part of the implementation.

Evaluation Question 2: What are the NAP programme contributions to the WPS framework 2016–2019 and 2020 three specific objectives and what change happened along the causal pathway - did assumptions hold?

As with the overall objective, the specific objectives of the Dutch NAP III were used by the programmes as a basis for designing their respective activities. There was however great diversity in terms of which specific objectives, how and with whom the various programmes sought to address these. Under the ‘protection’ specific objective, the programmes by and large defined the scope of protection narrowly by focusing on gender-based violence (GBV) – the exception was the additional integration of the protection of women activists in the Colombia programme. Although the scope was narrowed to mainly addressing GBV, within this, the approaches covered not only protection but also awareness-raising, prevention and response. Under ‘gender norms’, the programmes employed a wide variety of approaches, from promoting individual change through dedicated curricula, over broad public awareness raising campaigns, to engaging with duty-bearers and community gate-keepers. Nonetheless, the programmes reported to have been able to build women’s and girls’ capacities to participate and all programmes, to some degree, were able to find entry points at various levels to increase women’s meaningful participation. A lack of baseline data and a focus on activity- rather than impact-level reporting beyond anecdotal evidence made it difficult to establish the degree of impact.

Evaluation Question 3: How far were NAP programmes aligned or responsive to:

- **WPS policy frameworks**
- **National and decentralised policies**
- **The needs of beneficiaries and target groups**
- **The changing context**

The programmes were closely aligned to the Dutch NAP III, and in the case of Afghanistan, also aimed to simultaneously contribute to the implementation of the Afghan WPS NAP. In those countries where it was possible to engage with national policies and institutions – Afghanistan, Colombia, DRC, Iraq, South Sudan and to a lesser degree Yemen – programmes reported to have engaged successfully with these and were in part able to shape these. In the case of Libya and Syria, the programmes had to often work with levels other than the national government, such as with UN processes. All programmes also engaged to some degree with local level formal and informal power structures. The programmes were based on and responsive to the needs of local beneficiaries, though the degree to which beneficiaries were involved differed. All programmes had to repeatedly adapt to changing contexts, ranging from changing political and security contexts to adapting to the Covid-19 pandemic.

Evaluation Question 4: What evidence is there to show the benefits of NAP programmes will be sustained beyond the life of the programme and how far, and in what ways did actions in the additional year (2020) contribute to strengthening sustainability?

The question of sustainability is somewhat skewed by the fact that with the exception of Afghanistan and Libya, and one of the consortium partners in Colombia, much of the programming has been carried over into new Dutch MFA-funded programmes. Thus, the question of whether they would have been sustainable is moot for most of the programmes. In Afghanistan, the Taliban take-over, which occurred during the evaluation phase, has also put the gains of the programme in question. In Colombia and Libya, the work of the implementing partners who did not get new Dutch MFA funding has continued, with some of the Libyan local partners now being engaged in designing a Libyan WPS NAP.

More broadly, however, the question of sustainability deserves more discussion and a differentiated conceptual approach. Programmes articulated indicators for sustainability in very different ways, and strengthening these could benefit from an open discussion between the Dutch MFA and the implementing partners as to what is realistically possible in terms of different kinds of sustainability – be it community, financial, organisational, strategic or other types of sustainability.

Evaluation Question 5: How did programme design and implementation affect achievements along the causal pathways of change, programme sustainability and, value for money? How did the dynamics between various consortium partners (including Dutch partners and local implementing partners) impact the implementation of the programme?

The ability of programmes to implement activities as had been originally intended in their design varied depending on the context. The more stable, comparatively speaking, the situation was, and the more they were able to engage with existing frameworks and structures, the more the original plans and intended pathways for change were able to be followed. The consortium model received mixed reviews from respondents involved with the programmes, ranging from celebrating the ‘richness in diversity’ to seeing them as ‘forced marriages.’ In the ideal case, consortia would be able to build on their members’ strengths and foster synergies, something which could be enhanced in the design phase. Although power imbalances are inevitable, they need to be addressed and mitigated in the consortia, as do differences of political outlook and opinion.

Questions of efficiency, value for money and sustainability were approached quite differently by the various programmes. The evaluation was able to identify several good practices in this respect, such as linking with and building on other existing structures, frameworks and partnerships, or building on synergies. However, the differences between the programmes suggest a need for a more in-depth and nuanced discussion between the Dutch MFA and on what kinds of sustainability or efficiency gains the programmes should aim for and what can be realistically expected in the given circumstances. Furthermore, programmes should consider questions of efficiency, value for money, sustainability and exit strategies more concretely already at the project design stage.

Evaluation Question 6: What lessons can be learned from the NAP programmes and how far do achievements and lessons learned align with the broader international WPS evidence base?

Many of the challenges of the Dutch NAP III-funded programmes identified in this evaluation are not unique to these, but rather widely-shared ones of in terms of WPS implementation globally. We have identified four emerging good practices from the Dutch NAP III programmes

Good practice 1: Adopting a broader and more nuanced approach to gender in WPS: The programmes across the board, but to differing degrees, expanded the scope of stakeholders when promoting WPS, but also nuanced their approach. Rather focusing on women only, the programmes actively sought to engage with men and boys on addressing gendered inequalities, and in the case of Colombia, also engaged with persons of diverse sexual orientations and gender identities. Women were also not treated as a

homogenous group, but rather differential needs and possibilities of different women based on their age, location, ethno-religious background and other factors were taken into account.

Good practice 2: Broadening the range of WPS programming approaches and themes. While the approaches used and themes the programmes engaged with are not necessarily new to the broader women's empowerment field, they do represent a significant, and in our view welcome, broadening of the WPS field.

Good practice 3: Aligning and linking WPS NAP implementation with national and local framework and processes. As discussed above under evaluation question 3, a number of the programmes were able to link their activities to other policy frameworks and strategies at various levels, helping to institutionalise them, increase the likelihood of sustainability and decrease the likelihood of duplication. This is in contrast to the often-siloed implementation of WPS programming. The Afghanistan programme was particularly notable in that it sought to simultaneously contribute to both the Dutch and Afghan WPS NAPs, a harmonisation which globally speaking tends to be an exception.

Good practice 4: Taking the need for flexibility seriously. Both the Dutch MFA and implementing partners had to show a high degree of flexibility and adaptability, not least to the Covid-19 pandemic, but also to local and national dynamics in the focus countries. Respondents from the programmes were highly appreciative of the Dutch MFA in this respect, while the programmes themselves were often able to react quickly to changing circumstances and arising opportunities, as well as re-calibrating approaches when needed.

Recommendations

The recommendations of the evaluation are grouped in three thematic areas, aimed at consortium partners and local partners on the one hand, and the Dutch MFA on the other:

Consortia and ways of working: Recommendations focus on improving consortium management, addressing internal power imbalances, investing more time into the consortium establishment phase, continuing with giving implementing partners the necessary flexibility for working in rapidly changing contexts, and encouraging reporting on when and why particular goals were not achieved.

Design, monitoring, evaluation and learning: The recommendations propose measures to address some of the shortcomings identified in the evaluation whilst retaining good practices. In terms of the design phase, this should be used for developing realistic and achievable theories of change, contextualising through improved baseline reporting, thinking through questions of sustainability and efficiency as well as do no harm. Monitoring and evaluation should be improved to address 'missing middles' and focusing more on outcomes and impact, while learning aspects should be improved.

Thematic focus areas: The recommendations call for a broader conceptualisation of the 'enhanced protection' specific objective; encouraging further work on transformative gender norm change and ensuring various levels of change are inter-linked; and assessing what is realistically possible in terms of better and meaningful participation, and continuing to create and utilise entry points at various levels as these arise.

1 Introduction

1.1 Evaluation purpose, scope, and target audience

The Taskforce on Women’s Rights and Gender Equality (TFVG) of the Dutch Ministry of Foreign Affairs (MFA) has commissioned Itad to conduct an external end-term evaluation of programmes under the Dutch Women, Peace, and Security (WPS) National Action Plan (NAP) 2016–2019 and WPS NAP 2020.¹ The 2016–2019 NAP was the third WPS NAP of the Government of the Netherlands and the NAP 2020 was a one-year extension, aiming at the implementation of United Nations (UN) Security Council Resolution 1325 and its subsequent ‘sister’ resolutions.² The objective of this end-term evaluation is to gain insight into the extent to which the main goals of the NAPs have been achieved through the eight in-country programmes, and if so, how. The eight programmes are one of the ways in which the MFA has, as a signatory, contributed to the implementation of the NAP, but they do not constitute the whole of the Dutch WPS NAP III.

The third Dutch WPS NAP 2016–2019 was drafted and signed by 64 signatories made up of the Netherlands MFA, Ministry of Defence, Ministry of Security and Justice, including the national police and Ministry of Education, Culture and Science, as well as WO=MEN, the Dutch Platform for Women Rights and Gender Equality that brings together over five dozen Dutch NGOs, individual experts and knowledge institutions. What was notable in the third NAP was the central role of Dutch civil society in co-designing, co-ordinating, and monitoring the plan through the community of NAP signatories and WO=MEN.³

Overview, rationale and objectives of the evaluation

This evaluation consisted of:

- A **meta-evaluation** of the country-programme evaluations that have been externally commissioned by the NAP consortia.
- **Additional desk research, literature review and data collection** to elaborate on and substantiate the conclusions from the country-programme evaluations that were externally commissioned by NAP consortia.
- Using **contribution analysis** to identify where programme outcomes support or challenge the theory of change (ToC) of the NAP.
- A **ToC workshop** with the implementing consortia on the ToC of the NAP.
- Two **critical reflection workshops** on the findings of the meta-evaluation and reflecting on underlying assumptions and causal pathways of the programmes.
- Three **case studies** (two country case studies and one thematic case study – see Annex 1) on programming conducted under the NAP, conducted remotely and in-country with national consultants. The two country case studies are Colombia and South Sudan, and the thematic case study focuses on the specific objective of the decrease of harmful gender norms.

Furthermore, the evaluation team conducted a validation workshop of the draft final report findings with the consortium partners and local partners, and presented the draft findings to a reference group which included the Dutch MFA and WO=MEN.⁴ The validation workshop participants were also given the opportunity to

¹ [The Netherlands National Action Plan on Women, Peace and Security 2016-2019](#)

² These include the UN Security Council Resolutions 1820 (2008), 1888 (2009), 1889 (2009), 1960 (2010), 2106 (2013), 2122 (2013), 2242 (2015), 2467 (2019), 2493 (2019) and 2538 (2020).

³ See, for example, Myrtilinen, Henri, Laura J Shepherd, and Hannah Wright (2020). *Implementing the Women, Peace and Security Agenda in the OSCE Region*. Vienna: Organisation for Security and Co-Operation in Europe and Trojanowska, Barbara K., Katrina Lee-Koo and Luke Johnson (2018). *National Action Plans on Women, Peace and Security: Eight Countries in Focus*. Canberra/Melbourne: Australian Civil-Military Centre/Monash Gender, Peace and Security Centre.

⁴ Most of the feedback in the validation workshop and the written inputs from the consortium partners and local partners pertained to minor technical clarifications. The main pushback was on the use of IoB criteria for assessing the evaluation reports of the programmes, as these criteria had not been shared at the time the evaluations were commissioned. The evaluation team however would like to stress that the use of IoB criteria is not a reflection on the quality of the programmes themselves, but was merely a tool used to identify data gaps (see also sections 2.3. and 3.1.)

provide written comments on the draft report. The feedback from this validation round was integrated into this final report.

This evaluation serves two key purposes. The first is to support the MFA and NAP consortia to **understand the impact** of the NAP programmes against their goals and if and how they have achieved them. The second is to identify **lessons** to inform future policy and programming. The findings will help inform the MFA and other ministries, as well as non-governmental organisations (NGOs) implementing WPS programmes, and contribute to global dialogue on improving WPS implementation. The primary audiences of this evaluation are the TFVG and other key stakeholders involved in implementing the Dutch NAP, including the eight consortia implementing programmes under the NAP. The report will also be made publicly available online to inform a global WPS policymaker and practitioner audience, as well as the Dutch NAP community.

The Itad evaluation team is fully independent, with no members of the team having been involved in the design, implementation, or previous reviews or evaluations of the third Dutch WPS NAP or the programmes of the eight consortia; nor have any of the evaluation team members been previously employees of the Dutch MFA or any of the implementing partners. The evaluation process was also guided by an external reference group, who were asked to give input on the draft final report. The evaluation process also involved key stakeholders at various stages to allow for a validation of interim and final findings.

The analytical approach employed by Itad for the evaluation is gender-responsive, conflict-sensitive, and attuned to the contextual differences between the countries where the different NAP programmes were implemented. The approach to gender used in the evaluation is one that is **comprehensive, intersectional, and relational**, meaning that it will take into account femininities, masculinities, and other gender identities; consider how gender interacts with other social identity markers such as socio-economic class, age, ethno-religious background, sexual orientation, disabilities, and other factors; and how gender norms, expectations, and identities are constructed in relation to one another.

Report structure

This report is structured as follows: this introductory section gives a brief contextual background on the focus countries in which the NAP III programmes were implemented, the programmes themselves and introduces the overall Theory of Change of the WPS NAP III. Section 2 gives a brief overview of the evaluation methodology, while section 3 covers the key findings of the various stages of the evaluation. Section 4 focuses on the six evaluation questions which guided this study, Section 5 provides a discussion of key themes and Section 6 concludes with key recommendations.

1.2 Contextual background

This evaluation focuses on the NAP programmes (2016-2020) that were funded by the Dutch MFA to contribute to the implementation of the third Dutch WPS NAP. The NAP programmes were implemented by consortia consisting of Dutch and local NGOs in eight countries: Afghanistan, Colombia, Democratic Republic of Congo (DRC), Iraq, Libya, South Sudan, Syria and Yemen. In each of these countries, one consortium consisting of Dutch or Netherlands-based international NGOs, Dutch knowledge institutions (e.g. in the case of Yemen) and local NGOs was funded by the MFA.

All of these eight WPS focus countries can be classified as conflict-affected, with all of them in varying states of active conflict and armed violence at the time of implementation. Most of the countries have seen decades of armed conflict and instability and there are areas in each of the countries that are not fully under the control of the central government.

State and civil society capacities vary greatly between the eight focus countries, ranging from Colombia, which has a comparatively strong state machinery and legislation as well as civil society presence more or less throughout the country, to Libya, Syria and Yemen which have areas controlled by rival powerholders and often minimal civil society space. Particular areas of some of the focus countries have very strong regional governments and active civil societies (e.g. Iraqi Kurdistan) or very large presence of UN agencies and international NGOs in addition to local civil society (e.g. North and South Kivu in DRC and parts of South Sudan).

Others have pockets where local civil society and state agencies are present in the capital and some cities (e.g. Afghanistan prior to the Taliban take-over) or extremely limited to non-existent space for civil society (e.g. Syria, with the partial exception of the Kurdish-controlled Northeast).

The possibilities for advocating for WPS or gender issues in these different contexts vary considerably, as does the respective central government's openness to engaging with them. In terms of the four WPS pillars of participation, prevention, protection, and relief and recovery, all eight countries face challenges to differing degrees. DRC and South Sudan have a comparatively high degree of openness to working on WPS in spite of entrenched patriarchal norms and persistent side-lining of women in politics and social life. Both countries have a comparatively large number of international and local NGOs working on women's empowerment and addressing gender-based violence (GBV). Conversely, in Afghanistan, Iraq, Libya, Syria, and Yemen, there are limited spaces for advocating for women's rights and a degree of government-approved (and/or donor-supported) 'state feminism', but also strong to virulent opposition to any form of increased gender equality, including from some armed opposition forces. Arguably the most extreme manifestation of this during the implementation period was the Islamic State, mainly present in Iraq and Syria, but also with branches in Afghanistan, Libya, and Yemen. Northeast Syria is an exception as the armed opposition group mainly in power during the NAP implementation period, the Kurdish *Partiya Yekîtiya Demokrat*, has explicitly promoted gender equality and women's empowerment. In Colombia, the political and social space is by far the greatest of the eight countries, and the peace agreement between the government and the *Fuerzas Armadas Revolucionarias de Colombia—Ejército del Pueblo* (FARC-EP) includes provisions on women's rights, sexual and gender diversity, as well as indigenous and Afro-Colombian rights, and the state machinery is at least on paper committed to promoting gender equality and civil society participation. Nonetheless, even in Colombia, the risks for women's and lesbian, gay, bisexual, trans, intersex, queer and other (LGBTIQ+) rights activists as well as others have increased over the years, with death threats and assassinations increasing after the peace accord.

Of the eight focus countries, Afghanistan, DRC, Iraq, and South Sudan had their own WPS NAP in place during the implementation period, and the Dutch WPS NAP programme in Afghanistan sought to support its local implementation. In Yemen, a NAP was adopted at the end of 2019 for the period 2020–2022, i.e. towards the very end of the third Dutch NAP.

All of the eight programmes were heavily impacted by the outbreak of the Covid-19 pandemic, and all had to one degree or another limit their activities in the face of restrictions on movement and gatherings. The possibilities of shifting planned in-person activities to online or telephone-based activities varied greatly between the countries as well as within the countries, depending on the existence and access to internet and mobile phone networks as well as electricity.

1.3 Programme background

The third NAP was 'outward-looking,' i.e. focused largely on WPS implementation outside of the Netherlands, with eight conflict-affected focus countries. As stated in NAP III, these eight were chosen based on the following criteria, though not all of these needed to be met:

- (1) *Countries that are in conflict or fragile states.*
- (2) *Countries that are the focus of a Dutch policy.*
- (3) *Countries in which signatories have sufficient capacity, local partners and a relevant track record.*
- (4) *Countries in which the Netherlands participates in a multilateral civil and/or military mission.*

It should, however, be kept in mind that the eight in-country programmes did not represent the whole of the Dutch WPS NAP III⁵, but rather were only a part of the activities and measures implemented by the Dutch government and civil society.

⁵ The NAP III programme also involved activities conducted by Dutch civil society organisations in the Netherlands as well as advocacy and promotion of WPS on international platforms.

A mid-term review (MTR) of the third NAP was finalised in 2019⁶ and while the fourth WPS NAP was developed in 2020, the eight NAP programmes continued with no-cost extensions into 2020 and/ or with additional funding (see table 1). Among the findings of the MTR were a recognition that the NAP was underpinned by a strong ToC (see also section 1.4) and that the MFA worked collaboratively with civil society. However, the MTR also identified challenges in working with a wide network of stakeholders that could have been more effectively co-ordinated. Notably, the MTR pointed out that while the consortia implementing the eight programmes that funded under, and formed part of, the WPS NAP III, had their own respective monitoring and evaluation frameworks, the WPS NAP as a whole did not have an over-arching one.

In addition to the consortia, a Country Group was formed for every focus country, to which all signatories to the NAP were invited. The exception was Afghanistan, where no new country group was established, but WPS and gender issues were a regular discussion item in the pre-existing Dutch NGO co-ordination group.⁷ These country groups were mostly chaired by consortium leads, though some were chaired by other consortium members or by NAP signatory organisations which were not part of the consortium. As noted by the MTR, their degree of activity varied greatly.⁸

The NAP consortia and programmes were, in brief:

Afghanistan: Safhe Jaded – Implementing the Afghan NAP 1325 by Linking Inclusive Security and Justice

The ‘Safhe Jaded (New Page)’ consortium was headed by Oxfam Novib and included Cordaid and the Afghan Women’s Educational Center (AWEC) as consortium partners and additionally Hamida Barmaki Organization for the Rule of Law (HBORL) and Women and Peace Studies Organization (WPSO) as implementing partners. The consortium focus was on the improved implementation of the participation and protection pillars of the Afghan WPS NAP by civil society and state actors, in particular in terms of the justice and security sectors, at the national and provincial levels. The engagement with the justice sector included traditional justice providers. The programme was implemented in Kabul and in six provinces.⁹

Colombia: Women as Central Agents for Peacebuilding in Colombia

‘Women as Central Agents for Peacebuilding’ was implemented by the Empowered Women Building Peace consortium, consisting of the Dutch NGOs the Dutch NGOs ICCO, Health Net TPO, and Mensen met een Missie and their Colombian partners Corporation Tamar, Red Tamar, Corporation Amiga Joven, Católicas por el Derecho a Decidir, Rempaz, Fundación Mencoldes, Ruta Pacifica de las Mujeres, Liga Internacional de Mujeres por la Paz y la Libertad (LIMPAL), Corporación Espacios de Mujer, and Corporación de Apoyo a Comunidades Populares (CODACOP). The programme focused heavily on capacity-building, strengthening of civil society networks, advocacy, participation, and monitoring of the peace process and GBV prevention, as well as changing harmful gender norms. The activities were implemented at the national, departmental, and municipal level in Bogotá and 11 other conflict-affected departments.¹⁰

DRC: Mwanamke, Amani na Usalama – Women Peace and Security

The ‘Mwanamke, Amani na Usalama’ programme was implemented in the eastern provinces of North and South Kivu in DRC by a consortium led by Mensen met een Missie with CARE Nederland and Tosangana. Local implementing partners included several women’s organisations (Le Directoire des Organisations Féminines pour les Actions de Paix – DIOFAP, Sautiya Mama Mukongomani – SMM, and Synergie des Femmes pour la Paix et la Réconciliation des Peuples des Grands Lacs – SPR), a youth organisation (Le Parlement des Enfants et des Jeunes – PARDE), and the Diocesan Commission “Justice et Paix” (CDJP) from Goma, Bukavu, Uvira and Butembo of the Catholic Church. In addition to the Kivus-based NGOs, the implementing partners also included two advocacy associations in Kinshasa, Femmes en Action and Voix des Faibles. The key areas of work

⁶ Herweijer, Rosien (2019). *Mid Term Review National Action Plan on Women Peace and Security 2016 - 2019*

⁷ Ibid. p. 24

⁸ Ibid. p. 18

⁹ Herat, Balkh, Parwan, Nangarhar, Daikundi, and Paktia Provinces

¹⁰ Antioquia, Bolivar, Cauca, Chocó, Cundinamarca, Meta, Putumayo, Santander, Sucre, Tolima, and Valle del Cauca

of the programme were GBV prevention, documentation, and response, changing harmful gender norms, supporting women's leadership and participation in peacebuilding and recovery efforts, as well as national-level advocacy.

Iraq: Engendering the Transition to Peace and Security in Iraq

The 'Engendering the Transition to Peace and Security in Iraq' programme was implemented by PAX and Impunity Watch as the Dutch consortium partners, as well as the Iraqi Al Amal Association and its local implementing partners. Its key areas of work were: GBV awareness-raising, documentation, prevention, and response; integrating gender perspectives and promoting women's active participation in security sector reform (SSR); justice and reconciliation processes; as well as shifting gender norms around women's participation. It was implemented at the national level, in Baghdad, Basra, Kirkuk and Salah ed-Din.

Libya: Women and Youth as Bridgebuilders – Strengthening Resilience in Libya

The 'Women and Youth as Bridgebuilders' programme in Libya was implemented by a consortium headed by Cordaid as well as Human Security Collective and Women Peacemakers Program on the Dutch side and eight Libyan implementing partners: Makers of Hope, Together We Build it, Tamazight Women Movement, Misrata: Dialogue and Debate Association, Goodness Brought us Together, Al Nour, Fezzan Libya Organization, and I am Libyan, But My Child is Not. The Women Peacemakers Program dropped out as a partner after it ceased operations at the end of 2017. The programme aimed at fostering women's participation and empowerment in Libya, as well as building the capacity of and supporting NGOs and activists to engage on human security, in peacebuilding processes, and on WPS. The programme included trainings on GBV as well as components on transforming harmful gender norms. The implementing partners worked in Tripoli, North/North-West Libya, Misrata, Al Bayda, Ghat, and Sabha.

South Sudan: Women and Girls for Change

The WPS programme in South Sudan's was implemented by the 'Women and Girls for Change Alliance' (WG4CA) headed by Plan Nederland as well as, PAX, HealthNet TPO and Support Trust for Africa Development (STAD), with local implementing partners AMA and EWO. Activities focused on improved protection of women and girls from GBV, including improved support services and engaging with traditional courts; work with men, women, boys and girls on transforming harmful gender norms; and improving women's participation in decision making around conflict prevention and resolution as well as peacebuilding. It was implemented in Eastern Equatoria, Lakes, and Unity States.

Syria: Syrian Women's Participation in Peace and Security – The Way Forward

The consortium implementing the 'Syrian Women's Participation in Peace and Security – The Way Forward' programme in Syria was led by Hivos and included the Centre for Civil Society (CCSD) and Kafa with Syrian Women for Peace and Syrian League for Citizenship as implementing partners. It aimed at increasing women's participation in the peace process and in local-level decision making, linking local-level participation with the national-level peace process, shifting inequitable gender norms and raising awareness of women's and girls' protection issues.

Yemen: Women Peace and Security in Yemen – From the Ground Up

The consortium implementing the 'Women Peace and Security in Yemen – From the Ground Up' programme was headed by CARE Netherlands with the Women Peacemakers Program, The Hague Academy for Local Governance (THA), and the local partners CARE Yemen and Youth Leadership Development Foundation (YLDF). As in the case of Libya, the Women Peacemakers Program dropped out after suspending operations in 2017. The programme was implemented mainly in the Sana'a and Taiz governorates; it aimed at increasing women's political participation at the local level, including in peacebuilding, changing harmful gender norms, and building local NGO capacity.

As noted by the MTR, none of the programmes were implemented by a Dutch women's rights organisation and while one (Iraq) was led by a Dutch peacebuilding organisation, seven of the eight consortia were led by large Dutch NGOs working more broadly on humanitarian aid and development co-operation.

The programmes were all granted either a no-cost extension and/or additional funding for 2020 after the end of the original implementation period from 2016–2019 (see Table 1). One of the aims of this additional year of funding was to allow for the programmes to improve their sustainability by consolidating and building upon results achieved in the period 2016-2019. However, these plans were complicated by the Covid-19 pandemic, as discussed further below.

Table 1: No-cost extensions and additional funding for 2020

Country programme	Received no cost extension 2020	Received additional funding 2020
Afghanistan	X	
Colombia		X
DRC		X
Libya	X	X
Iraq	X	
South Sudan	X	X
Syria		X
Yemen	X	X

1.4 NAP theory of change

The Dutch WPS NAP III had an overarching ToC (Figure 1), but beyond this there was no overall monitoring, evaluation, and learning (MEL) plan for the NAP. All consortia implementing NAP programmes funded by the MFA had their own, contextualised ToC that was based on the overall ToC of the NAP. The overall objective of the ToC was:

“Together we contribute to an enabling environment for women’s participation and empowerment in conflict and post-conflict environments, so they can meaningfully participate in conflict prevention, resolution, peacebuilding, protection, relief and recovery.”

This objective was to be reached by work on the following specific objectives:

- **Enhanced protection** – Better protect women and girls in conflict and post-conflict situations from violence and violations of their rights.
- **Decrease of harmful gender norms** – Subvert harmful underlying gender norms, which are obstacles to sustainable peace.
- **Equal leverage in conflict prevention, resolution, peacebuilding, relief and recovery** – Ensure that women have equal leverage in conflict prevention and resolution, peacebuilding, relief and recovery at all levels, and that their efforts are acknowledged and supported.

These were to be achieved through three parallel approaches (‘pathways of change’):

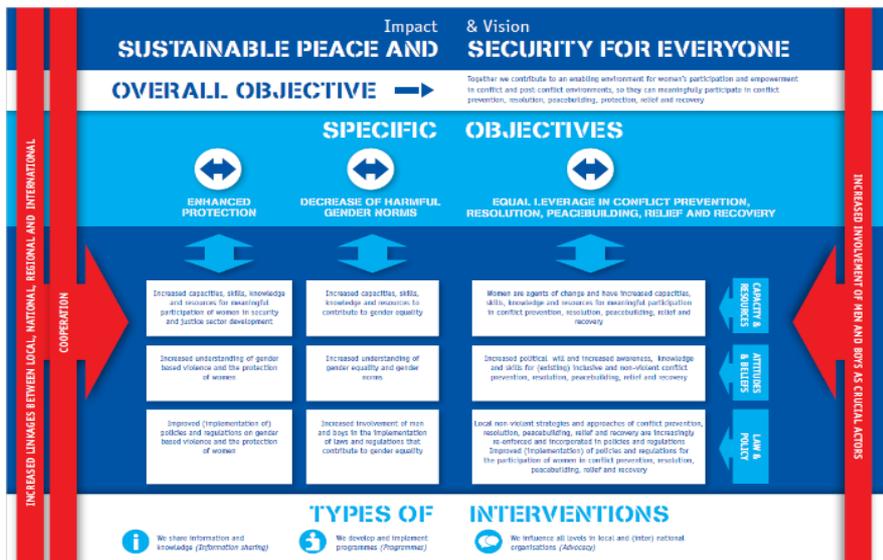
- i. **Capacity building and resources,**
- ii. **Changing attitudes and beliefs, and**
- iii. **Laws and policies.**

Two further overarching ways of enhancing positive change are cross-cutting:

1. **Increasing linkages between local, national, regional, and international co-operation**
2. **Increased engagement with men and boys.**

This overall ToC was used as a basis for the eight in-country programmes, who designed their activities and own ToCs based on it.

Figure 1: NAP 2016-2019 Theory of change



2 Methodology

This evaluation was conducted by a team of three lead evaluators, two in-country researchers and key support staff from Itad, in July-October 2021. More information on the data collection tools, case study approach and sampling can be found in Annex 2.

2.1 Evaluation design

Our methodology for this evaluation consisted of three modules: 1) quick-scan 2) meta-evaluation and 3) case studies as shown in Figure 2. The evaluation questions were addressed across all three modules of the evaluation with increasing depth of analysis.

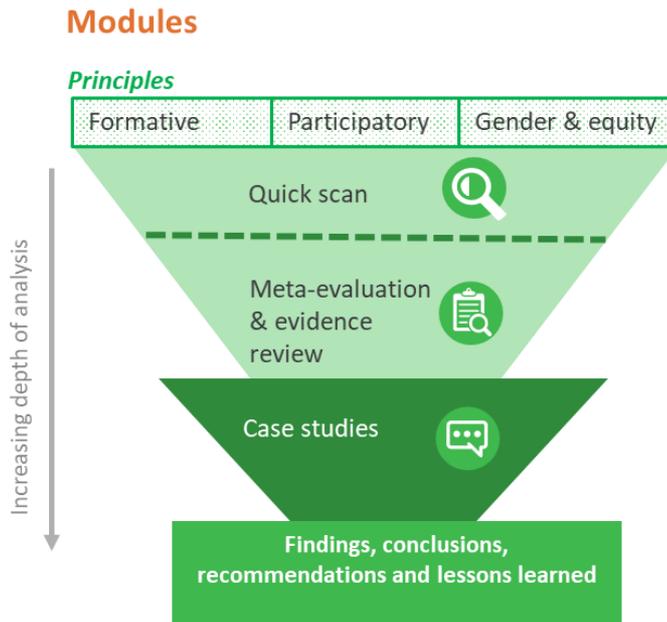


Figure 2: Evaluation Modules

2.2 Modules 1 and 2: A theory-based approach

Our analytical framework draws on a mixed-methods theory-based approach. This means we used the ToC, the causal pathways (between activities, outputs, and outcomes), and the assumptions as a map against which we were to collate and assess evidence. In order to do this, we used contribution analysis¹¹ as it enables a systematic and robust way to explore each programme contribution to observed impact against the three specific objectives of the NAP 2016–2019 and 2020 ToC.

Contribution analysis does not traditionally involve unintended outcomes outside the ToC, therefore, we adapted the approach to also gather both positive and negative unintended outcomes. In order to further unpack *how* change happened, we introduced process evaluation components¹², which helped us gauge what planning and actions took place to secure the sustainability of programmes. Our methodology followed these steps:

¹¹ Contribution Analysis is an approach for assessing causal questions and inferring causality in real-life program evaluations. It offers a step-by-step approach designed to help managers, researchers, and policymakers arrive at conclusions about the contribution their program has made (or is currently making) to particular outcomes.

¹² Process evaluation methods assess how well programme plans were put into action. Process evaluation components enabled us to address other evaluation criteria – for example process questions on programme design to gauge what planning and actions took place to secure programme sustainability.

Module 1: Quick-scan analysis

As part of the inception phase, we conducted a quick-scan of seven country programme evaluations and corresponding terms of reference, where available.¹³ The programme evaluations were judged against the 26 IOB evaluation criteria,¹⁴ which were used to assess the quality and reliability of the existing evaluation material and identify weaknesses and strengths in the evaluations. This fed into our decision making for the selection of case-study focus countries and evaluation design, explained in more detail in section 2.3 as well as Annex 2. After this, we followed the steps of contribution analysis, though necessarily we had to make some adjustments to this process due to the limited nature of the documentary evidence available. The adjusted process is outlined below.

Module 2: Meta-evaluation and contribution analysis

i) Validation of the NAP III ToC

We validated the ToC through a remote workshop with representatives from across NAP III programme Dutch consortia partners. The workshop discussed how the pathways worked in reality, unpacked underlying assumptions, and assessed which pathways were priorities to generate a more elaborated ToC. As the discussion unfolded contributions from different country programme representatives were captured in real-time using an interactive whiteboard against the ToC. The whiteboard is depicted below in Figure 4 and Annex 5.

The discussion from the ToC workshop then fed into our data collection tools for the document review. For example, economic empowerment and GBV service delivery were highlighted as specific programme themes that were not specified in the ToC but key to programme interventions. We therefore added this to our document review matrix to ensure that we captured evidence on it.

ii) Gather existing evidence from the country-programme evaluations

Building upon the quick-scan analysis, we mapped evidence from seven country-programme evaluations against the ToC and collated unintended outcomes. From the quick-scan, we were aware that the quality of evidence from the evaluations was weak and therefore required validation and triangulation with programme documentation.

iii) Gather existing evidence from the programme documentation

Alongside mapping evaluation findings, we also conducted a desk review to gather and analyse evidence from the programme documents. The data was mapped within an evaluation assessment framework against the ToC and evaluation questions. The evaluation assessment framework can be found in Annex 3.

iv) Assemble and assess the contribution story using the existing evidence

Existing evidence from the above three steps revealed that there was not sufficient data of results against the NAP ToC outcomes to assemble a robust contribution story and, that the large variety of interventions across different country programmes meant that the potential pathways of change were profuse and patchy.

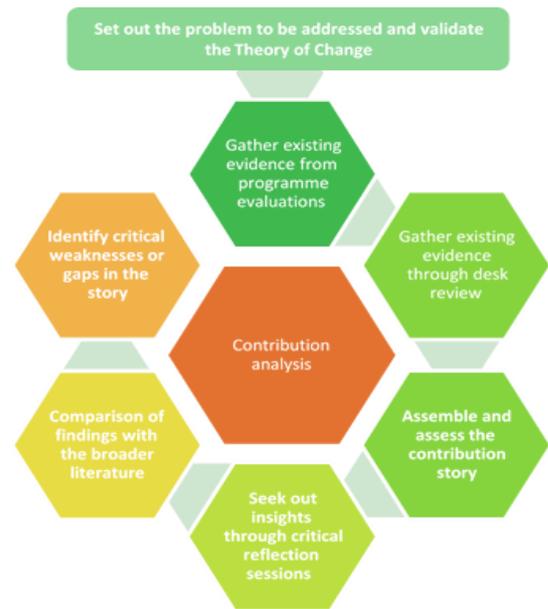


Figure 3: Contribution analysis steps

¹³ The Afghanistan programme evaluation was still in progress whilst we conducted the quick-scan. Therefore, the assessment was done on a draft version as of 14/07/2021. The Syria programme was not included in the quick-scan as there is no external evaluation available.

¹⁴ IOB Evaluation quality criteria 2020.

Therefore, instead of assembling a contribution story against which to gather additional case study evidence, we used existing evidence to identify a ‘missing middle’ within the ToC (see Figure 4 below and Annex 5).

This ‘missing middle’ is a set of intermediate outcomes that although not explicitly articulated in the ToC were in part identified by the evaluation team through the document review, and further elaborated by consortia members during the ToC and critical reflection workshops. They were seen by implementing partners as important steps towards achieving outcomes which went beyond their successful achievement of change at the output level but were not otherwise well captured within the ToC¹⁵. We then recorded evidence gaps in line with this and focused on synthesising and comparing findings across country programmes.

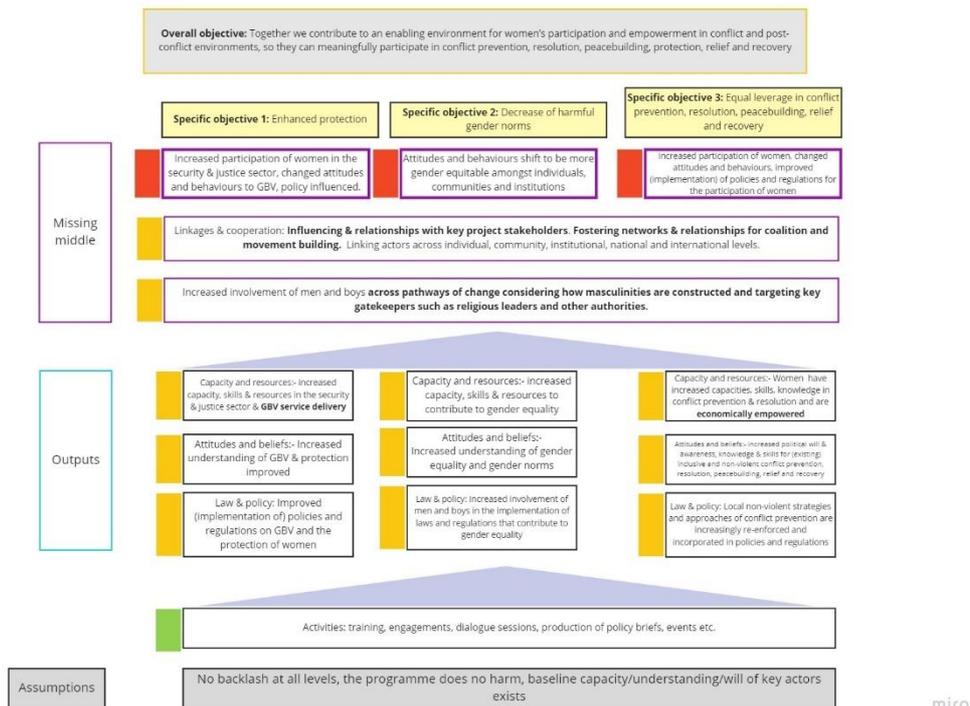


Figure 4: Theory of Change workshop feedback captured in an interactive whiteboard

v) Seek out insights through critical reflection sessions

The purpose of the critical reflection sessions was to create a safe space for key stakeholders to honestly reflect on the contribution story and aggregated findings to confirm or challenge them. As we were unable to construct a contribution story,¹⁶ we instead shared our key findings so far with a selection of both international and national members of the consortia. We then asked for participant reflections before conducting break-out groups per country to discuss key questions and ideas coming out of the document review thus far. These questions are available in Annex 2.

vi) Comparison of findings with the broader literature

At this stage we planned to compare synthesised NAP programme findings against wider WPS literature on ‘what works’ and what does not. This aimed to position learning from MFA NAP programmes within international research to understand where there are resonances or unique findings. Given the paucity of

¹⁵ In order to avoid the use of complex MEL jargon during workshops with partners it was agreed we would refer to these intermediate outcomes as the ‘missing middle’ – a concept which resonated very well with consortia representatives.

¹⁶ We were unable to construct a contribution story as there was an absence or weakness of data against some components of the causal pathways and a ‘missing middle’ of evidence, therefore creating a leap between activities and outcomes.

robust outcome-level evidence, we delayed this activity to run in conjunction with the case studies in order to increase its utility and respond to key findings coming out of the primary data collection.

vii) Identify critical weaknesses or gaps in the story and/or particular findings of interest

We reviewed findings from the ToC workshop, document review and critical reflection sessions together to guide the design of case study data collection in two focus countries and one thematic case study. The absence or weakness of data against some components of the causal pathways in the ToC – particularly at the outcome level – presented considerable challenges for our ability to robustly assess all the pathways through the case study data collection. We did however assess that case studies could still be used to understand how interventions have contributed to changes outlined in the ToC where possible, and in addition, collect evidence against the ‘missing middle’ identified i.e. intermediate outcomes which may not be explicit in the three ToC pathways, as well as understand more about programmatic and contextual assumptions within the current causal pathways.

2.3 Module 3: Case studies

The country case studies were chosen following the quick-scan review during the inception phase. As part of the quick-scan, we mapped the country programmes by the strength of evidence expressed by the quality and credibility of the reports of the externally-commissioned country programme evaluations (see Table 2) as well as theme and contextual considerations in order to select countries for the case study focus (see full table at the end of Annex 1).

Table 2: Programme evaluations assessed against the IOB criteria

Country	Score out of 26 IOB criteria ¹⁷	Key gaps
Afghanistan	15	Use of Outcome Harvesting; ¹⁸ unclear conclusions and recommendations
Colombia	12	Focus on activities and outputs, not impact; no triangulation of findings
DRC	13	Evaluation team not fully independent, and though data had been collected from third parties, there was little triangulation evident in the report
Iraq	9	Unclear ToC; lack of transparency on data collection
Libya	14	Lack of independence of evaluation; no triangulation of findings
South Sudan	18	Lack of baseline data; use of Outcome Harvesting; incomplete data collection
Yemen	12	Unclear ToC; no conclusions in evaluation report

It is important to note that the evaluation team are of the view that Outcome Harvesting is an extremely useful and widely used method for monitoring progress in conflict-affected settings where pathways to change are often unpredictable. However, this is not considered a sufficiently robust methodology by the MFA to yield reliable findings. As noted in section “2.5 limitations and challenges”, the lack of rigour in the country programme evaluations poses limitations to the strength of findings and conclusions this evaluation has been

¹⁷ Number of criteria that were rated good or satisfactory.

¹⁸ The IOB recommends explicitly against the use of Outcome Harvesting as an independent, external evaluation method. The IOB states that ‘in practice, this method is not appropriate to evaluate effectiveness and unable to validly establish the contribution of interventions to observed outcomes. In addition, the method is not in the spirit of several other evaluation quality criteria, specifically regarding the independence of evaluators (criterion 2), sufficient independent sources (criterion 17), triangulation (criterion 18), and avoidance of bias (criterion 19)’. However, it is worth noting that this is referring to external evaluations, rather than internal programme monitoring and evidence collation.

able to draw from them, nevertheless, the evaluation team consider the country programme evaluations to have enough valuable information and learning to warrant their use when triangulated with other data.

As per our inception report, the case study selection was based on the following criteria:

- Which countries have the weakest data (as ranked in the quick-scan)
- Countries where new programmes are being implemented from 2021 onwards
- Geographical spread of country case studies
- Thematic spread of country case studies
- Possibility to engage stakeholders given the country context and risk

Considering this, we proposed to conduct **two country case studies** in South Sudan and Colombia. This was agreed in consultation with the MFA. South Sudan ranked the highest in the quick-scan in terms of the IoB criteria and Colombia amongst the lowest. We elected high and low scoring programmes to enable a contrast of programme efficiencies around data and evidence. These countries were also chosen as we considered them low risk for data collection in-country and both programmes had continued Dutch funding after 2020 (2021-2025). They also provided a geographical and thematic spread.

As a third case study, we proposed a **thematic case study** on what worked in terms of shifting harmful social norms that underpin gender inequality, which was an element in six of the country-level interventions and is one of the goals of the NAP – to decrease harmful gender norms. Of the three goals of NAP III, this is arguably the one where globally the understanding of ‘what works’ is still evolving the most and thus the opportunity for learning and contributing to broader debates is the greatest.¹⁹ This case study focused on data collection from the programme level and involved Dutch NGOs that led consortia in the focus countries where this was an area of focus. The thematic case study relied on existing evidence on social norm change interventions, the data within the available documentation, and key informant interviews (KIIs) with representatives of the Dutch NGOs involved in the WPS NAP.

We originally planned to utilise an outcome case study methodology, however, the document review did not find sufficient evidence on outcomes as they are articulated in the ToC to do this. Our understanding from both documentation and conversations with consortia representatives was that this is the result of two key factors. Firstly, the sense that the NAP ToC outcomes were overly-ambitious and therefore unlikely to be achievable during the programme timeframe, and secondly because consortia were not explicitly asked to report against them, instead having their own individual ToCs to report against. Therefore, we adapted the case study methodology in co-ordination with the MFA to gather more information through primary data collection on intermediate results across the three strategic outcomes of the ToC.

2.4 Sampling

Throughout the evaluation we took a purposive and snowball sampling approach to select respondents for data collection at each stage. Participants for the ToC workshop and critical reflection session were self-selecting. See Annex 2 for a full description of our sampling approach.

2.5 Limitations and challenges

The evaluation has **inherent limitations regarding the quality of the country programme evaluations and country-level reporting**. The quick-scan revealed that none of the country programme evaluations met the MFA IoB quality criteria (see table 2) and therefore, the data from these evaluations was not sufficiently robust to draw concrete findings about the extent to which programmes had achieved their stated outcomes. It was felt that despite the lack of rigour, the country programme evaluations still held valuable findings and learning

¹⁹ While there is a growing literature on gender norm change, the evidence base is less developed than for increasing women’s meaningful political, social and economic participation; or the extensive evidence base on ‘what works’ on for example gender-based violence prevention

for the MFA and wider WPS programming and it was therefore decided to continue with the evaluation to attempt to triangulate some of these findings, despite these limitations.

Also, as discussed above, the paucity of evidence around outcome-level change from the documents available to the evaluation team meant we were required to adapt our approach in order to maximise the utility of any data collection. Interviews with consortium partners and local partners and discussions with them in the workshops further highlighted gaps in the reporting, including at the activity level. This makes it difficult to robustly assess impact and change along all of the pathways. We have nonetheless sought, with the time and resources available, to fill as many of the gaps as possible and trace outcomes and change to the degree possible. There was also inconsistency in the way budgets were reported which meant it was not possible to conduct a comparative budget analysis across programmes, we have, however, drawn some analysis from the data available.

The evaluation team is confident that this has not affected the quality of the evidence this evaluation has produced against change at the intermediate outcome level; the triangulation of sources has enabled us to ground our findings in sufficiently credible evidence. However, rather it has served to highlight the trade-offs which exist between a necessarily broad ToC, and the need to robustly generate evidence of programme effectiveness in achieving an ambitious set of outcomes.

The document review did not find sufficient evidence on outcomes ... which has served to highlight the trade-offs which exist between a necessarily broad ToC, and the need to robustly generate evidence of programme effectiveness in achieving an ambitious set of outcomes.

The lack of independent data sources has also been a limitation. Whilst the evaluation team has endeavoured to speak to as many relevant participants as possible, we are aware there are some gaps in participation and primary data was mostly collected with programme implementers themselves. This is partly because the evaluation team did not feel that it was appropriate to interview some beneficiaries for case study KIIs as some have already participated in the previously conducted country programme evaluations and the risks do not outweigh the benefit of the data that would have been captured through further KIIs. For instance, it might be an inconvenience or disrupt livelihoods of beneficiaries to participate in meetings for a programme they are no longer benefitting from, and asking beneficiaries to respond again to similar questions as during previous internal and external monitoring and evaluation processes risks contributing to research fatigue.

Gaps in participation from implementing partners are largely a result of significant political upheaval in particular contexts, due to other work commitments, because the organisation was no longer part of a NAP-implementing consortium or due to staff turnover. A number of key personnel who had been involved were no longer present and available for interviews or workshops, while their replacements did not always have full knowledge of the implementation phase. Furthermore, even with key staff members who did have the relevant institutional knowledge, often availability and at times access to communication was an issue.

Afghanistan presented a particular challenge, as midway through the evaluation process the elected government collapsed and the Taliban took control of the country, precipitating a mass evacuation from the country, especially of people at risk from the new regime. This meant that Afghan implementing partners of the NAP were no longer able to participate in the evaluation process, nor were Dutch consortium members who were heavily involved in the evacuation and response.

The scope of the country case studies was by necessity limited by several factors, in particular safety/security, availability and willingness of interviewees to engage, availability of data, and time and resources. In terms of security and Covid-19 risks of carrying out the country case studies in Colombia and South Sudan, risk assessments were undertaken by Itad's Global Safety and Security Team with input from the consultants. The evaluators adhered to all national Covid-19 regulations and global good practice, and conducted online interviews where visits were not possible due to limitations of time or Covid-19 restrictions. For instance, in Colombia, it is the norm to carry out meetings virtually, so meetings were held online for KIIs.

In South Sudan, it was only possible to visit one site outside of Juba, the capital city, due to the time required to travel to the communities where the programme activities were implemented.

In terms of availability and willingness to be interviewed, the in-country researchers for the case studies faced challenges in securing KIIs with representatives from local NGOs that have left the programme or where there have been changes in staff. Nonetheless, every effort was made to identify participants for KIIs who could provide meaningful information for the case studies, and reducing selection bias to the degree possible.

We also recognise that there may be some sampling bias. We may have encountered desirability bias²⁰ amongst respondents, especially if they are still involved in consortia implementing successors of the NAP programmes within the context of the WPS instrument under the Strengthening Civil Society subsidy framework. We mitigated this as much as possible by making it clear that responses will not affect their role in these programmes and indicated that no person would be linked by name or position to specific opinions or statements in the final report. As per best practice, such approaches serve to reduce, although not completely eliminate desirability bias. Selection bias was also a potential risk for the interviews and online workshops, as invitations to join would likely not have reached individuals or organisations disengaged with the consortia. Although interviews and interactions during the online workshops were frank and open, there is always the possibility that participants did not want to disclose negative experiences with the programme but, as an evaluation team, we see this risk as small. Snowball sampling in the case study research may also have resulted in selection bias, however, this is the most appropriate way to find respondents given the very short data collection period and resource constraints.

²⁰ Desirability bias is the tendency of research or evaluation respondents to respond to questions in a manner that will be viewed favourably by others. It can take the form of over-reporting good results or under-reporting negative or undesirable results.

3. Findings

This section presents the key findings from each step in the evaluation process.



Figure 5: Steps in the evaluation

3.1. Inception and quick-scan

During the inception phase, we conducted a quick-scan of seven country-programme evaluations²¹ and a light-touch review of key documents in order to refine our approach to the evaluation. The quick-scan assessed the quality and reliability of the existing evaluation material by reviewing the evaluations against 26 IOB (2020) evaluation criteria.²² The IOB criteria are used by the MFA as a guide for evaluation good practice.

It is worth noting that it was not a formal requirement for partners to fulfil the IOB criteria in the country programme evaluations, and that the IOB criteria were shared with them at a relatively late stage, and not all evaluation teams were aware of these criteria. The assessment is therefore not an accountability exercise. Instead, it informed the evaluation team that, as a starting-point for the meta-evaluation, the overall evidence available was weak and provides some interesting learning points on how to strengthen the quality of evaluations (see also section 5).

Key finding 1: Whilst the country evaluations were rich in information on activities, and sometimes anecdotal outputs and outcomes, evidence was generally weak (not validated or triangulated), and focused on activities and outputs rather than outcomes and impact. The methodologies used were not rigorous, making it difficult to provide results, well-founded conclusions and draw lessons from them.

Furthermore, activity reports and evaluations of the country programmes did not always capture the full range of programme activities implemented by partners.

In assessing the country evaluations, we found that all of them scored poorly against the IOB criteria. The evaluations of the WPS interventions in South Sudan, Afghanistan, and Libya were the highest ranking, with Iraq scoring lowest based on the IOB criteria.

None of the evaluations scored ‘good’ or ‘sufficient’ across 23 of the 26 criteria, which is required by the IOB to consider the evaluation to be of good standing. The evaluations lost marks across the key criteria including:

Quality control and description of the background and objective of the evaluation

Five of the seven programme evaluations²³ did not fulfil the IOB quality criteria to have independent evaluators, an external reference group, and were missing validation workshops in the evaluation process. Whilst most of the evaluations sufficiently described the objective of the evaluation and the context,²⁴ only the South Sudan and DRC evaluations sufficiently discussed the programme ToC and validated its assumptions.

Methodology of the evaluations

Despite all evaluations including suitable evaluation questions, the methods selected to determine if the country-programmes were effective lacked rigour. IOB standards require a theory-based approach using the

²¹ The Afghanistan programme evaluation was still in progress whilst we conducted the quick-scan. Therefore, the assessment was done on a draft version as of 14/07/2021. A final version of the evaluation was shared in August and was reviewed as part of the document analysis.

²² IOB Evaluation quality criteria 2020.

²³ Iraq, South Sudan, Libya, Iraq and DRC programme evaluations did not fulfil formal requirements

²⁴ All except the Iraq evaluation.

ToC as the evaluative framework yet, none of the evaluations used appropriate indicators to explore the ToC pathways of change. Moreover, sufficient independent evidence sources and triangulation of data proved weak across the evaluations and generally poorly referenced.

Subsequent interviews and discussions with consortia members in the critical reflection workshops also highlighted that often a range of activities or impacts had not sufficiently been covered by the evaluations. In at least one case, one implementing partner felt so under-represented in the country programme's mid-term evaluation that they commissioned a separate external evaluation of their own programme component.

Conclusions and recommendations

The conclusions drawn from the evaluations were of varying quality, ranging from insufficient when conclusions were entirely absent (Iraq and Yemen) to sufficient (Afghanistan, Colombia, and DRC) to good (Libya and South Sudan). Similarly, the recommendations in the evaluations were of variable quality.

The quick-scan was useful in highlighting early the low-quality evidence available which enabled us to adapt our methodological approach appropriately to focus less on outcomes (as there were no well-evidenced outcomes available) and gather information at output and intermediate outcome level, as well as explore country programme M&E systems that underpin the issues with data quality. Therefore despite the lack of robustness and subsequent limitations to using country programme evaluation data, it was warranted worthwhile to move forward drawing out key learnings and triangulating evidence where possible through the document review and primary data collection.

3.2. Theory of Change workshop

Key finding 2: Implementing partners considered the ToC broad and flexible enough to cover a variety of different approaches taken by the different country programmes; however, it was thought to be overly ambitious in its objectives and the logical pathways to change were found not well expounded.

After completing the quick-scan, the evaluation team organised an online ToC workshop with representatives from all eight country programmes, including staff both from Dutch and implementation country NGOs. The aim of the 1.5-hour workshop was to examine the overall ToC of the WPS NAP 2016-2019 in detail, discuss in how far it guided the respective consortium-level theories of change of the programmes, how the pathways of change worked in reality and reflect on some of the underlying assumptions.

The consensus among the participants of the workshop was that the overall ToC had been useful in giving guidance when designing the respective country-level theories of change, had reflected realities and was broad enough to be able to design programmes for very different country contexts and to allow for the flexibility that implementation at the country level required. However, as beneficial as this broad approach is in terms of allowing for a wide spectrum of activities, it did have the drawback that comparisons between programmes were difficult. The linking of the work on the different specific objectives and across their three respective pathways of change was furthermore complicated by the fact that the activities were in part carried out by different partners within the programme, focused on different actors or regions and were in many instances siloed rather than actively seeking complementarity. Nonetheless, there were also numerous cases where implementing partners complemented each other and built on synergies through their different approaches.

For the most part, the ToC was only used in the initial planning and design phase, and was not something that was revisited on a regular basis. There was also consensus that the ToC was too ambitious to achieve the objectives in the lifetime of the country programmes. For example, one participant highlighted that under the 'decrease of harmful norms' objective, the timeframe only allowed shifts at an individual level, not at community or institutional levels. Further, another participant noted that the first two years of their country programme was relationship building, an essential prerequisite to any change.

From the discussion it became apparent that country programmes differed in several ways: i) which pathways they prioritised; ii) the ways in which they interpreted the NAP III ToC into country-level programme theories of change and; iii) the approaches used to implement the country-level theory. Whilst broadly participants felt that the three over-arching specific objectives covered their work, it was noted that some vital components underneath these objectives were not well expounded. For example, economic empowerment was seen as an important precondition to some of the other pathways (such as women’s ability to participate in peacebuilding); GBV service delivery strengthening, working broadly with civil servants and policymakers and addressing harmful masculinities were also mentioned as outputs participants felt were important but not reflected in the ToC.

One of the analytical tools we use in this evaluation is that of a ‘missing middle’ (see also section 2.2.), which we will re-visit in different sections of this report. In terms of the overall ToC and the country-level programme theories of change, this missing middle manifested itself in the gap between the ambitious goals of the overall objective and specific objectives, and actual implemented activities of the programmes, which were by necessity much smaller in scale and scope.

While it is not uncommon for programmes to aspire to ambitious overall objectives, it is important to have clear limits of accountability indicating what outcomes the implementing partner can influence, realistically achieve, monitor and be held accountable for – and what lies beyond. Bridging the missing middle requires mapping concretely what in-between steps and pre-conditions are necessary for the activities (e.g. a workshop for local police explaining the law on preventing domestic violence and a public awareness campaign on GBV) to contribute to achieving the stated objective (e.g. women in the community are better protected from GBV). As was evident from the discussions with consortium partners and local partners as well as from the ToC and critical reflection workshops with them (see section 3.4.), they were very well aware of the missing middle-gap between what they could do and achieve through their activities and the much more ambitious objectives.

Underlying assumptions were also discussed in the ToC workshop. As country programmes were thematically widely varied, there were more specific assumptions at country programme level. At the NAP ToC level, the ToC is not articulated enough to allow assumptions along the pathways of change, for example, between output outcome change, yet some common assumptions could be aggregated. These include;

- Assumption that backlash would not halt the programme or that the programme would have capability and capacity to adapt where necessary to uphold do no harm principles.
- Assumption that there is will and capacity amongst key programme stakeholders to participate in the programme.
- Civil society groups focussed on women’s rights are aligned in their approaches and priorities for enhanced protection, decreasing harmful gender norms and equal leverage

3.3. Document review

Following the ToC workshop, we conducted a document review of all country partner reports.²⁵

Throughout the course of the NAP III programme, across country programmes a total of 55 organisations were involved.²⁶ This includes 16 Dutch organisations and 39 local organisations. A full list of organisations involved in the implementation of the NAP III programme can be found in Annex 6. With such a large number of organisations implementing the programme, capacities and skills, expertise and interpretations of the NAP III ToC will consequently vary.

²⁵ This does not include local partner reports as this is out of the scope of the evaluation.

²⁶ Although some organisations were no longer able to operate during programme implementation including: STAD and their implementing partners EWO in South Sudan; Women Peacemakers Program in Yemen and Libya.

There are clear thematic differences per country approach to the NAP III programmes as depicted in Table 3 below. The thematic focus areas are based on the respective programme documents.

Table 3: Thematic focus areas of the eight NAP programmes²⁷

Country programme	Security & justice	GBV services	Women's participation	Youth participation	Gender norms
<i>Afghanistan</i>	X		X	X	
<i>Colombia</i>		X	X	X	
<i>DRC</i>		X	X		X
<i>Iraq</i>	X	X	X		
<i>Libya</i>		X	X	X	X
<i>South Sudan</i>	X	X	X	X	X
<i>Syria</i>			X		
<i>Yemen</i>			X	X	X

We understand from the workshops together with the document review that the thematic focus pursued by the different country programmes was shaped by the country context and the different partner experiences, expertise and strengths. Apart from the varied thematic focuses per country programme, where country programmes were similar in theme, they differed in approach. Examples derived from the country-programme evaluations, other documents, and interviews and workshop are provided below.

Security and justice: The Afghanistan, Colombia, DRC, Iraq and South Sudan programmes focused on working with formal and informal security and justice actors. Both Afghanistan and Iraq programmes aimed to work with justice systems but were met with institutional challenges, particularly when dealing with informal justice mechanisms. The Iraq programme withdrew from working in the justice sector to focus on security but the Afghanistan programme using the model of embedded advisors continued to influence national justice systems within the Ministry of Justice with some anecdotal evidence of successfully influencing them.²⁸ In South Sudan, local justice systems also known as the ABC courts were sensitised to issues of gender and GBV, with anecdotal success of increased representation of women court members and better referrals of GBV cases.²⁹ In DRC, the programme worked with a range of formal and informal justice and security actors, such as village chiefs, police and community leaders to increase knowledge about national laws on GBV prevention and response. These targeted efforts were supported by broader public awareness-raising campaigns.

In Colombia, work on women's access to justice included educating women about existing justice mechanisms, including engaging closely with two of the transitional justice mechanisms set up under the peace accord, the

²⁷ Thematic areas distilled from the programme documents

²⁸ Anecdotal evidence of influencing the Ministry of Justice to recruit a consultant to do inclusive justice training (Oxfam Novib (2018) Safhe Jaded Annual narrative report 2018); Ministry of Justice support for justice sector training manuals Oxfam Novib (2017) Safhe Jaded Annual narrative report 2017.

²⁹ Plan Annual Reports (2018, 2019 and 2020) claim that more gender-related cases are being brought to court. Plan 2018 Annual report notes that in 2018, the Nyang chairperson of the ABC courts ensured two women are included in the court, which previously consisted of only men. This happened after he participated in a training by AMA in June 2018 which included the 35% quota for women representation. In Sudan People's Liberation Army – In Opposition (SPLA-IO)-held territories such as Nyang only local informal justice systems are present

Truth Commission (*Comisión para el Esclarecimiento de la Verdad, la Convivencia y la No Repetición – CEV*) and the Special Jurisdiction for Peace (*Jurisdicción Especial para la Paz – JEP*). This included organising hearings of survivors with both mechanisms, successfully advocating for a stronger gender strategy for the CEV and helping the JEP improve its gender-responsiveness, including with respect to indigenous women and LGBTIQ+ persons.³⁰

Approaches to working with security actors also differed. The Afghanistan and Iraq programmes aimed to increase representation of women in the police force. Both worked at national level with Ministries of Interior to institutionalise these changes, and relationships with the Ministries in both country programmes are presented as well developed with some evidence of awareness-raising and attitude change in Iraq.³¹ The Afghanistan programme model of embedded advisors within the Ministry of Interior demonstrated more success with reports that the training and support resulted in increased reporting of harassment against women within the police force,³² facilities built for women police officers³³ and revision of two core policies (Gender Policy, Policy on Prevention of Harassment Against Women).³⁴ In Afghanistan, DRC, Iraq and South Sudan, the programmes worked to sensitise police officers to issues of GBV and gender, including training them on existing national legal frameworks.

In South Sudan, the programme trained police who were responsible for running the gender desk at police stations. Reporting provides very little information about this, however, some successes in case referral were noted in the case study (see South Sudan case study, Annex 1).

GBV services: Colombia, Libya, DRC and South Sudan programmes all had a GBV service provision component. Libya and Colombia programmes focused on GBV prevention through training of women, awareness-raising and training of local actors to empower communities to address GBV themselves. In the case of DRC and South Sudan, there is a stronger focus on developing referral pathways to enable women access to necessary services. In Colombia, implementing partners worked in several locations on creating, mapping, and strengthening care, support and referral pathways for survivors/victims of GBV, as well as educating women and officials about these.³⁵ In Colombia and South Sudan, Healthnet TPO together with its local implementing partners helped establish local-level psycho-social support mechanisms and networks especially for GBV survivors in target communities. These were based on locally-articulated needs and locally available resources mapped through the Resources Mapping and Mobilization approach.³⁶ The work on GBV thus was more pronounced in terms of response rather than prevention, and the latter tended to focus mostly on awareness-raising, including on laws. More transformative work on tackling the root causes of GBV through gender norm change was however also included in the Colombia, DRC, and South Sudan cases (see also Annex 1).

On the prevention side, there is sparse anecdotal evidence from the Colombia and Libya programmes to suggest that training on GBV awareness was effective in the *prevention* of (as opposed to response to or awareness of) GBV within communities.³⁷ Whilst reporting provides the number of persons trained, there is no evidence of knowledge or attitude change as a result. Evidencing prevention or reduction in GBV would

³⁰ ICCO, Mensen met een Missie and Healthnet TPO (2020). *Women as Central Agents for Peacebuilding in Colombia* - Final Report NAP III.

³¹ For example, in Iraq, the PAX 2017 annual report notes that head of the Police Training Directorate of the Ministry of Interior showed signs of increased awareness on dealing with the past and truth seeking related to SGBV through seeking continued dialogue with al-Amal Association on gender related training matters. In 2021, which falls outside of the scope of this evaluation, the consortium was able to influence the Ministry of Interior Code of Conduct, which was seen as a major advocacy win (KII with Iraq implementing partner

³² Oxfam Novib (2019). *Safhe Jaded Annual narrative report 2019*.

³³ ARM consulting (2021). *Final Evaluation of Safh-e-jadid programme (final version)*.

³⁴ *Ibid.*

³⁵ ICCO, Mensen met een Missie, Healthnet TPO (2020) *Women as Central Agents for Peacebuilding in Colombia* - Final Report NAP III.

³⁶ Healthnet TPO (2019). *Evaluation of the Resources Mapping and Mobilization Approach in the frame of the project 'Women as Central Agents for Peacebuilding In Colombia,'* and ELSconsults (2020) *End evaluation, WG4C project 2016-2020*.

³⁷ It should be noted that evidencing prevention or reduction of GBV would require extensive monitoring frameworks, starting with robust baseline prevalence data

require rigorous baseline and endline studies that would require quantitative and qualitative research, for which the resources were not available.³⁸

From DRC and South Sudan, there is some anecdotal evidence that the referral system is being used to address GBV and issues surrounding GBV. For example in South Sudan, the referral system set-up through the psycho-social focal points has made 111 referrals across four areas in South Sudan, including cases of child marriage, alcohol abuse and GBV³⁹ (although this is the tip of the iceberg in relation to the scale of the problem). In DRC, around 200 duty bearers were trained on national legal frameworks on GBV prevention and response, trainings which were backed up by advocacy and awareness-raising activities by women's rights groups.⁴⁰ However, there is no strong evidence to show that the referral mechanisms are increasing the number of referrals or indeed, that the experience of the survivors is improved by using these pathways. In Colombia, the data from the final report and evaluation reports documents the number of pathways created, as well as the number of women and public officials trained on the existence of these.⁴¹ However, there is little information available on the impact this had on GBV prevention or survivors' experiences of accessing these pathways. An unexpected positive outcome has been that these pathways have acted as blueprints and have been replicated in other parts of Colombia.⁴²

Women's participation: All programmes aimed to increase women's participation in decision-making spaces. This was mostly done through building the capacity of individual women, of local initiatives, women's groups and CSOs. This included raising the awareness on women's rights, gender norms and UNSCR 1325 and building capacity of these groups for advocacy and lobbying activities. These often reached a high number of women, with for example the South Sudan programme reporting around 1,200 women and girls trained,⁴³ the Colombia programme reporting around 2,000 trained,⁴⁴ and in DRC by 2019 almost 7,000 persons had been trained on women's rights and UN Security Council Resolution 1325 and a further 5,800 on understanding harmful gender norms.⁴⁵ Other programmes took a more focused approach, with those in Libya, Syria, and Yemen for example doing less broad outreach but concentrating on building the skills and capacities of implementing partners and more select groups of activists and local CSO beneficiaries.

There are few anecdotal stories where these activities have resulted in increased representation of women in decision-making spaces, for example in Afghanistan,⁴⁶ DRC,⁴⁷ Iraq,⁴⁸ Yemen,⁴⁹ and Colombia.⁵⁰ However, there is no evidence available on whether once in these positions, women use their power to advocate for more gender-sensitive decisions.

There is some evidence in the Afghanistan and Syria programme of advocacy efforts conducted by CSOs and women's groups being successful in creating change. One example in Afghanistan involves a result of successful advocacy against harassment in a hospital of Parwan, several male doctors were fired in 2019, and

³⁸ Furthermore, to complicate matters, increasing the understanding of GBV and building trust in GBV-prevention and response mechanisms often leads to initially higher reporting rates.

³⁹ ELSconsults (2020). End evaluation, WG4C project 2016-2020.

⁴⁰ Mensen met een Missie, CARE and Stichting Tosangana (2019). Mwanamke, Amani na Usalama/ Women, Peace and Security – Democratic Republic of Congo (DRC) Final report November 2016 - December 2019.

⁴¹ ICCO et al. (2020), Rodriguez, Adriana, Yeny Álvarez, and Raúl Bernal (2019). Final evaluation of the project "Women as Central Agents for Peacebuilding in Colombia" - Final report

⁴² Rodriguez et al. (2019)

⁴³ ELSconsults (2020). End evaluation, WG4C project 2016-2020.

⁴⁴ ICCO, Mensen met een Missie and Healthnet. TPO (2020).

⁴⁵ Mensen met een Missie, CARE and Stichting Tosangana (2019). It is unclear from the report, however, if there is overlap between these two figures, as well as between these participants and 2,300 men and 1,800 boys who received training on positive masculinities.

⁴⁶ A group of AWEC advocated for the inclusion of women in the traffic department. As a result of several meetings with key government officials and local governance bodies, they convinced the officials to hire three women officers in the traffic department (ARM consulting (2021). Final Evaluation of Safh-e-jadid programme).

⁴⁷ Mensen met een Missie, CARE and Stichting Tosangana (2019).

⁴⁸ Interview with Dutch consortium partner, 20 September 2021.

⁴⁹ A water management committee has been established in Habbat sub-district where four of the eight members elected were women (Insight Source Centre (2021). Final Evaluation NAP Yemen programme).

⁵⁰ Multiple examples of women's participation in governance structures (ICCO (2018). Annual Narrative Report).

more female nurses were hired. In Syria, a programme participant successfully advocated for the establishment of a support centre in Idlib for former female detainees and for female family members of detainees.⁵¹

The South Sudan, Yemen, Syria and Libya programmes aimed to increase the linkages between women at the grassroots level and national level and in the case of Yemen and Libya, to participate in international fora.

Youth participation: Youth participation was an important aspect of the Afghanistan, Colombia, DRC, South Sudan and Yemen programmes. Despite core youth activities within these programmes, the ways in which youth contribute to the desired outcomes is not well developed in the programme theories, nor is the category of ‘youth’ clearly defined. Approaches to youth participation are largely similar in the ways that they aimed to challenge gender norms amongst young people and capacitate the youth groups to advocate for women and youth participation through awareness-raising and campaigning. In Yemen, youth-led initiatives were supported in building their capacity, and supporting them in promoting gender equality and peacebuilding, including through arts- and sports-focused as well as participatory approaches.

Gender norms: All of the programmes included work on transforming gender norms to some degree, as also discussed further in the thematic case study (see Annex 1). Some of this work also fell under other categories discussed above, such as engaging with security and justice sector actors, building women’s and girls’ capacities to participate and promoting youth participation. The majority of the gender norms work tended to be about promoting and teaching about women’s rights, WPS and GBV, be it through public awareness campaigns, trainings or capacity-building and discussions within the consortium. The Colombia, DRC, Libya, and South Sudan programmes had more dedicated curricula and approaches aiming at transforming gender norms, in particular with men. These included dedicated trainings on positive masculinities implemented by local partners in Colombia that reached 812 men and boys.⁵² In DRC, the Men Engage⁵³ approach developed by CARE International was used, in which 851 of the approximately 2,300 men who had received positive masculinities training enrolled.⁵⁴ In South Sudan, an adapted version of Plan International’s Champions of Change has been rolled out, with additional modules on peacebuilding and slightly different curricula for girls/young women and boys/young men.⁵⁵ In the Libya programme, gender norm change work was also conducted amongst participating CSOs, using transformative approaches applied especially by Libyan gender rights activists.

Key finding 3: Country programme reporting captures activity and some output-level results to demonstrate what the programme has achieved. However, **evidence of outcome-level change is sparse, anecdotal and poorly evidenced.** Within reporting, **there is a ‘missing middle’ in evidencing intermediary changes** which are critical on the pathway towards more ambitious outcome-level change.

We conducted a mapping of available evidence from the documents onto the NAP III ToC. We found the strength of evidence was low across the country programme reporting (see Figure 6). Figure 6 shows that strong (green) evidence is patchy across the country programmes with most evidence from documentation being amber i.e. not triangulated or validated or only presenting evidence that an activity happened or in some cases, output-level change. There was also reporting that was anecdotal or statements without evidence and are therefore weak (red). The grey colour means that no evidence was found; however, this may be because the country programme did not address this strategic objective in the ToC.

This weakness of evidence is not to say that there was no change, rather that the change that reports were trying to capture was overly ambitious in relation to ToC outcomes and any incremental progress towards the

⁵¹ Hivos (2020). Programme proposal 2020.

⁵² ICCO, Mensen met een Missie and Healthnet TPO (2020).

⁵³ This approach should not be confused with the international network of feminist men’s organisations of the same name.

⁵⁴ Mensen met een Missie, CARE and Stichting Tosangana (2019).

⁵⁵ ELSconsult (2021), Interview with Plan Nederland.

Coalition building – Colombia, Syria and Libya programmes: Another issue which emerged from the interviews and workshops with the implementing partners but was largely absent from the reporting, with the exception of Libya,⁵⁸ were the efforts invested into building, managing, and maintaining a coalition which spanned a wide range of varying positions, both on politics and gender issues. While coalition building is frequently mentioned in WPS literature as a key element for successful civil society engagement, it is often depicted as a mostly positive process, rather than an often messy and difficult reality where compromises need to be made and sides have to ‘agree to disagree.’ This was the case in at least Colombia, Libya and Syria, where implementing partners ranged from more radical feminist and pro-LGBTIQ+ rights organisations and activists to more socially conservative and religious organisations.⁵⁹ In Libya, respondents also shared their experiences of working on shifting gender norms within the consortium through transformational trainings held for implementing partners, which in a number of cases led to fundamental shifts at the individual participant level.⁶⁰

 *Globally speaking, the transformational engagement with men and boys as part of WPS programming is comparatively unique.* 

Addressing barriers to change including working with men and gatekeepers: As also discussed in the social norms case study (see Annex 1), a number of the programmes, in particular Colombia, DRC and South Sudan, included elements of working with men for transformative change as well as engaging with male gatekeepers in communities. While the former was through dedicated programming based on gender transformative change curricula, the latter was more based on individual engagements, in part to overcome unexpected resistances to working on WPS, as for example in South Sudan.⁶¹ Globally speaking, this transformational engagement with men and boys as part of WPS programming is comparatively unique.⁶² However, beyond being mentioned as an activity among others, reporting on these approaches did not go into much detail on this issue, nor was the important question answered of how men’s resistance to the work on WPS in the communities was overcome.

Flexibility and adaptation

Key finding 4: Country programming has been **agile and adaptive** to difficult and dynamic contexts which has enabled implementing partners to be opportunistic in generating change and continue work in adversity. However, **longer-term thinking for programme sustainability** was generally not well planned for across programmes.

In the lifetime of the NAP III programme, several country programmes have experienced contextual shifts between times of relative stability and conflict as well as facing the current Covid-19 pandemic and other contextual challenges. In general, programmes have been adaptive to this and shifted their approaches accordingly.

Libya and Iraq programmes successfully shifted to contextual challenges they were met with once the programme started. For example, the Iraq programme faced serious institutional blockages when working with justice systems and so adapted to focus on security institutions. On the other hand, Libya faced issues

⁵⁸ The reported efforts in the activity report, however, focused on bridging political differences rather than on gender issues.

⁵⁹ Interviews with Dutch and local consortium members, August-October 2021.

⁶⁰ Interviews with Libya consortium member, October 2021.

⁶¹ Plan International (2020).

⁶² Myrntinen, Henri (2019). “Locating Masculinities in Women, Peace and Security,” in Davis, Sarah and True, Jacqui (eds.), The Oxford University Press Handbook on WPS. Oxford: Oxford University Press.

with rifts between different implementing partners due to the worsening security situation and adapted to introduce inclusive coalition building to help build bridges.⁶³

The Afghanistan and Yemen programmes were able to capitalise on shifts within the context. In Afghanistan, the programme adapted its activities in line with up-and-coming elections. It held workshops for parliamentary and district council candidates in Daikundi parliament which claims some contribution to Daikundi being the first and only province where women achieved equal seats to men and where women's turnout to vote was highest in the country.⁶⁴ In Yemen, the shift to working online due to Covid-19 meant they were able to reach a broader audience which coupled with changes in the conflict dynamics enabled implementing partners to reach more women and youth groups and local authorities freely.⁶⁵

Sustainability

On the whole, long-term planning for sustainability was not well developed – or if it was, this was not clearly articulated in the reporting. Sustainability was conceptualised differently between and within the programmes. Measures listed in annual and final reports as contributing to sustainability included building on pre-existing, long-term relationships, seeking to work with existing structures to institutionalise change, building capacity of local actors, providing seed funds to CSOs, working with youth, or ensuring community buy-in for activities. In 2020, the NAP III bridge fund placed a greater emphasis on the importance of sustainability as some programmes were coming to an end.

Until 2020 none of the country-programmes had exit strategies in place to clearly lay out how the activities, outputs and outcomes would be sustained. Several of the programmes reflected on exit strategies but did not elaborate them further, in at least one case because they knew that their funding would be continued.⁶⁶ Afghanistan and South Sudan developed such exit plans: The South Sudan exit plan in the Final report, 2020 thoroughly considers each component of the country-programme in line with the ToC and how it can be made more sustainable. This can be used as a good example if applied at programme-design stage. These plans would have increased effectiveness if implemented from the beginning.

Other country-programmes did not use an exit strategy. The country-programme documents suggest that sustainability of results are inherent within the programme design, however, as noted in the ToC workshop, due to the short implementation period, the change engendered by the country programmes was mostly at an individual level (for example individual attitude change on harmful gender norms) as opposed to group, community or institutional changes which would generate more sustainable results. Therefore, how activities translate into sustainable change was mostly poorly reported across programmes, including in external evaluations.

What could help improve thinking through and integrating sustainability into programme design and implementation would be a discussion between the Dutch MFA and grantees on what sustainability could look like in volatile, conflict-affected contexts – which was indeed suggested in one of the critical reflection workshops. This discussion could include a more detailed discussion of different kinds of sustainability (Figure 7), what is realistic in the given circumstances, and what would need to be in place to ensure sustainability.

Figure 7: Types of sustainability

⁶³ Cordaid (2019). Annual Narrative Report 2019.

⁶⁴ Oxfam Novib (2018). Safhe Jaded Annual narrative report 2018.

⁶⁵ Insight Source Centre (2021). Final Evaluation NAP Yemen programme.

⁶⁶ E-mail communication with Dutch implementing partner

1. **Community sustainability** (e.g. enhanced resilience, access to services) achieved

2. **Financial sustainability** (recipient has independent access to resources) achieved

3. **Organisational/institutional sustainability** (e.g. independent functioning, networking) achieved

4. **Regulatory sustainability** (new laws, regulations are changing established behaviours) achieved

5. **Political influencing** (sustained change in decision making) achieved

6. **Strategic/contextual sustainability** (e.g. in security) achieved

Discussions of sustainability should also include reflections on implicit assumptions (e.g. will continued funding or continuing with more of the same approaches lead to sustainability) as well as with how to counter risks to sustainability to the degree possible, such as countering roll-back on women's rights or more incremental slide-back on gender equality.

2.6 Critical reflection workshops

The evaluation team conducted two critical reflection workshops with the implementing partners in August 2021. The purpose of these was to create a safe space for participants to openly reflect on the evaluation team's findings from the document review against the ToC, as well as have more in-depth discussions on underlying assumptions, whether these proved to be correct, and examining how change was brought about.

Country-based programme partners were divided into two groups based on the main focus of their work:

- Group A: Libya, Yemen, Syria and Colombia – programmes focusing on building networks and primarily working with civil society (though Libya could not attend and a separate meeting was held).
- Group B: Afghanistan, DRC, Iraq, and South Sudan – programmes working with government in context with national WPS NAPs (though Afghanistan could not attend).

The critical reflection session participants included both local partners as well as Dutch consortia members, ideally one each, though some programmes only had one participant and others invited a number of local partners. WO=MEN also joined both critical reflection workshops. Due to unfolding events in Afghanistan, neither the Dutch consortium lead nor local partners were able to participate. A separate meeting was held with representatives of the Libya programme as they were unable to attend the critical reflection session.

We explored some of the programmatic pathways, impacts, and underlying assumptions during critical reflection sessions, where participants were able to reflect on the following initial findings from the document review discussed above in country programme-based breakout groups:

1. There is a wide variety of approaches across the country programmes (see also key finding 3).
2. Programming has been flexible and adaptive to challenging and dynamic contexts.
3. There is a leap between activities and outputs and ambitious outcomes which has led to a 'missing middle' evidence gap.
4. Defining what can realistically be achieved and evidenced as sustainability requires more careful consideration.

 *An issue which emerged in both critical reflection workshops was the mental and emotional toll that lack of progress and setbacks on WPS issues took on implementing staff, and how time, efforts and resources that were invested into trying and failing to achieve a particular goal were not able to be reflected in reporting frameworks.* 

Feedback received at critical reflection sessions supported the evaluation team's initial findings, particularly that the overall objectives were too ambitious for the duration of the programmes, though partners were able

to adapt to dynamic and challenging operating contexts. Some participants felt that this ambitiousness was embedded into the overall ToC, especially in terms of the pathways of change under the specific objectives. This was then reproduced in programme-level theories of change, which were aspirational and hopeful, but not in all cases achievable in the existing contexts. There was widespread support for the finding that there is a ‘missing middle’ of evidence, which would demonstrate progress had been achieved towards the three objectives and would link the activities of programme partners to the intended outcomes. In terms of the question of sustainability, several participants raised the wish to have a discussion with the Dutch MFA about this at the outset to ensure both sides understand what can be expected realistically in dynamic and conflict-affected settings. Lastly, an issue which emerged in both critical reflection workshops was the mental and emotional toll that lack of progress and setbacks on WPS issues took on implementing staff, and how time, efforts and resources that were invested into trying and failing to achieve a particular goal were not able to be reflected in reporting frameworks.

3.3. Case studies and survey

The final component of the evaluation were the three case studies, two focusing on country-level programmes (Colombia and South Sudan) and one thematic case study. The full case studies can be found in Annex 1 with the headline findings presented in the boxes below. The case studies aimed to interrogate the ‘missing middle’ or intermediate outcomes from the programme, however, as noted in section “2.5 limitations and challenges” the document evidence base is low-quality and the primary data collection scope was limited with limited independent sources. Therefore, the case studies do not present robust evidence of outcome-level change but instead present successes identified by implementing partners and beneficiaries⁶⁷ that are under-reported or absent in the country-programme reports which should be considered in future WPS ToC design, reporting and MEL systems.

As part of the case studies, we further conducted a survey which interrogated the functioning of the consortium model in Colombia and South Sudan. The models were different in each country with the consortium in South Sudan led by Plan and an ‘alliance’ of core members including PAX, HealthNet TPO and STAD with implementing partners AMA and EWO.⁶⁸ Colombia on the other hand has more actors involved, with ICCO leading a consortium of two other Dutch partners Healthnet TPO and Mensen Met een Missie and ten local partners as implementing partners.

The survey was sent to representatives from both Dutch consortia and local partners. This is a total of 17 organisations.⁶⁹ There were a total of 13 responses. Amongst these, respondents shared the benefits and challenges of working within a consortium. Benefits included:

- The opportunity the consortium offered to learn from each other, share experience, knowledge and expertise.
- It enabled a greater reach in terms of geographical spread and access.
- It offered potential for joint advocacy and mutual support.
- And it offered opportunities to build solidarity around the agenda.

The challenges mentioned include:

- Challenges to find common ground amongst such diverse partners with different capacities and approaches with risk of siloed approaches.
- The power differentiation between Dutch and local partners, and how this impacted representation and decision making.
- Challenges with the MEL systems due to lack of co-ordination amongst pen-holders to build a common system of reporting.

⁶⁷ Triangulated between KIIs and FGDs

⁶⁸ PAX, Healthnet TPO and STAD are Netherlands-based organisations. STAD is a diaspora NGO. EWO is the local implementing partner of STAD and AMA the local implementing partner of PAX.

⁶⁹ The survey was not sent to EWO as their office closed.

- Lack of clear chain of accountability and responsibility.
- Budget constraints posed by a consortium approach as it requires more co-ordination time creating issues of cost-efficiency.

Solutions provided to these challenges included: bringing people together in annual gatherings or learning events to build consensus, improving coordination from the lead and capacity building with members of the consortia. Even though in the new WPS programmes under the Strengthening Civil Society tender the consortia members have shifted, all respondents noted that NAP III consortium partners and local partners remain in close contact.

Headlines from the South Sudan case study

The ‘missing middle’ outputs and intermediary outcomes that are under-reported in the country-programme reports include:

1. Success in fostering trust and resilience in communities supporting the functioning of GBV referral mechanisms.
2. The country programme helped form relationships and networks for individual and group empowerment amongst women.
3. Consciousness raising in individuals, including men and boys and changing dynamics at household level.

Key challenges included upholding ‘do no harm’ (DNH) and ensuring the safety of staff and participants when addressing sensitive themes in patriarchal contexts (see the note on DNH below). Also, managing weaknesses in government institutions and staff turnover.

Efficiencies in the way the country programme was delivered were found in partners sharing their expertise and knowledge; however, this was undercut by ongoing tensions and power dynamics created in the consortia model.

The **sustainability** of the programme results were prepared for by shifting from targeting individuals to targeting groups in 2020 to empower local groups to continue activity beyond the lifetime of the programme. However, this was not planned for from the beginning and therefore is limited in efficacy.

Headlines from the Colombia case study

The ‘missing middle’ outputs and intermediary outcomes that are under-reported in the country-programme reports include:

1. Activities had a significant multiplier effect in the communities and were able to reach more people.
2. The country programme created spaces of forgiveness and reconciliation, opening up opportunities for local level reparations.
3. Advocacy spaces were created at the institutional level providing in-roads to influence policy and policy implementation.

Key challenges included upholding ‘do no harm’ and ensuring the safety of staff and participants when addressing sensitive themes in patriarchal contexts and increasing violence against rights defenders (see the note on DNH below), as well as Covid-19.

There were **efficiencies** in having both Dutch and local partners for shared learning and geographical reach. However, with such a large number of diverse partners, it was challenging to create consensus and there were issues with budget allocation meaning that some local partners received considerably less than others. There was also a lack of clarity and translation issues around the ToC and MEL framework amongst local partners.

The **sustainability** of programme results is not guaranteed. The country programme did not plan rigorously for sustainability and results are seen mainly at an individual level as opposed to community or institutional levels, reducing the likelihood of enduring change.

A note on Do No Harm (DNH)

‘Do No Harm’ (DNH) is both a principle and a practice that involves understanding a context and the impacts of aid interventions on it and making sure that unintended negative effects are prevented or minimised. DNH is based on the awareness that actors and organisations (through their actions and behaviours) have an influence on the context they operate in and any intervention can have unintended or unexpected impacts that could reinforce, or create, conflict or tensions and undermine social cohesion.

DNH recognises that any intervention should, at a minimum, aim to minimise negative impacts (avoid causing harm) and, where possible, maximise positive impacts (contribute to peace, stability and social cohesion).

The DNH approach consists of the following key elements^[1]:

- Understanding the context, its dividers and connectors (by undertaking a context analysis)
- Understanding how the intervention interacts with the context and what potential (negative or positive) impacts it could have on it
- Making adjustments to the intervention (by developing options and supporting opportunities) so that it has more positive and fewer negative impacts.

It is important to note that DNH is not about avoiding to act on issues that some might see as controversial (such as addressing gendered power imbalances) because there is the slightest possibility of doing harm or a risk of negative outcomes. DNH is about understanding the complexities of the environment in which one works and think about different ways of doing things to have better effects.^[2] DNH should be integrated at different levels, from the individual level (i.e. making sure that staff understands DNH and knows how to apply it) to the policy level (i.e. DNH is a guiding principle that is referred to in policy documents)

^[1] For details on how to apply DNH, please see CDA Collaborative Learning Projects (2015), *From Principle to Practice: A User’s Guide to Do No Harm*

^[2] *Ibid*

4. Analysis of findings

This section synthesises and analyses findings across document review, workshops and case studies. It is organised by the main evaluation questions (EQs) in the Terms of Reference for this evaluation and refined/operationalised in the inception report.

4.1. NAP programme contributions to impact and unintended outcomes

Evaluation Question 1: What are the NAP programme contributions to the overall objective of WPS framework 2016–2019 and 2020? What are the programme contributions to other positive and negative outcomes (intended and unintended and sensitive to the ‘do no harm’ principle) and were there any significant differences in this respect between different consortium members/implementing partners?

Taking a very literal reading of the overall objective of the Dutch WPS NAP III, the programmes can be said to have been successful, as all of them did ‘contribute to an enabling environment for women’s participation and empowerment in conflict and post-conflict environments.’ However, it is difficult to give an overall verdict on how successful they were in this respect. This is in part because of the wide variety of programmatic approaches and activities, the contextual differences between the various countries and due to factors beyond the scope of the programmes, such as political and security developments that sometimes enabled and sometimes negated the efforts of the programmes. However, the difficulties in gauging the extent of the overall impact also arise from the lack of an overall MEL framework,⁷⁰ and from programme-level reporting being more focused on activities rather than systematically tackling and evidencing outcomes and impacts.

🔗🔗 Greatest successes were achieved where there were pre-existing mechanisms and institutions that could be engaged with and where there was a social and political openness to the interventions. Nonetheless, that does not mean that work in the more challenging contexts, where change was more incremental, was for naught. On the contrary, the small, at times very micro-level changes achieved may prove to be important initial openings and often were quite substantial, at a personal level, for those involved. 🔄🔄

Partners reported achieved outcomes at various levels, ranging from the individual and micro levels, at community level and within institutions and in part at the national level, where in some cases programmes contributed to processes of change. Examining the three specific objectives of enhanced protection, decrease of harmful gender norms and equal leverage in conflict prevention, resolution, peacebuilding, relief and recovery, the greatest successes were achieved where there were pre-existing mechanisms and institutions that could be engaged with and where there was a social and political openness to the interventions. Nonetheless, that does not mean that work in the more challenging contexts, where change was more incremental, was for naught. On the contrary, the small, at times very micro-level changes achieved may prove to be important initial openings and often were quite substantial, at a personal level, for those involved.

In terms of unintended or unexpected consequences, in some cases, such as Colombia, DRC and Libya, the programmes had unexpected ‘spin-off’ effects as beneficiaries involved with the programmes were inspired to set up their own local level projects. In Yemen, young people involved were reported as having become more active on social issues on social media while at least in Colombia and Iraq, women who had been beneficiaries of the programme went into local- and national-level politics. Some of the consortia have, even if in a changed set-up, continued to work beyond the lifespan of the NAP and some of the local implementation partners were reported as continuing to co-operate with each other. This includes a number of the local

⁷⁰ Having an overarching MEL framework for programmes designed largely independent of each other and operating at different levels through different approaches, as they were in NAP III, is also challenging to impossible. For the NAP IV, the overall M&E framework also does not track overall or joint progress/ impact on NAP IV outcomes by civil society signatories

implementing partners in Libya, which did not receive funding beyond 2020, now working together in the process to design the Libyan WPS NAP.⁷¹

However, the programmes have also faced resistances and backlash, be it against women speaking out on political issues, women participating in programmes, against men choosing to question assumptions about what it means to be a man in a given society, or against programme participants seen as speaking with others from across the conflict divide. In at least one case, programme staff also reported of local resentment against what was seen as the privilege of programme participants who were able to travel to workshop meetings. In part, the resistance and backlash had been anticipated and factored into risk mitigation plans. However, some of the negative reactions were either not anticipated or not anticipated to the degree with which they emerged.⁷² In Colombia, Libya and Yemen at least, the social space for working on WPS issues was perceived to be rapidly shrinking during the implementation period, and the risks and threats against women activists rising, as documented in detail in the Colombia case study.⁷³ This shrinking of social and political space increases the risks of negative unexpected consequences of activities that previously might have been seen as innocuous, necessitating a constant review of risk assessments and mitigation plans.

Given the sensitive issues addressed by the programmes, the volatile conflict-affected situations they operated in, the often-shrinking space for civil society activists and activism as well as the risks of backlash, programmes adopted a range of risk mitigation and ‘do no harm’ approaches. These included seeking to proactively engage with gatekeepers, incorporating mitigating measures, understanding and respecting local sensitivities without compromising on gender equality goals, and providing support to women – to the degree possible – to women activists coming under pressure for their work. A key new facet of ‘do no harm’ that was integrated into some of the programmes was to include self-care elements to avoid burn-out and support staff, including community volunteers, working on emotionally and/or physically taxing issues.

4.2. NAP programme contributions to the WPS framework three specific objectives

Evaluation Question 2: What are the NAP programme contributions to the WPS framework 2016–2019 and 2020 three specific objectives and what change happened along the causal pathway - did assumptions hold?

As with EQ 1, the lack of an overall MEL system for the whole of the NAP and reporting gaps in the programmes, including a lack of baseline data and poor evidencing of change and impact, make answering this EQ difficult. Ideally, from a MEL perspective, such a system would have been able to capture both the individual causal pathways of the various components of the programmes but also show how these different parts cohered and added up to a greater whole. From a programmatic and practical point of view, however, such a system may quickly become overly complex, resource-heavy and unworkable. Not designing a robust MEL system during the design phase of the WPS framework 2016–2019 and 2020 can be considered as a missed opportunity to monitor and manage the implementation of the programme and its constituent country-programmes and report on successes and failures.

In terms of the three specific objectives:

- ***Enhanced protection*** – *Better protect women and girls in conflict and post-conflict situations from violence and violations of their rights.*

Much of the work falling under ‘enhanced protection’ in those contexts where there is a functioning and responsive state apparatus focused on improving the implementation of existing laws on GBV and ensuring formal and informal actors take the issue seriously, as well as on improving responses and support services available for survivors of GBV. Efforts were undertaken across the programmes to raise awareness of women and girls of GBV but also of existing rights. Whilst all of this work is

⁷¹ Communication with Libyan implementing partner

⁷² See for example Plan International Nederland (2020)

⁷³ ICCO, Mensen met een Missie and Healthnet TPO (2020).

extremely important and should continue to be supported, much of it does not necessarily constitute *protection* per se, but is more about response or civic education.

As often in WPS work, the focus of the programmes tended to be on GBV rather than protection from all forms of violence *and* from violations of their rights more broadly. Even in taking the narrower approach of focusing on GBV rather than on violence more broadly, however, programmes of the size of these NAP III programmes can only contribute to prevention and response systems. Ideally, these efforts on awareness-raising and response should be linked to simultaneous work that focuses on addressing the root causes of violence such as harmful gender norms and unequal power relations.

The causal pathways of change differed between programmes, but most often these included awareness-raising on GBV among women beneficiaries, including on what GBV is and on their rights; awareness-raising and advocacy among justice and security sector providers at either the national or local level; and the creation of different kinds of support and referral pathways. Some of the implemented activities did link GBV prevention with root causes (thus linking the first and second specific objectives – see below and Annex 1), but other activities were narrower, such as explaining existing national legal frameworks on domestic violence to duty-bearers, which were mostly men. Some of the programmes also were able to set up crucial support systems for survivors which were previously non-existent, such as psycho-social support services in the Colombian and South Sudan cases. The levels linked by these activities were thus mainly the individual, local and national levels. In terms of numbers of intended beneficiaries reached, the awareness-raising campaigns were by far the most prominent activity under this specific objective. However, given an absence of reliable baseline data (e.g. on prevalence and reporting of GBV), it is not possible to gauge how much of an impact these activities had in terms of reducing violence or improving service provision. Anecdotal and qualitative evidence does nonetheless point to increased levels of awareness of beneficiaries of their rights.

A protection element that was particularly prominent in the Colombian programme, and which went beyond GBV in the narrow sense, was increasing protection for women rights defenders and other socially active women, including for the staff of the CSOs involved in programme implementation. This included training on protection and self-protection, the development of protection plans and protocols, as well as public awareness-raising and social media campaigns on this issue.⁷⁴ Given the shrinking space for women activists in many contexts, and increasing risks for those working on gender equality and WPS issues in different contexts, these measures could be integrated as good practice in other WPS programming.

• ***Decrease of harmful gender norms*** – *Subvert harmful underlying gender norms, which are obstacles to sustainable peace.*

Tackling harmful gender norms was approached in a variety of ways, from public awareness-raising through radio shows and posters reaching out to communities to dedicated transformational change curricula for a specific targeted group of individuals. Again, a lack of reliable baseline data and lack of tracing of impact makes it difficult to gauge overall impact, but what evidence was available does point to at least partially successful efforts (this is discussed further in the gender norms case study in Annex 1). As echoed in literature on gender norms change, the most successful efforts are the ones which seek to affect change at multiple levels, i.e. with individuals, their social environment, but also institutional and policy level; which invest sustained time and efforts into the change processes and address multiple aspects of norm change simultaneously. Here, the more comprehensive approaches used in Colombia, DRC and South Sudan as well as the internal efforts in the Libya case are the most promising.

The causal pathways of change under this specific objective for the most part focused on the individual and community levels, seeking to both raise awareness and open discussions on harmful gender

⁷⁴ ICCO et al. (2020)

norms, and in part transform these. Of the specific objectives, it was the one which most centrally engaged with men and boys to bring about change, as discussed further in Annex 1. In some cases, such as in South Sudan, the initial, more transformative approach was adjusted so as to spend more time raising basic awareness first and then tackling harmful gender norms at a later stage of the implementation. Beyond the individual beneficiaries and their communities, most programmes - Afghanistan, Colombia, DRC, Iraq, South Sudan and Yemen - also sought to shift gender norms at the institutional level, be it local-level, sub-national (e.g. provincial, district or governorate-level) and national actors. Often, the minimum aim was to raise awareness among these actors of national legal frameworks and improve their implementation of these, linking to the 'protection' objective. The degree to which duty bearers and their institutions were open to engaging with civil society actors in general, or on gender norm changed varied, and at times required a shifting of the focus to more responsive institutions or regions, as in the case of Iraq and Yemen.

• Equal leverage in conflict prevention, resolution, peacebuilding, relief and recovery – *Ensure that women have equal leverage in conflict prevention and resolution, peacebuilding, relief and recovery at all levels, and that their efforts are acknowledged and supported.*

Taking a literal reading of this specific objective, this could be deemed to be the least successful one, as in no country was equal leverage achieved. However, if one takes this as an aspirational goal rather than a hard target, many of the programmes were able to create new and important openings for women, and actively support women in participating meaningfully and/or interacting with justice and security providers. The parameters within which this was possible varied however greatly depending on the country context. In Colombia, at one end of the spectrum, it was possible to push not just for women's participation but rather ensure that diverse women were able to participate, at various levels and across a range of institutions and processes. At the other end of the spectrum, in Libya and Syria for example, programmes struggled with getting any women or women's perspectives into political processes, including peace negotiations. Where access to national-level participation was blocked, however, the programmes then sought other entry points at the local or international levels, such as Dutch parliamentarians or the UN Human Rights Council.

As detailed above in section 3.3, the programmes across all eight programmes increased the capacities and skills of thousands of women in the target countries to be able to participate better and more meaningfully, and to have a better understanding of existing systems and processes.

The most common pathway for change under this specific objective was that of increasing women's capacities to participate actively and meaningfully, which was an element to one degree or another in all of the programmes. The focus of who was supported and how differed between the programmes, from broader community level engagement such as in Colombia, DRC, South Sudan or Yemen; to targeted support to women active in civil society such as in Colombia, Libya or Syria to supporting women who were active in local or national level politics, such as in Colombia and Iraq. Given the challenges of engaging at the national level, the Libya and Syria programmes especially also sought to find entry points at the international level, such as at the United Nations level. An important part of this work, which was not always reflected much in the reporting, was collation-building between women's organisations to increase leverage and find common ground on WPS issues, as for example in Colombia, Libya and Syria.

By and large, assumptions on which the programmes were based on held, even if changing circumstances required flexibility and adaptation in all cases. What was under-estimated in some of the programmes was the degree of reluctance of key actors to engage on gender issues and the resistance that work on WPS might face at different levels. The programmes which utilised participatory methods to map local-level needs, available resources and potential challenges were able to also have the most grounded assumptions upon which to design their programmes.

4.3. Alignment with other frameworks and contextual relevance

Evaluation Question 3: How far were NAP programmes aligned or responsive to:

- ***WPS policy frameworks***
- ***National and decentralised policies***
- ***The needs of beneficiaries and target groups***
- ***The changing context***

The degree to which the various programmes engaged or aligned with national government policies, in particular on WPS, varied greatly, in part because most of the countries did not have their own WPS NAPs. The spectrum of engagement spans from Syria, where there was no engagement of the programme with the central government under Bashar al-Assad to Afghanistan, where the Safhe Jaded programme was explicitly designed to support the implementation of the national WPS NAP of Afghanistan. In Afghanistan, Colombia, DRC, Iraq and South Sudan, the programmes actively engaged with national and local governance frameworks and actors, including formal and informal justice and security sector actors. This kind of engagement was not possible in Syria, and only to a limited degree in Libya and Yemen. Rather, the programmes had to find other kinds of entry points, including in the Libya and Syria cases engaging more in advocacy with Dutch decision makers, other donor governments and the UN rather than the national governments. All of the programmes designed their own, national-level ToCs to align with the overall WPS NAP III ToC, but, as discussed above in section 3, emphasised different thematic areas and chose different entry points, be it at the local, sub-national, national or international level.

Apart from the national level, sub-national and local levels were a key site of affecting change for all programmes other than Syria. Again, the approaches varied greatly, from local level women's economic empowerment in Libya over psycho-social support in Colombia and South Sudan to increasing spaces for women's participation at the governorate level in Yemen and concerted efforts at changing men's attitudes and behaviour in DRC (see also Annex 1). Local-level implementation was in some cases informed by local-level needs assessments, as for example in DRC, Colombia and Libya, which allowed for implementing partners to respond to particular community needs as articulated by beneficiaries and/or utilise openings such as around increasing women's economic participation as entry points for broader change on gender equality issues or linking WPS activities to local level development plans. In other cases, such as Afghanistan and Iraq, the programmes were designed more centrally, which may have given local implementers less space to define their own approaches but helped make the programme overall a more coherent whole.

One of the cross-cutting approaches of the ToC of the NAP was increasing linkages between the local, national and international levels, and this was again implemented differently depending on the country context. In Afghanistan, Colombia, DRC, Iraq and South Sudan, this involved a 'national-to-local' level flow of information where beneficiary women but also service providers were informed of national-level laws and mechanisms (e.g. on GBV prevention and response or transitional justice). The programmes did also to differing degrees seek to facilitate the flow of information from the local to the national level, though this was often more difficult and time-consuming. Nonetheless, long-term engagement did lead to important national-level gains, such as the engagements with national-level ministries in Afghanistan and Iraq, and the transitional justice process in Colombia discussed above in section 3.3. In the Libya and Syria cases, but also to an extent with other programmes such as Colombia or DRC, it was at times easier for the consortia to voice local concerns at the international level rather than at the national level. In other cases, such as Colombia, Libya and Yemen, much of the success was at the sub-national or municipal level.

As also detailed in the case studies (see also Annex 1), local-level implementation also entailed addressing resistances to increased gender equality and navigating the risks of backlash. This resistance manifested itself at different levels. At the family and community level, husbands or parents at times objected to their wives or daughters participating in programme activities. Community leaders felt the programmes were undermining established family roles and power relations, and people who had engaged with gender equality programming at times faced pushback or ridicule from other community members. At an institutional level, security and justice providers, for example in Afghanistan and Iraq, were at times reluctant or opposed to engaging with

women's rights organisations or civil society more broadly. In at least Yemen and Libya, security concerns and shrinking space for any civil society activities impacted upon the programmes, and participants from Colombia also reflected on the need for increased vigilance. In the case of Yemen, the shrinking space for the work of the consortium led to a shifting of the geographical focus from Sana'a governorate to Taiz governorate.

Throughout the implementation of the various programmes, the consortia had to respond and react to changing circumstances, including political changes and security-related concerns, finding suitable entry points for programming and advocacy or tensions between consortium members. The most dramatic impact arguably was the outbreak of the Covid-19 pandemic, which not only led to a sudden need to re-design implementation approaches but also saw several local-level implementing partners pivoting from WPS-focused activities to Covid-19 response in several countries, or shifting to addressing increasing levels of GBV during the lockdowns. The pandemic and ensuing restrictions hampered activities relying on direct interaction, such as lobby and advocacy efforts, training and capacity building and public awareness-raising campaigns, in particular in countries and with communities that do not have reliable and easily accessible communications networks.

4.4. Sustainability

Evaluation Question 4: What evidence is there to show the benefits of NAP programmes will be sustained beyond the life of the programme and how far, and in what ways did actions in the additional year (2020) contribute to strengthening sustainability?

As discussed in section 3.3., consortia listed a wide range of measures under the term sustainability in their reporting, but for the most part did not pursue these issues further in the document. In other words, the fact that a particular activity was continued, for example, or a first step towards institutional up-take was billed as evidence of sustainability. Ideally, this would be the beginning rather than the end of a discussion on sustainability: in what way does the continuation of that activity contribute to sustainability and at which level (see Fig. 7); what follow-up is needed to ensure that the first step towards institutional take-up can be sustained and built upon?

Measuring the sustainability of the various programmes was skewed in part due to the fact that with the exception of Afghanistan and Libya, and one of the consortium partners in Colombia, much of the programming has been carried over into new Dutch MFA-funded programmes. While the ability to attract continued funding can itself be seen as a form of sustainability, it is difficult to prove or disprove the counterfactual of whether the programmes would have been able to continue without this funding. In the case of the work of Mensen met een Missie and its partners in Colombia, which is not being funded through the Dutch MFA, the activities have indeed been financed through other sources.

The additional year of funding in 2020 was used by some, but not all consortia, to at least reflect upon, or develop sustainability strategies and exit plans. To a large degree, however, consortia and implementing partners had to use much of 2020 adapting to the Covid-19 pandemic and adjusting programme implementation accordingly, leaving less space for sustainability planning.

The different country contexts allow for very different degrees of achieving sustainable outcomes, with those most affected by armed conflict facing the greatest challenges in this respect. The most dramatic – and tragic – example of this is Afghanistan, where the fall of the state to the Taliban during the evaluation period will, based on the evaluation team's prognosis and initial reports from the country, most likely mean at least a partial roll-back the gains made during NAP.

One of the areas where programmes across the board faced challenges in terms of sustainability was in engaging with state actors, or in the case of Syria and Libya with UN institutions and donor governments. Ideally, the linking of WPS programming with state structures at the local, sub-national and national levels can allow for an institutionalisation of the intended change, so that it can be sustained beyond the end of the programme. The challenges faced in engaging with state structures were two-fold: first, state institutions, in particular those working on security issues, are not necessarily open to engaging with civil society actors

and/or issues which are seen by them as ‘women’s issues.’ Second, when rapport was built up and there is engagement, state functionaries, but also diplomats or UN staff, are often rotated and/or political priorities shift, necessitating a renewal of efforts by the civil society actors. A common frustration voiced in the critical reflection workshops was that implementing staff often felt that they too often had to be reactive to events and changes rather than being able to be pro-active vis-à-vis representatives of the state and/or the international community.⁷⁵ However, in some of the programme locations in Colombia and DRC, implementing partners were able to build up good working relations with local government structures, who were keen to take WPS issues forward.⁷⁶

In at least Colombia, DRC, Libya, and Yemen, some of the beneficiaries continued engagement on their own independently of the programmes, including local level women’s lobby groups in Colombia and groups of male gender champions in DRC, as well as Libyan local partners who are now working together on the Libyan WPS NAP. Whilst some of the civil society networks created through the consortia continue to co-operate with each other, others such as in Yemen became dormant once the direct funding of the co-operation ended. In Afghanistan, the Taliban takeover of power has, at least in the short run, reduced civil society space.

4.5. Efficiency and the consortium model

Evaluation Question 5: How did programme design and implementation affect achievements along the causal pathways of change, programme sustainability and, value for money? How did the dynamics between various consortium partners (including Dutch partners and local implementing partners) impact the implementation of the programme?

The ability of programmes to implement activities as had been originally intended in their design varied depending on the context. The more stable, comparatively speaking, the situation was, the more the original plans and intended pathways for change were able to be followed. As discussed above, adaptation and flexibility were nonetheless required from all consortia. Based on their reporting consortia were able to largely reach or even surpass their respective activity and output goals, with some indicators, such as participation rates in some activities in Colombia, surpassing the projected figures manifold. However, as discussed, outcome and impact were, for the most part, either not at all or not adequately documented. This points to a gap in the design of the programmes between the ‘action’ phase and its follow-up. While the input and activity sections were designed, adapted when necessary and implemented in a way that mostly met the set targets, structures and processes were not in place to capture what this led to, and how sustainable these changes were. In terms of efficiency and value for money, the evaluation team did not conduct a detailed budgetary analysis, and the diversity in types of programming and the respective resources they require make comparisons impossible.⁷⁷ However, given the challenges highlighted in this evaluation around MEL systems and the gaps in capturing impact, this may well be an area where resources could be spent more effectively by designing systems which, from the outset, are geared towards better capturing outcomes and impacts.

The implementation of the programmes through consortia of Dutch and national NGOs received mixed reviews from those involved. While many respondents did feel that the broad approach brought a ‘richness in diversity’, others did refer to the consortia as ‘forced marriages.’ Whilst in the best of cases, consortia were able to create synergies (e.g. Local implementing partners complementing each other’s skill sets and mutually enhancing capacities), in other cases there was a sense of opportunities lost. Around half a dozen respondents reported tensions within their consortia, be it due to real or perceived differences in power, size and funding; whether or not partners felt that their voices were being heard, but also thematic and political differences –

⁷⁵ This was especially the case in more volatile contexts, where implementing partners had to react to changing political and security situations beyond their control, but also in situations where international, national, or sub-national political priorities changed, for example following a changed in the elected governments.

⁷⁶ This was due to a variety of factors, including an openness on the side of the local government structures to engage on these issues, long-term advocacy work by local women activists and women’s rights organisations, existing national and local policy frameworks mandating an engagement on gender/WPS issues and the use of multiple avenues by implementing partners to engage with these structures, including awareness-raising, capacity-building, direct advocacy and increasing broader public awareness and interest on these issues.

⁷⁷ For example, the resources needed for training community-level psycho-social health workers to reach out to, say, 500 survivors are much higher than for a public awareness-raising radio spot that might reach 10,000 people but have less profound impact.

be it on the conflict in the country or on gender-related issues. In the case of Libya, the consortium pro-actively used some of these tensions as an opportunity for working through them together as a coalition-building process while in Syria, differences were navigated through a reaffirmation of core principles in spite of differences, plus engaged mediation. In at least the cases of Libya and South Sudan, one source of power imbalances and tensions within the consortium was that not all local partners or diaspora organisations were able to fulfil the formal criteria to join. Communication between consortium members in the Netherlands, between their representatives in-country and local implementing partners proved challenging in some cases as well.

Lessons learned from the ways in which the various consortium models operated were the need for:

- Good internal management of consortia, including responsiveness to members’ concerns and good conflict management.
- Ensuring the consortia are not too unwieldy in size.
- A common sense of purpose, including taking the time to allow members to get to know and understand each other at the outset.
- Ensuring that power imbalances are acknowledged and reduced.
- Roles and responsibilities being clearly delineated and defined.
- Ensuring that decision-making processes are clear from the outset.

Given that there is often a power imbalance between larger, global North-based (I)NGOs and national partners, it is important that there is not an onus on the national and local partners to prove their worth and added value, and that the involvement of the larger INGOs also brings added value to the smaller partners.

In addition to the consortia, country groups were established as part of the Dutch NAP Community, which included non-consortium members and Dutch MFA and/or the respective Embassy of Netherlands’ representatives. An unexpected impact of Covid-19 and the subsequent relocation of country group meetings online was that this allowed for much broader participation in these meetings, both from the Netherlands and implementation countries, in those cases where the country group was functioning well.

Budget analysis & Value for Money

Where data was available to compare budget vs actual variance⁷⁸, we found that Colombia, South Sudan and Yemen had very low variances in both the 2016-2019 programme and the 2020 extension with marginal underspend against the projected budget. In contrast, Afghanistan and Iraq programmes showed a higher variance (Afghanistan: 29% and Iraq: 32%). We speculate the reason for this is due to programme adaptations in response to rapidly changing context. The Iraq budget particularly points to challenges in security preventing or delaying activities (see figure 8 below).

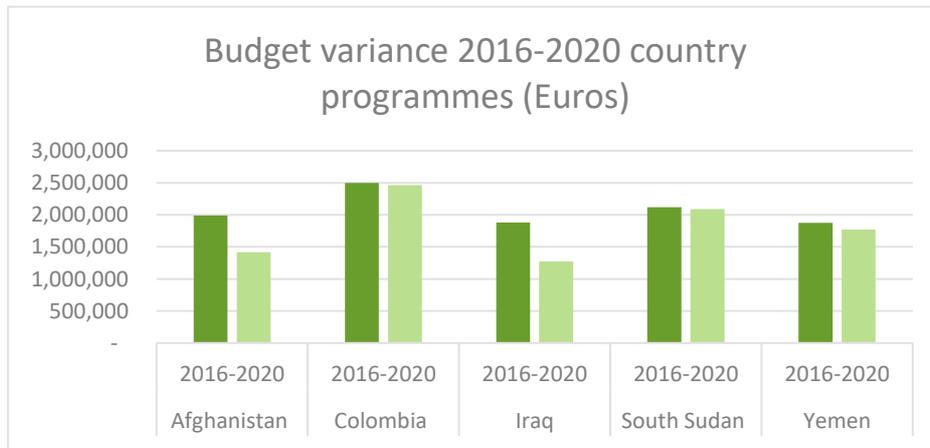


Figure 8: NAP III country programme budget vs actuals 2016-2020

⁷⁸ All countries except DRC and Libya where there was no aggregate of budget vs expenditure by year available. Data was only available for Syria 2020 programme.

Total overhead costs varied between country programmes and were described differently within the budgets but generally covered HR, PMEL and Admin costs. On average, the overhead costs were high. Afghanistan, Colombia, South Sudan, Yemen and Syria⁷⁹ all had overhead costs which accounted for over 30% of the overall budget. Overhead costs for Colombia were the highest at 46%. Higher overhead costs in Colombia may be a factor of the consortium being much bigger than to the other country programmes and therefore requiring more co-ordination, however, this is our assumption and differing costs may also be due to a variance in the ways in which overheads are calculated per country. Widely differing from the other country programmes, the Iraq overhead costs are extremely low at only 3%. This huge difference requires further investigation but again may be due to a difference in the way overheads are calculated and presented.

Within programme delivery, a number of the country programmes used approaches which increased efficiencies and value for money. These included for example adapting programmes that had proven to be successful elsewhere to the country context rather than developing one from scratch (e.g. the DRC programme adapting the Men Engage approach previously developed in Uganda and Burundi); being flexible in changing to other counterparts or advocacy targets if the original ones prove unresponsive, as in Iraq; and exploring synergies between implementing partners in the consortium design phase. However, value for money and efficiency should not only be seen in purely financial terms. For example, in South Sudan better compensating psycho-social volunteer workers both improved their effectiveness and the sustainability, while in Libya and in Syria investing more into internal coalition-building made the consortia more resilient and sustainable.

4.6. Lessons learned

Evaluation Question 6: What lessons can be learned from the NAP programmes and how far do achievements and lessons learned align with the broader international WPS evidence base?

The challenges of the NAP programmes listed here in tracking and evidencing change, of bringing together a variety of programming approaches to make a coherent whole, of shifting deep-seated gender norms, or of getting reluctant institutions to take gender seriously are not unique to the Dutch WPS NAP III. Rather, they are reflective of widely-shared challenges of WPS implementation. In spite of the various challenges, the implementation of these NAP III programmes did successfully advance important changes to WPS work. We see four key lessons learned that are of relevance to global WPS debates.

Good practice 1: Adopting a broader and more nuanced approach to gender in WPS. Compared with WPS NAPs globally, one striking feature of the overall NAP and of the programmes, was in the understanding of gender. This moved beyond equating gender with women, seeing women as a homogenous group, or equating women with victimhood or innate peacefulness only. Both the overall NAP and the country-level programmes explicitly sought to work with both men and women to advance WPS goals and gender equality, and this was anchored in the NAP's ToC. In the case of Colombia, this broader approach also included raising issues of diverse sexual orientations and gender identities. The programmes also, to differing degrees, took intersectionality and women's diversity into account. In Colombia, Iraq, Libya and Syria, for example, programmes explicitly sought to engage with the differential needs and possibilities for agency of women from different ethnic, socio-economic, and regional backgrounds, including indigenous women, but also their different political standpoints. The Colombia, DRC, Libya, South Sudan and Yemen also explicitly addressed differences between women of different age groups. Women were also not cast as victims or objects of charity, but rather the programmes sought to engage with and strengthen women who were already active in different spheres of public life, be it in the security sector (e.g. Afghanistan and Iraq), at different levels of politics and community engagement (all programmes), or women in economy (e.g. Libya). This broader understanding of women's agency did not however mean that women's needs were left unexamined. Rather, as for example in Colombia, DRC and South Sudan, these approaches happened in parallel to improving services for women survivors of violence. This broader and more nuanced approach to gender allows for a more realistic engagement with women and their differential needs and capacities for action than previous,

⁷⁹ Information available for Syria 2020 financial reports only

narrower approaches. Including men in the programming also opens up ways of engaging with key gatekeepers at the family, community and state level.

Good practice 2: Broadening the range of WPS programming approaches and themes. A further successful area of innovation was the broad variety of approaches used at the implementation level, and broadening the thematic scope of WPS programming. Thematically, the country-level programmes took a broad approach to WPS, including for example elements of economic empowerment, working in schools, engaging with men and boys and including mental health care issues. While these are critical issues faced by women in conflict-affected communities, they have, globally speaking, seldom been included under the umbrella of WPS NAPs. The programming approaches included using arts- and sports-based methods for increasing youth and women's participation, employing novel gender transformative approaches (see also annex 1), developing community-based psycho-social support systems, or developing intersectional approaches to transitional justice processes. While these approaches are not necessarily new to work in other areas such as women's and youth empowerment, GBV prevention or rural development, they are comparatively new to the WPS field.

Good practice 3: Aligning and linking WPS NAP implementation with national and local framework and processes. A further area of good practice was the linking of the programmes to existing national-level and local-level frameworks and processes. The Afghanistan case was unique in its own way, in that it explicitly linked the Dutch WPS NAP to the implementation of the Afghan WPS NAP. This is globally speaking quite rare, as often NAPs are implemented in parallel to each other rather than building on synergies, and NAP implementation is often separate from the implementation of other strategies and action plans (e.g. gender equality strategies or GBV reduction action plans). Thus, the efforts in a number of the programmes evaluated here in aligning the WPS NAP programming with municipal-, sub-national- and national level policy frameworks and strategies, such as in Colombia, DRC and in Iraq should be viewed as good practice as it increases sustainability, institutional anchoring and policy coherence while avoiding a duplication of efforts. This should however not mean that all NAP implementation needs to always be directly linked to other frameworks, as there might be gaps which these frameworks do not cover but the NAP could address, as well as the risk that any delays or challenges in implementing the other frameworks might also negatively impact NAP programming.

Good practice 4: Taking the need for flexibility seriously. As discussed in section 3.3, all of the programmes – and the Dutch MFA as a donor – had to show a great deal of flexibility and adaptability during the implementation phase, be it in responding to programming challenges, shifting national or sub-national contexts or the Covid-19 pandemic. Having this flexibility is essential to successfully working in volatile and conflict-affected contexts, and is essential for risk response and mitigation. It also helps in ensuring better effectiveness and efficiency, as implementing partners were able to seek out new entry points for engagement if the original plan prove unworkable or new opportunities arise.

5. Conclusions

Due to the limitations of the evaluation, it is not possible to draw rigorous conclusions on the NAP III programme outcome-level achievements. Instead, this evaluation report offers important considerations on MEL systems and programme efficiencies, contrasts between different programmes and reflections from implementing partners drawn from country programme evaluations and programme documentation which, whilst low quality were able to be triangulated to an extent through a small number of KIIs and FGDs. This provides important learning for NAP IV programme design. Based on the activity reports of the programmes, the Dutch NAP III programmes reported success at many levels. The rich variety in approaches and the diversity of the contexts and implementing partners was very much a positive aspect of these programmes. The programmes navigated extremely challenging contexts and unexpected developments, such as the global Covid-19 pandemic. On the whole, partners noted that the consortia worked well, even if there was room for improvement. The overall ToC was able to successfully guide the design and implementation of the programmes, even if the articulated specific objectives proved to be overly ambitious.

The deficiencies in the evaluations and MEL systems, in particular the lack of baseline data and of evidencing of results at the outcome and impact levels hampered the evaluation team in seeking to ‘understand the impact of the NAP frameworks against their goals and if and how they have achieved them’, as outlined in the evaluation’s Terms of Reference. The identified gap between the programme activities and the outcome-level change is what we have termed as the ‘missing middle’. The missing middle is made up of outputs and intermediary outcomes that are essential markers on the pathway to outcome-level change, and these were often not articulated or captured. Furthermore, many pathways of action towards outcome-level change were making significant progress during programme implementation, yet were absent or not captured well within reporting. We have sought here to address some of these gaps and ‘missing middles’ to the degree possible. However, in future NAP programming, these should be avoided from the outset by thinking through pathways of change, intended impacts and how to document these.

All of the work of the consortia can be seen as having been relevant to both the country contexts and the overall objective of the Dutch NAP III, but at the specific-objective level, activities were only able to partially contribute towards the ambitious goals. In terms of enhanced protection, a broader ‘prevention, protection, and response’ objective might have made more sense. Its scope could have been better delineated, either by explicitly narrowing it to GBV or, alternately, actively encouraging implementers to consider violence and violations of rights more broadly. The use of local participatory mapping methods should continue to be encouraged, so that programmes are able to respond to on-the-ground needs as well as capitalise on entry points and existing resources.

In terms of effectiveness and efficiency, the fluidity and volatility of the contexts which the programmes worked in makes it difficult to predict at what level, how and through which entry points impact can realistically be achieved. In order to minimise the risk of wasting efforts and not having any impact, work often needs to happen at multiple levels and with multiple targets simultaneously, and have the flexibility of seeking out alternative pathways when necessary. Effective programming can also require difficult trade-offs, such as whether to use resources to, for example, reach a broader group of people with a simple, but not highly transformative messages or focus on a smaller group and achieve more profound and possibly more sustainable results. Here, different consortia opted for quite different approaches, both with their own benefits and drawbacks. Consortia were in some cases, such as in Colombia, DRC and South Sudan, able to build on synergies between implementing partners. However, even in those programmes in which synergies were utilised, there were also instances of siloed implementation. These could and should be addressed at the design stage and when consortia are formed, but also revisited and, if need be, addressed during implementation. As highlighted in section 4.6., one of the positive lessons learned

In terms of sustainability, more could and should be done in future programming in incorporating thinking through different kinds of sustainability as well as exit planning into the initial design phase of programmes. While all programmes had addressed sustainability issues in some way, the way in which this was

conceptualised differed. Here, it might be useful for the MFA to have structured discussions with implementing partners as to what can be realistically expected from such programming in conflict-affected contexts. The additional year in 2020 did not lead all programmes to further develop sustainability plans and exit strategies beyond initial consultations, in part due to having to adapt to Covid -19 and in part as they already knew whether or not they could get additional funding in 2021.

6. Recommendations

We have divided our recommendations into three broad thematic groups:

- **Consortia and ways of working**
- **Design, monitoring, evaluation and learning**
- **Thematic focus areas**

Consortia and ways of working

On the whole, the consortia approach worked well, though we have identified several ways in which these could be improved in section 4. Key recommendations include:

For consortium partners and local partners:

1. Acknowledge the time and resources for building, managing and maintaining consortia and coalitions, especially if these span differences in political views, ethno-religious or regional divisions, or differences in stances on gender issues.
2. Allow for more time and resources for consortium members to get to know each other.
3. Link up partners in-country to support each other and capitalise on synergies and improve efficiency.
4. Continue to allow for the necessary flexibility in the implementation of programmes in conflict-affected contexts.
5. Acknowledge and seek to mitigate power imbalances, ensuring all consortium and local partners feel appreciated and heard.
6. Continue to ensure 'do no harm' approaches and plan for preventing and responding to potential backlash and shrinking spaces, including by investing in mapping risks and updating risk registers.

For the Dutch MFA:

1. Encourage the linking of Dutch WPS NAP activities with implementation country WPS NAPs and other relevant local, sub-national and national gender equality frameworks where this is feasible.
2. Continue to encourage the engagement of implementing partners with Dutch actors, including embassies, in particular when and where it is challenging to engage with national or international level actors.
3. Encourage openness in reporting, including on when and why objectives were not met, and allow the consortia the necessary flexibility for implementation.

Design, monitoring, evaluation and learning

Some of the main challenges highlighted in this evaluation are linked to programme design, monitoring, evaluation and learning. These include:

For consortium partners and local partners:

1. **Theories of change:** Take time to develop a well thought through intervention logic to ensure the level of ambition and scope of the programme ToC is realistically achievable, and be specific about what the programme is aiming to achieve and how, including any intermediate steps towards outcomes and critical assumptions.
2. **Collaboration and integrated working:** Consortia should use the theory of change development process as an opportunity to bring implementers together to build positive relationships between them; to collectively understand the common goals of a programme; how those goals connect; how each implementer can contribute to those goals, and to identify ways in which different parts of the programme and different implementers fit together and can support each other to achieve objectives.
3. **Context:** Where feasible, invest more in localised, thorough context analyses and collect baseline data for programme components. It may be possible for consortium partners and local partners to share and combine their existing contextual analyses, or for the programme to use and build on those which are already publicly available. Baseline data may also already be available in some cases for certain programme components.
4. **Flexibility and realism in MEL:** Ensure the room for flexibility and adaptation that is necessary at the programme-implementation level is reflected in a simple MEL framework. Improve MEL frameworks to address 'missing middles' and focus more on outcomes and impact. Test and use monitoring methodologies which account for dynamic contexts - such as outcome harvesting – that are also straightforward to understand for all implementing partners regardless of their MEL capacity and experience. Where possible invest resources in periodically capturing intermediate change (i.e. beyond outputs) and aim to do this directly with programme beneficiaries to triangulate evidence.
5. **Sustainability and exit planning:** Consider from the outset if and how the programme can be sustainable and what resources and outcomes that will require. Where possible, involve all consortium partners and local partners in this discussion as there will be differing understandings of sustainability depending on the context and what is possible.
6. **Efficiency:** Map possibilities for increasing efficiency in the project design phase by drawing on existing resources and adapting these where possible. Map synergies both amongst implementing partners and between the programme and other frameworks, strategies and being implemented by other actors, and evaluate how to best engage with these.
7. **Do No Harm:** Ensure DNH context analyses and approaches are integrated into the design of all stages of the programme cycle, and that these are reviewed and updated on a regular basis. Wider structures, policies and ways of working should support and encourage DNH in order for DNH to be effective.

For the Dutch MFA:

1. **Reporting:** Ensure clarity over where the responsibility lies for reporting evidence at different levels, especially if partners are working to different theories of change. Consider providing guidance on potential reporting mechanisms within a larger MEL framework to help harmonise the way evidence is generated (whilst acknowledging types of evidence may be wide-ranging).
2. **Evaluation:** Ensure that when evaluations are conducted, those carrying them out are aware of quality criteria, e.g. those of the IOB. Consider harmonising suggested evaluation approaches to improve chances of comparability across programmes; for example a focus on contribution rather than attribution; a focus at the intermediate level of change.
3. **Sustainability:** Encourage discussions of what sustainability could look like in a given context as well as exit strategies at the inception phase of programmes.
4. **Learning:** Provide opportunities for consortia to come together to discuss contextual challenges and share learning about the types of programming and interventions that are working and why.

For all of the WPS NAP partners:

1. Improve on the learning aspects of programming, including, linking with other WPS donors and implementers, and good practices emerging from the programmes more broadly, including between country programmes and third parties, such as UN agencies.

Thematic focus areas

In terms of the thematic focus areas covered by the NAP, more could be done to ensure different aspects and approaches of a programme support and enhance each other, as several of the programmes did. Change often needs to happen at various, mutually-reinforcing levels simultaneously.⁸⁰ While the possibilities and scope of working on the different specific objectives varied between the country contexts, broad recommendations are:

For the Dutch MFA, consortium partners and local partners:

1. Broaden the scope of protection to also include prevention and response, as well as encouraging thinking of protection/prevention beyond GBV, i.e. prevention of/protection from all forms of violence and threats to women's and girls' rights.
2. Consider more carefully which gender norms are tackled and what assumptions these intended change processes rest upon; ensure gender norm change is transformative and goes beyond one-off trainings or campaigns, and engages with men's but also women's resistance to change, and links individual change to community-, institutional- and policy-level change processes.
3. Consider what can be realistically achieved in different contexts in terms of diverse and meaningful participation, how this can be achieved and impacts measured beyond the number of women in the room; continue to create and utilise entry points at various levels as these arise.

⁸⁰ For example, individual-level change should ideally be reinforced by the surrounding community and be backed by laws and state institutions; increased gender equality in participation can lead to increased prevention and protection, whilst improved protection and prevention of violence also increases diverse participation.

7. Annexes

Annex 1 – Case studies

Annex 2 – Data collection tools, case study approach and sampling

Annex 3 – Evaluation matrix

Annex 4 – Quick-scan results

Annex 5 – NAP III ToC, reconstructed ToC and ToC workshop interactive whiteboard capture

Annex 6 – Full list of consortium partners and local partners

Annex 7 – List of stakeholders consulted

Annex 8 - Bibliography

Annex 9 – Full Terms of Reference



Itad is a global organisation. Our strategy, monitoring, evaluation and learning services work to make international development more effective. We generate evidence on important issues – from malnutrition to migration – to support our partners to make informed decisions and improve lives.

Itad.com

 [@ItadLtd](https://twitter.com/ItadLtd)

mail@itad.com

Itad Ltd

Preece House, Davigdor
Road, Hove, East Sussex,
BN3 1RE, UK

+44(0)1273 765 250

Itad Inc

1110 Vermont Ave NW,
Suite 500, Washington, D.C
20005, USA

Tel: +1 (301) 814-1492



Annex 1: Case studies

**External evaluation of the Netherlands
WPS 2016-2019 and WPS 2020
programmes**

Lessons from the Women as central agents for peacebuilding programme, Colombia

Whilst **outcome-level results were not well captured** in the ‘women as central agents for peacebuilding programme’, the **country-programme had positive impacts at the individual, community, municipal and national levels**. Lessons can be learned from unintended consequences, and the strengths and drawbacks to the consortia model and the diversity of partners going forward.

Introduction

The Dutch WPS NAP III programme in Colombia was implemented between 2017 and 2020 (NAP III including its 2020 extension), and was named “Women as central agents for peacebuilding in Colombia”. This project was implemented by a consortium of three Dutch organisations: ICCO Cooperation, Mensen met een Missie and HealthNet TPO, and 10 Colombian organisations: *Ruta Pacífica de las Mujeres* and *Mencoldes* (ICCO partners), *Corporation Amiga Joven*, *Católicas por el Derecho a Decidir* (CDD), *CODACOP*, *Corporation SerVoz* (previously Corp. Tamar), *Espacios de Mujer*, *REMPAZ* and *Red Tamar CRC* (Mensen met een Missie partners) and *LIMPAL* (HealthNet TPO partner).

The general objective of the consortium was to contribute to an environment conducive to **women’s empowerment** and **participation in peacebuilding** by significantly increasing women’s participation in **conflict prevention, peace consolidation, protection** and **recovery** in different geographical areas in Colombia as depicted in Figure 1.

The activities focused on three main objectives (Theory of Change - ToC):

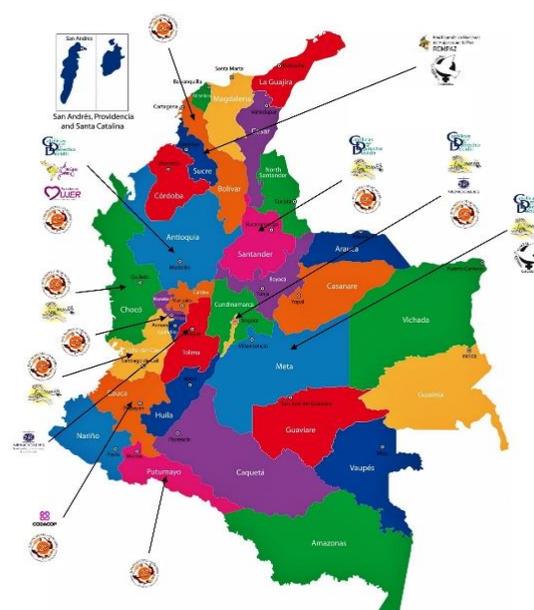


Figure 1: Project implementation locations and implementing partner

1. Enhanced **protection** for women and girls
2. Decrease of **harmful gender norms**
3. Equal leverage in (local) decision making related to **conflict prevention** and **resolution, peacebuilding**, relief and recovery peace and security.

This case study builds on the two NAP III Colombia external evaluations conducted in 2019 and 2021 in order to further explore the programme results and interrogate its

efficiency, relevance to the context and sustainability. It is part of the meta-evaluation of the NAP III programmes and should be read alongside the main report.

Methods

As part of the meta-evaluation, we conducted a theory of change workshop, a desk review and critical reflection sessions.¹ This helped us to identify key questions and evidence gaps in the various country programme findings. We selected two country programmes to focus on for further exploration and primary data collection: Colombia and South Sudan. Primary data collection in Colombia was conducted by a national consultant based in Bogotá and remotely for Dutch consortia members. Ten key informant interviews (KIIs) were conducted (seven with participants in Colombia and three with participants in the Netherlands). Limitations of low-quality documentation and a low number of independent sources means findings are not robust evidence of outcome level change but instead programme success identified by programme partners and beneficiaries that are under-reported.

Unpacking the missing middle: How change happened in the NAP III programme, Colombia.

The ToC for both the NAP III and the Colombia programme were identified as overly ambitious in the documentation and ToC workshop.² Much of the country programme MEL aims to capture change against outcome level ambitions that in reality are not yet reflected on the ground. Programme results are often sparsely reported, poorly evidenced and anecdotal – which also meant that many of the achievements have gone undocumented.³

However, many pathways of action towards outcome-level change were making significant

progress during programme implementation yet were absent or not captured well within reporting. The progress between the programme activities and the outcome-level change we have termed as the **'missing middle'**. The missing middle is made up of outputs and intermediary outcomes that are essential markers on the pathway to outcome-level change.⁴ See Figure 2.

Figure 2: The missing middle between activities and outcomes



In this section, we account for the important progress of results from the 'missing middle' in the lifetime of the NAP III Colombia programme.

Missing middle: Significant multiplier effect in the communities



The participation of several local Colombian organisations helped the consolidation of a big network of women that did not know each other before the project and had the possibility to work together within the NAP III programme implementation.⁵ This gave them the opportunity to strengthen their incidence and collaboration in different territories where they conducted a joint process. Likewise, through alliances among the consortium's organisations and with other actors, women and their organisations have been able to influence the implementation of the Peace Agreement and the 51 gender measures which were included in it through their participation in advocacy, the formulation of proposals, and their engagement with the transitional justice system.⁶ In addition, local organizations have monitored the implementation of the Agreement and have denounced the

¹ The ToC workshop and critical reflection sessions were held with Dutch and local consortia members to reflect on the programme and desk-review findings.

² ToC workshop held 28/07/2021.

³ KII 4 with Dutch implementing partner

⁴ It is worth noting that several outputs in the WG4C ToC are pitched at outcome level. Outputs are still within the control of the programme while outcomes are not.

⁵ KII 1, 2, 3, 6, 7, 8 with local implementing partners.

⁶ KII 1 and 3 with local implementing partners.

persecution, threats and assassinations of social leaders.⁷

Moreover, according to all of the KIIs with local partners, the NAP III programme not only opened a window of opportunity to articulate different efforts around one common purpose, women's empowerment and their participation in peacebuilding, but it also allowed more people to be part of the different activities and processes carried out by the organisations. In this sense, many women and men were part of the project, producing genuinely an important multiplying effect.

"It was a nice surprise. In three years, more was achieved than could have been expected."⁸

Missing middle: Increasing self-confidence and participation in the public sphere



For the local partners of the project,⁹ a visible impact of the NAP III programme was strengthening the self-confidence of women, and even some men who participated, on their capabilities to impact their communities. After the interventions of the project, many participants ran for elected office as mayors, local councils and other political participation spaces. In Montes de Maria, a highly affected conflict zone in the Caribbean region, 27 women and 3 men who joined the programme were able to improve their mediation skills, and they became important actors in the local elections and even in the daily lives of the communities.¹⁰

Missing middle: Spaces of forgiveness and reconciliation



Many activities implemented within the project allowed victims and former combatants to gather in the same spaces,¹¹ opening the opportunity of a deep collective

dialogue about the damages, pain and grievances produced by a more than 50 year-long armed conflict. This opened a space to process the hatred and rage contained by the victims, allowing the recovery and construction of collective memory, redefining womanhood and opening opportunities for reparations at the local level.¹²

Missing middle: Advocacy spaces at the institutional level

The tools developed during the project were recognised within the institutional level, which expanded the incidence of the organisations' work from the local to the national arena. As an example, *Espacios de Mujer* and *SerVoz* worked together on a public policy document called "Assessment of the implementation of anti-human trafficking policies in Colombia". The document has been used in the national Congress on several occasions to design, restructure and implement policies to prevent and combat trafficking in persons, especially women and children.¹³

Unintended outcomes and barriers

The programme came across several unintended outcomes as well as barriers to implementation. This includes:

Challenges to upholding 'do no harm' principles: Security risks and expectations management

Even if in 2016 the Peace Agreement between the Colombian government and the Revolutionary Armed Forces of Colombia (FARC) was signed, the power vacuum left by the guerrilla was replaced by many non-state groups. Within the years, confrontations between these groups have increased considerably the levels of violence in the different territories of the country. In this context, working with communities that are caught in the crossfire poses further risks to programme participants, including threats to

⁷ KII 3 with local implementing partner.

⁸ KII 8 with local implementing partner.

⁹ KII 1, 2, 3 and 8 with local implementing partners.

¹⁰ KII 8 with local implementing partner.

¹¹ KII 2 with local implementing partner.

¹² KII 2 with local implementing partner.

¹³ KII 6 and 7 with local implementing partners.

their well-being and safety. The weak presence of state institutions in the rural areas and the continuous assassinations of social leaders have affected the confidence and security of women leaders.¹⁴

“Every social leader who is murdered, threatened or displaced is the fracture of an entire community process”¹⁵

According to one KII,¹⁶ although they did not receive a pamphlet with direct threats, the risk was constant for the participating women leaders and therefore, it was important to minimise the risks through different security protocols, and protection plans were developed under the NAP III programme. However, even if the local organisations have adopted various measures to protect their staff and programme participants on the ground, this problem goes beyond their scope of action and requires a greater commitment from the national government.

Other risk situations were also present in the private sphere. The project had positive impacts for many women who started to see themselves as rights-bearers. However, at times this empowerment also led to negative unintended consequences: some male partners were violent towards female participants as they started to challenge the socially-established gender norms that they have had to obey.¹⁷ Thus, a more active and integrated participation of men in these programmes is key to achieving sustained results and reduce the risk for participating women (see also Case Study 3 on Gender Norm Change).

Furthermore, participants living in more marginalised communities not only face violence, but an increasing socio-economic vulnerability due to the absence of state

institutions. For this reason, their expectations with interventions such as this programme can be enormous, and when it comes to an end, expectations for a follow-up are high.¹⁸ Even if the organisations continued working with the communities, some of them faced many challenges as they did not continue in the consortium that will implement a new WPS programme under the Strengthening Civil Society tender. Subsequently, these organisations, which are mainly small and locally based, had difficulties in raising funding and support for continuing their work in the communities.

COVID-19: An unexpected challenge to WPS NAP III implementation

In 2020, additional challenges arose due to the health emergency caused by COVID-19, which led to mandatory confinement throughout the country and the need to move activities from the physical sphere into virtual. This necessitated a comprehensive rethinking of how to implement activities under the programme.

Apart from adapting activities to the virtual space,¹⁹ the main challenge during the pandemic was maintaining women's participation in programme activities. The new barriers faced by women were three-fold. First, the burden of care work for women increased considerably. Second, many participants either lost or were suspended from their jobs, increasing economic precariousness. Third, there was an increase of gender-based violence (GBV), especially in the domestic sphere.²⁰ Lastly, many of the participants did not have a good mobile device or stable internet connection or did not have basic digital skills.²¹ The programme sought to adapt its interventions and counter-act and mitigate these challenges, including the resultant emotional impacts on women of these multiple stress factors.

¹⁴ KII 2 and 8 local and 5 with Dutch implementing partners.

¹⁵ KII 2 with local implementing partner.

¹⁶ KII 3 with local implementing partner.

¹⁷ KII 1 with local implementing partner.

¹⁸ KII 1 and 8 with local implementing partners.

¹⁹ KII 3 and 8 with local implementing partners.

²⁰ KII 3 with local implementing partner.

²¹ KII 3, 8 KII 3 with local implementing partners and KII 5 with Dutch partner.

Strengths and weaknesses of how the programme was implemented

In the Colombian case, there are several key things to learn from the way in which the programme was designed and implemented.

The consortia model

Strength: Joint work between Dutch and Colombian organisations

The joint work between Dutch and several Colombian organisations allowed for a positive exchange between the different members of the consortium. On the one hand, on the Colombian side, small locally-based organisations were able to be part of the project and to learn from organisations that have historically more expertise and acknowledgement in the country. On the other hand, the presence of these small organisations made it possible to reach remote regions of the country. The only way to reach remote areas is to work with these organisations. The most significant impact of the project was precisely to make these areas and their inhabitants visible, give them a voice and allow the women who live in them not to feel forgotten.²²

Weakness: A vertical decision-making process

Many of the local organisations argued that they were not considered during the design process of the project. The approach was seen as being a top-down (from the Dutch partners to the Colombian organisations) rather than a joint effort to think and structure the objectives and interventions of the programme. Therefore, it was felt that it was a vertical decision-making process instead of a horizontal one, which would have encouraged a joint planning and implementation process. Interviewees felt that these power dynamics reduced the independence of the Colombian organisations, as they only executed the measures adopted from the Netherlands.

²² KII 3, 6 and 8 with local implementing partners.

²³ KII 9 with Dutch implementing partner

²⁴ KII 5 with Dutch implementing and e-mail exchange with local implementing partner

Diversity of partners

Strength: Sharing knowledge and skills

As stated previously, the joint work between different local organisations allowed them to know of each other's work and to articulate different individual strategies not only at the local, but also at the national level. In part, Colombian partners were able to build and draw on each other's strengths and expertise.²³ They continue working together on some activities despite some of them being no longer part of the consortium that will implement a new WPS programme under the Strengthening Civil Society tender.

While, as discussed in the next section, there were initial co-ordination challenges in the consortium, some respondents stressed that over the years the consortium was able to grow in to a more collaborative space for exchanging knowledge, capacity strengthening, and alliance-building.²⁴

Weakness: Coordination and budget management

All of the KIIs highlighted the difficulty in the beginning to coordinate the project considering the high number of partners and the heterogeneity of the organisations involved within it.²⁵ Even if with time they managed to coordinate their actions and to have better communication, it is essential for the organisations to maintain a dialogue among themselves, in which, in addition to addressing technical and project planning issues, processes of exchange, solidarity and mutual support are strengthened.²⁶ However, this was hard to achieve. The partners coordinated among the group, but not entirely with the whole consortium.²⁷

However, the budget was distributed equally among the three Dutch NGOs regardless of the number of local partners they worked with. For this reason, in the case of Mensen met een Missie partners, they received less

²⁵ KII 5 with Dutch implementing partner.

²⁶ KII 3 with local implementing partner.

²⁷ KIIs 1, 2 and 3 with local implementing partners.

resources²⁸ and were able to implement fewer activities.²⁹ On the contrary, ICCO's and HealthNet TPO's partners conducted a variety of activities on the ground.³⁰

Impacts, monitoring and evaluation reporting

Weakness: Lack of clarity

The indicators for measuring the impacts were not clear. They were formulated in the Netherlands and local organisations did not understand them completely. For instance, the ToC did not have a clear translation into Spanish which made the consolidation of the indicators difficult. Because of this, each organisation defined their own criteria for measuring impact. This made the aggregation for the MEL of the NAP programme in Colombia very challenging.

The wide variety of activities implemented by the ten participating local organisations made the aggregation of results even more challenging. Furthermore, larger and better-established organisations had advantages over smaller, locally-based organisations in that they could implement larger projects and document these better, making it hard to compare the achieved results.

Lastly, the project emphasised collecting quantitative rather than qualitative monitoring data.³¹ While this is important in terms of understanding the breadth of the reach of the programme (e.g. in terms of beneficiaries reached through trainings), quantitative approaches say little about the type of impact that these activities have achieved. While the collection and analysis of qualitative monitoring data is more time-consuming and resource-intensive, there are various, comparatively 'light touch' approaches which could be used in future programming. This would help all of the parties involved have a better understanding of the changes brought about by the activities, but also see where adjustments may need to be made.

²⁸ KII 5 with Dutch partner, KII 2 and 8 with local implementing partners.

²⁹ This was even more difficult for Mensen met een Missie partners, as they are mostly small locally-based organisations,

Prospects for sustainable change

Regarding sustainability, new community resources and systems have been established (e.g. some income generation activities, women association, civic oversight of referral pathways for survivors of GBV, tools of public advocacy, training workshops, campaigns, mobilizations), as a result of the project.

However, long-term impacts and their sustainability depends on the individual efforts of each implementing partner rather than the collective initiatives of the consortium. The continuation of community networks once donor's funding ends depends on multiple factors, but the future use of the results by project beneficiaries will mainly take place at an individual level.

Key lessons learned

- i. Positive sustainable impacts require time. Even if the way in which the programme was implemented had many challenges, the co-operation between different organisations is key not only to reach remote areas and to have a multiplier effect in the communities, but also to have long-lasting, sustainable effects.
- ii. Due to the increased risk of women leaders and human rights defenders during the pandemic, there was a need identified for organisations to maintain contact with these women after the end of the project, as a measure of protection against the materialisation of the risk.
- iii. It is necessary to reinforce the participation of men in the programmes, so they can be part of the gender norms change and can be supportive towards women in their communities.
- iv. Covid 19 represented a huge challenge to reach the participants so that they could participate in the activities virtually. It is

whose work depend exclusively on donors and international cooperation resources (KII 8 with local implementing partner).

³⁰ KII 1 and 3 with local implementing partners.

³¹ KII 2 with local implementing partner.

necessary to improve women's knowledge in the use of technological tools in order to reduce the digital divide they experienced by living in remote areas without access to internet.

Lessons from the Women and Girls for Change project in South Sudan

As part of the Dutch WPS NAP III programme

The Women and Girls for Change project implemented in South Sudan **generated change towards its goals, particularly amongst individuals and within communities**. However, this **progress was not well captured**. This case study aims to better represent this progress whilst drawing out key lessons from how the programme was implemented in a challenging and dynamic context.

Introduction

The NAP III programme in South Sudan was implemented between 2015 and 2020 by the Women and Girls for Change Alliance (WG4CA) headed by Plan International Nederland as well as, PAX, Healthnet TPO and Support Trust for Africa Development (STAD), with local implementing partners AMA and EWO. Implementing partners operated in different geographical areas including Eastern Equatoria, Lakes, and Unity States as depicted in Figure 3.

Programme activities focused on: (i) **improved protection of women and girls** from gender-based violence (GBV); (ii) improving support services and **engaging with traditional courts and police**; (iii) working with men, women, boys and girls on **transforming harmful gender norms**; and (iv) **improving women's participation in decision-making at local and national level** around conflict prevention and resolution as well as peacebuilding.

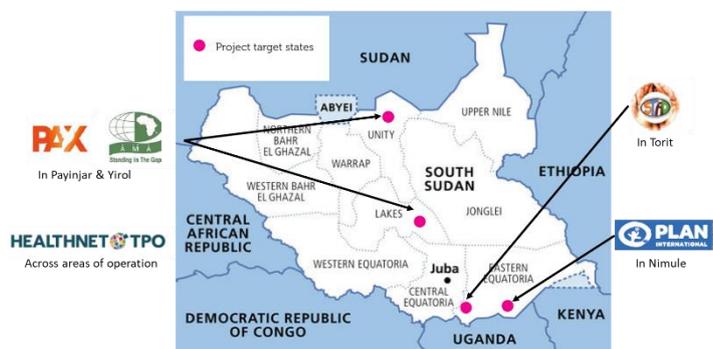


Figure 3: Project implementation locations and implementing partner

This case study builds on the NAP III South Sudan external evaluation conducted in 2020 in order to further explore the programme results and interrogate its efficiency, relevance to the context and sustainability. It is part of the meta-evaluation of NAP III programmes and should be read alongside the main report.

Methods

As part of the meta-evaluation we conducted a theory of change (ToC) workshop, a desk

review and critical reflection sessions.³² This helped us to identify key questions and evidence gaps in the various country programme findings. We selected two country programmes to focus on for further exploration and primary data collection: Colombia and South Sudan.

Primary data collection in South Sudan was conducted by a national consultant based in Juba and remotely for Dutch consortia members. Twelve key informant interviews (KIIs) were conducted (one in Juba, six in Nimule, and five remotely) and one Focus Group Discussion (FGD) in Nimule with eight participants. Limitations of low-quality documentation and a low number of independent sources means findings are not robust evidence of outcome level change but instead programme success identified by programme partners and beneficiaries that are under-reported.

Unpacking the missing middle: How change happened in the NAP III programme, South Sudan.

The ToC for both the NAP III and the South Sudan programme were identified as overly ambitious in the documentation³³ and ToC workshop.³⁴ Much of the country programme MEL aims to capture change against outcome level ambitions that in reality aren't yet reflected on the ground. As a result programme results are sparsely reported, poorly evidenced and anecdotal.

However, many pathways of action towards outcome-level change were making significant progress during programme implementation, yet were absent or not captured well within reporting. The progress between the programme activities and the outcome-level change we have termed as the '**missing middle**'. The missing middle is made up of

outputs and intermediary outcomes that are essential markers on the pathway to outcome-level change.³⁵ See Figure 4 below.



Figure 4: The missing middle between activities and outcomes

In this section, we account for the important progress of results from the 'missing middle' in the lifetime of the NAP III South Sudan programme.

ToC Outcome 1: Women and girls have access to and use psycho-social and legal protection services

Missing middle: Fostering trust and resilience within communities



A key component of the programme was to train community-based psycho-social focal points (PFPs) to become first responders and referral points for persons and families experiencing mental health and gender-based violence (GBV) issues. The planned outcome of this work was to improve the access of women and girls to psycho-social and legal support. However, a critical intermediary step towards this outcome is to successfully develop trust within communities and throughout the referral chain between community members, PFPs and other service providers, including the police and local courts. This is something the programme achieved well.³⁶

Community trust in the PFPs is indicated in the 111 cases were reported during the project³⁷ alongside anecdotal stories of the critical support they provided; for example, the 15-year old girl in Torit who reported to a PFP her parent's plans to marry her to an elderly man.³⁸ As noted by an implementing

³² The ToC workshop and critical reflection sessions were held with Dutch and local consortia members to reflect on the programme and desk-review findings.

³³ Plan (2021) Final report 2016-2020, p.18; ELSConsults (2020) End evaluation, WG4C programme, p.31.

³⁴ ToC workshop held 28/07/2021.

³⁵ It is worth noting that several outputs in the WG4C ToC are pitched at outcome level. Outputs are still within the control of the programme while outcomes are not.

³⁶ KII 1 with implementing partner, KII 3 with implementing partner.

³⁷ ELSconsults (2020) End evaluation, WG4C project 2016-2020.

³⁸ Plan (2019) Annual report 2019, p. 11.

partner, the trust stems from PFPs being local community members who sometimes, are survivors themselves. This is not always the case with health workers.

Further, the partner noted that PFPs would bring for discussion key issues experienced in their work to the wider community. This fostered a wider community response and ownership to provide support to those in need.³⁹

The strong community foundations enabled referral mechanisms to operate in both directions with PFPs sign-posting to authorities and authorities also referring cases to PFPs.⁴⁰ Referral mechanisms fed directly into the programme's work with the ABC courts and the police where there is some anecdotal evidence that GBV case referrals were increasing and more appropriately dealt with.⁴¹ Whilst evidence of institutional change is weak, the evidence that the community component of the referral mechanism functions well is much stronger. The programme has an opportunity in NAP IV to address the institutional barriers more as the community mechanisms are in place and functioning well.

ToC Outcome 2: Women show leadership in peacebuilding

Missing middle: Forming relationships and networks for individual and group empowerment



Another planned programme outcome was to improve women's leadership in peacebuilding and increase local actor support for this. An important step towards developing this leadership is building the confidence and support network of individuals and groups of women at different social levels (e.g. grass roots women and members of parliament). Increasing linkages is a 'cross-cutting' component in the NAP III ToC yet warrants unpacking within this pathway of change to

³⁹ KII 3 with implementing partner.

⁴⁰ KII 1 with implementing partner.

⁴¹ Including local courts and police gender desk more responsive and sensitive to GBV cases: KII 13 women's group respondent, FGD, and Plan (2021) Final report 2016-2020.

⁴² FGD, KII 13 women's group member.

trace the journeys women travel on the routes to empowerment and leadership.

The formation of women's groups such as the Vision Women advocacy group in Nimule and advocacy group in Ganyiel are indicative of growing solidarity and support networks amongst women at the local level as ignited by the programme activities.⁴² In Nimule, the programme targeted existing women's group such as the Kokura and Alezoka women's village saving and loans groups. Following training through the programme, these groups along with others decided to form the advocacy group Vision Women and were further supported by a lawyer provided by Plan to register with the government.⁴³ As women from the Vision Women network asserted in the FGD, joining women together and raising their awareness of their rights has helped them "fight for themselves rather than others fighting for them".⁴⁴

The mutual support and empowerment the programme engenders is not limited to the local level within the project. Several key informants noted how powerful the national women's forum was on connecting grassroots women with elite women.⁴⁵ These events allowed women from different spheres to really 'know each other'⁴⁶ for the first time and communicate their differing experiences.

ToC Outcome 3: Women and girls are empowered and claim their rights

Missing middle: Consciousness raising in individuals, including men and boys and changing dynamics in the home



Shifting harmful social gender norms is another key intended outcome of the project. It is well understood that changing norms is a long-term endeavour and not something that can occur within a short four-year programme. Yet, incidences of change at individual and household levels signal the

⁴³ Plan (2021) Final report 2016-2020.

⁴⁴ FGD with women and youth group members and training participants Nimule.

⁴⁵ KII 2, 4 & 5 implementing partners.

⁴⁶ KII 5 with implementing partner.

beginnings of this shift. The programme has some evidence of this happening.

In several of the KIIs women and men noted how the training and programme activities had changed their attitude thereby changing relationships within the home. One women's group member noted that:

*"My husband used to be rebellious, fight with me in front of my children and does [sic] not provide food. Then, I got courage and ways to fix our differences."*⁴⁷

Another male participant of UNSCR 1325 training noted that:

*"I restrict my wife from attending meetings... (she) missed this very workshop that changed my life. I will give a maximum chance for my wife to attend meetings with rest of the other women."*⁴⁸

These examples are anecdotal and not robust evidence, yet they are an indication that changes within the home are a good place to look further for seeds of social norm change.

Similar changes can be seen amongst youth group members - an important step towards developing young people into lifelong advocates for gender equality. One member of the children and young people's parliament noted:

*"We can question harmful gender norms that are encouraged by our parents."*⁴⁹

Voices from the Plan reports and from KIIs and FGDs speak about the champions of change methods aimed at youth and the resulting set-up of the children and young people's parliament in changing the role of youth in the community. Anecdotal stories of attitude change and empowerment amongst young people could be better captured to understand how these changes connect with better protection and representation of women and girls.

Unintended outcomes and barriers

The programme came across several positive and negative unintended outcomes as well as barriers to implementation. This includes:

Positive unintended outcome: Addressing the stigma on mental health

The WG4C programme had ambitions to change harmful gender norms. Yet, another important change elicited by the programme that isn't mentioned in the ToC are community attitude changes to mental health. There is anecdotal evidence that the work through the PFPs also contributed to reducing mental health stigma within communities.⁵⁰ The ways in which social norms around mental health and gender intersect would be an interesting component of the programme to interrogate further.

Challenges to upholding 'do no harm' principles

Working within conflicted contexts poses further risks to programme implementers and participants, including threats to well-being and safety.

For example, the PFPs were initially recruited as community volunteers, however, it became apparent that the role required significant time and skill. There was an issue raised with PFP self-care and their remuneration for the role. Healthnet adapted to this by introducing a self-care module in the training and arranging remuneration where possible.⁵¹ However, there are ongoing risks to be

⁴⁷ KII 15 with women's group member.

⁴⁸ FGD with women and youth groups and training participants Nimule.

⁴⁹ FGD with women and youth group members Nimule.

⁵⁰ KII 14, psycho-social focal point.

⁵¹ KII 3 implementing partner.

checked on well-being and voluntary labour exploitation.

The safety of participants and implementing partners was also raised as an area of concern.⁵² There was an incident of a partner staff member that required evacuation to the Netherlands due to security concerns due to the sensitivity of the themes the projects deal with.⁵³ Backlash from men in the communities was also a concern. One group of men threatened to divorce their wives who were participating in the programme.⁵⁴ These risks should be managed and mitigated throughout the programme lifecycle.

Contextual barriers to programme success

A key barrier limiting impact was the active armed conflict that broke out in 2017. The WG4C programme was being implemented in extremely volatile environments where beneficiaries, staff and partner organisation were displaced due to armed conflict.

The infrastructure within South Sudan was cited as a further barrier to programme implementation and success. For example, for specialist mental health referrals, provision is only provided in one hospital in Juba that has a limited capacity to take 12 persons at a time.⁵⁵ There is no government-provided service provision for transport, for example, to police stations. One respondent from the police gender desk noted that the special gender unit doesn't have transport to be able to respond. There is also no support to cover legal or medical fees which also blocked referral pathways.⁵⁶

Also on the referral pathways, the turnover of staff was noted to be an issue. Including turnover; amongst PFPs, amongst the police and amongst development actors who only provide short-term services for the duration of a project.⁵⁷ This required further resources to re-train new staff and adapt.

Strengths and weaknesses of how the programme was implemented

There are several learnings from the way in which the programme was implemented.

Efficiency in partners sharing their technical knowledge and skills

It was identified in reporting and in the KIIs that Healthnet TPO's approach of technically capacitating both community members and implementing partners across geographical areas was successful in providing cohesion across programme strands. It is described as the 'red wire throughout the programme'.⁵⁸ There was a missed opportunity in not sharing further other partner expertise and learning within the consortia. This led to the programme in the Unity and Lakes where PAX and AMA were leading the programme looking quite different to the programme in East Equatoria where Plan International were leading the programme.

Power-dynamics and the consortia model

The structure of the consortia in the South Sudan programme offers particular learning around power-sharing within an alliance made up of international, national and diaspora partners.

The partnerships were described by partners as an uncomfortable marriage initially.⁵⁹ The challenge of varying partner capacities and unclear responsibilities created a tension throughout implementation. This was laid bare by a safeguarding incident with a local partner which was difficult to manage from the Dutch pen-holder side as direct oversight of partner was through another consortia member. This incident highlights the requirement to think through the risks and trade-offs between power-sharing and required oversight and management and, what measures need to be in place to create cohesion and consensus with consortia partners from programme inception.

⁵² KII 3 and KII 2 implementing partners.

⁵³ KII 2 implementing partner.

⁵⁴ KII 3 implementing partner.

⁵⁵ KII 3 implementing partner and KII 1 implementing partner.

⁵⁶ Plan (2021) Final report 2016-2020.

⁵⁷ KII 3 implementing partner.

⁵⁸ Plan (2021) Final report 2016-2020.

⁵⁹ KII 2, 4 and 5 implementing partners.

Prospects for sustainable change

In 2020 the programme focus turned to sustainability as requested by the MFA. The programme developed an exit plan which is something other partners can learn from. However, such exit plans should be created from the beginning of a programme. The strategy for sustainability in the programme proposal insufficiently explains the ways in which results will be sustained or programme funding could be continued.

In 2020, there was a shift away from focusing on individuals to groups, for example improving governance structures of women's groups and linking them with other actors so they endured beyond the lifetime of the programme.⁶⁰ Also, PFP supervisors were put in place who were made up of implementing partner staff members to continue this component once funding ended.⁶¹ If sustainability had been on the agenda from the beginning activities like this would have been in place over a longer period and thereby more effective. Much of the programme activities continue through the new 'leaders of peace programme' in NAP IV, as due to Covid-19 restrictions on gathering, and the freedom of movement, it was impossible to implement the exit strategy as planned, which limited the intended focus of moving from the individual level to the group level

Key lessons learned

- i. The South Sudan NAP III programme made significant progress towards outcome-level change amongst individuals and within communities – and to a much lesser extent at institutional level. However, this change was not captured and forms a 'missing middle' of evidence. Looking forward, partners should consider how these changes can be measured to tell a better story of change.
- ii. The safety of participants and partners must be of utmost importance to the project implementation and contextual barriers and risks continually reviewed and mitigated as best possible.
- iii. Management processes within the consortia and partner relationships should be prioritised to maximise efficiencies and ensure robust safeguarding mechanisms are in place.

⁶⁰ KII 2 implementing partner.

⁶¹ KII 1, 2 and 3 implementing partners.

Gender Norm Change: Thematic case

Gender norm change

Changing **harmful gender norms** was one of the specific objectives of the Third Netherlands Women, Peace and Security National Action Plan 2016-2019 and 2020. Work on this was integrated in different ways into all of the eight country-level programmes, ranging from dedicated curricula to public awareness campaigns. Whilst transformative impacts were achieved at different levels, lack of robust tracing of change made it difficult to ascertain how widespread and sustainable these were.

Introduction

One of the three specific objectives of the Dutch WPS NAP III was the **decrease of harmful gender norms**, and to achieve this, the following three pathways to change were formulated in the NAP's overall theory of change (ToC):

- i. *Increased capacities, skills, knowledge and resources to contribute to gender equality.*
- ii. *Increased understanding of gender equality and gender norms.*
- iii. *Increased involvement of men and boys in the implementation of laws and regulations that contribute to gender equality.*

The specific objective is closely linked to the cross-cutting element of **increased involvement of men and boys as crucial actors**. It is, however, also closely connected to the other two specific objectives of **enhanced protection** and **increased leverage in conflict prevention, resolution, peacebuilding and relief and recovery**.

Many of the consortia combined the work on these specific objectives, for example by working on norm and attitude change with respect to the prevention of gender-based violence (GBV) with formal and informal

justice and/or security sector actors (e.g. Afghanistan, DR Congo, Iraq, South Sudan and Yemen). Another common approach was working with women and women's groups to increase their understanding of and capacity to work on gender issues, including on GBV prevention and response as well as peacebuilding (e.g. Colombia, Iraq, Libya, and Syria).

The rationale for choosing this thematic focus is that it focuses on root causes that affect all other work on advancing WPS, but also one where the global evidence base of 'what works' is still evolving. Furthermore, all of the programmes integrated gender norm change components in different ways, allowing for a degree of comparison across different contexts and programmatic approaches. It should be noted that limitations of low-quality documentation and a low number of independent sources means findings are not robust evidence of outcome level change but instead programme success identified by programme partners and beneficiaries that are under-reported.

All of the key challenges which emerged in the document review were reflected in the norm change approach as well, and these are in part

interlinked. Thus, across the various programmes:

- i. There is a wide **variety of approaches** across and within the programmes.
- ii. Programming has had to be **flexible and adaptive** to challenging and dynamic contexts.
- iii. All programmes faced challenges in **evidencing impact**.
- iv. There was often a leap between activities and outputs and ambitious outcomes which has led to a **'missing middle'** evidence gap; and
- v. There is a need to consider **sustainability** more carefully.

Furthermore, the programmes had to deal with resistances and led to **unexpected positive and negative consequences**, highlighting a need to carefully **(re-)consider assumptions** around social norm change. These six issues will be discussed in detail below.

Variety in approaches



The ways in which the specific objective of social norm change was integrated into the programmes of the consortia differed greatly, and often multiple approaches were used simultaneously within a programme. These approaches ranged from comprehensive, dedicated curricula focusing on shifting norms, over more individualised coaching and accompaniment, to broader public awareness raising on gender norms (e.g. through arts, sports, radio and poster campaigns) and the integration of sensitisation on norms into other components of programmes.

Research on best practices on norm change, for example in relation to GBV prevention,⁶² has highlighted how individual-level change can best be sustained when it is supported by similar, reinforcing messaging coming from

⁶² Kerr-Wilson, A.; Gibbs, A.; McAslan Fraser, E.; Ramsoomar, L.; Parke, A.; Khuwaja, HMA.; and Jewkes, R (2020). *A rigorous global evidence review of interventions to prevent violence against women and girls*. Pretoria: South African Medical Research Council.

other sources, such as the media or from other community members. At times, there were such concerted efforts in the programmes to bring different approaches together and create synergies. For example, in the DRC programme, efforts to change attitudes of individual men in the communities through the Men Engage⁶³ approach and those of formal and informal justice and security sector decision makers were complemented by public awareness-raising campaigns and inviting the decision makers to workshops held by women's rights organisations.⁶⁴ At other times, however, the individual norm change components did not work in concert, reducing possible synergies.

The overall ToC did not specify which harmful norms would be targeted for transformation, and in part the country programmes did not specify which these were. The degree to which the approaches were indeed transformative varied. As a number of respondents highlighted, in a number of the programmes the focus was more on raising general awareness about women rights, women's participation, GBV, understanding existing laws and implementing these rather than fundamental change. Especially some of the larger public awareness-raising campaigns kept to a more general level of messaging while coaching and extensive training programmes could go further.

Need for flexibility and adaptation



A number of different factors and dynamics required the programmes to adapt their approaches. The most impactful of these external factors were the Covid-19 pandemic and changes in the respective security situation. However, some of the programmes also had to shift their focus in terms of the intended target audiences (see also below on resistances and unexpected outcomes). For

⁶³ This should not be confused with the global network of pro-feminist organisations working with transforming masculinities called MenEngage

⁶⁴ Mawawo, Josée and Nzotsi Paluku (2020) Evaluation externe finale du programme MAnU «Mwanamke, Amani na Usalama», interviews with DRC consortium members, 14 September 2021.

example, in Yemen, the geographical focus of the implementation was shifted to another governorate due to shrinking space for CSOs;⁶⁵ in Iraq, there was a successful shift in the ministries which the programme primarily engaged with;⁶⁶ and in South Sudan, unexpectedly strong resistance by men in the communities required greater engagement with male gatekeepers and community leaders.⁶⁷ In several of the consortia, there was also a need to engage internally with gender norms, as implementing partners had differing views on particular social norms and on how radical or not their approaches could or should be.

The overall approaches of shifting attitudes on gender norms were also adapted to the particular audience, using a variety of approaches seen as appropriate for these. For example, while participatory and innovative methods such as role play or arts-based approaches that are more pedagogically effective can be used with communities, engaging with officials required more formalised and didactic lecture-style approaches. In Colombia and Libya, approaches were also adapted to address the different lived realities, for example of indigenous and ethnic minority women.

Tracking impact



Tracing impact for gender norms change is challenging, requiring ideally mixed-methods, some degree of monitoring attitudes and practice before and after interventions, as well as longer-term monitoring to determine medium-term impacts. Often, however, baseline data on harmful gendered norms and practices in the programmes was missing. Problematically, it was unclear at times which norms were being targeted and why, and there was often little to no data to evidence sustained impact of the interventions beyond individual stories of change. An emerging

good practice in this respect that was used in Colombia, DRC and South Sudan by some of the implementing partners was to work with intended beneficiary communities to determine which social norms were seen as being the most harmful.⁶⁸

Rather than tracking impact, the focus in the reporting was often more on the activities and outputs themselves. In some cases, this was due to a lack of capacity of local implementing partners in collecting more impact-focused data, and in other cases the MEL systems were designed and resourced with mainly quantitative indicators in mind. The tracing of impact is especially challenging in terms of more diffuse awareness-raising activities, where attribution and contribution are difficult to impossible to trace, even with significant additional resources. Individual level impact can be traced easier, for example through pre- and post-testing, but this is susceptible to desirability bias and may only be capturing a short-term change in increased knowledge rather than a fundamental, sustained shift in attitudes and practices.

The degree to which it makes sense to invest large amounts of resources into trying to capture elusive social norm change is a valid discussion that the implementing partners and the Dutch MFA should have. Given the already tight funding environment for women's rights organisations and the pre-existing reporting requirements on local partners, shifting resources into more complex and onerous MEL system would not be the answer. However, if there is a will to capture potentially important but diffuse shifts in norms and attitudes, also as an early warning marker if these become more restrictive, there is a need to invest in developing more qualitative data collection methods. These should be as light as possible, and not be driven by expectations of having to show success.

⁶⁵ KII, Yemen consortium partner, 06 September 2021.

⁶⁶ KII with Iraq consortium member, 04 October 2021.

⁶⁷ Plan International Nederland (2020). Women and Girls For Change - Building sustainable peace and gender equality in South Sudan. Final Report.

⁶⁸ In DRC, for example, these were identified as attitudinal barriers to women's meaningful participation, women's lack of access to land and inheritance, as well as early marriage. Interview with DRC consortium members, 14 September 2021 and Colombia/South Sudan consortium members, 05 October 2021

An elaborate example with built-in elements of monitoring and accountability is the Men Engage approach implemented by CARE in its DRC programme, which draws on similar approaches employed elsewhere.⁶⁹ In it, men choose to undergo a transformative change process and are monitored by their peers, and if they pass the trial period successfully, they are ‘inducted’ into the group in a public ceremony. The men then develop and implement local-level action plans for furthering gender equality.

Missing middles



As discussed in the main report and country case studies as well, there is often a ‘missing middle’ and a jump from one level to the next in the reporting on the gender norm change approaches of the programmes. As discussed above, the reporting tends to focus on activities such as workshops, a training session for police or justice actors on gender norms, or a public awareness campaign. There is an implicit assumption in much of the reporting that this participation in and of itself firstly does constitute gender norm change, that secondly this change from a workshop or training will sustain itself over time, and thirdly that the individual change will lead to changes in the family, community, respective state institution and so on.

“In contrast to implicit, linear assumptions in the reporting, all respondents were fully aware of the non-linearity of change.”

In contrast to these implicit, linear assumptions in the reporting, all of the respondents interviewed for the case study were fully aware of the non-linearity of change, the real-life challenges faced in enacting sustained change in highly fluid contexts, the resistances to change at

community level and the need for long-term, continued and repeated engagement.

This ‘missing middle’-gap between the high-level ambitions of the ToC and what is implemented (and what is possible) in reality also creates a conundrum in the reporting. The individual-level change, and the courage this takes, should rightly be celebrated, as should important outcomes such as the increased self-esteem of women, of fighting for and achieving to set up meetings between women and recalcitrant power holders or of ministries adapting individual policies. These are often hard-won, hard-earned and real victories. However, they are often well short of the higher-level impact formulated in the ToC.

Sustainability and longer-term impact



The missing middle also points to questions of the sustainability of gender norm changes. As highlighted by respondents, social norm change requires both individual change and changes in society more broadly, with especially the latter being a long-term process. Some of the respondents were quite explicit in that the changes they had been able to enact were at the individual level only, or that what had been achieved at the community level so far was raising gender awareness, rather than a transformation of harmful norms.

In DRC and Colombia, the gender norm change approaches were in part able to ‘cascade’ impact, as approaches were taken on by individuals willing to go through a personal transformation process and/or programme activities led to women setting up their own groups at the local level.⁷⁰ A good example of successful, sustained engagement that opened up a previously reluctant

⁶⁹ Mwanamke, Amani na Usalama (MANU) - Femmes, Paix Et Sécurité (2020). “Quels rôles jouent les hommes et les garçons engagés dans la promotion de la participation des femmes et filles dans la prise de décisions autour de paix et sécurité?” Learning Brief. While these approaches have proven successful, they do also carry the risk of centering male agency and sidelining women’s roles (El-Bushra, Judy; Myrntinen, Henri and

Naujoks, Jana (2014). *Re-Negotiating the ‘Ideal’ Society: Gender and Peacebuilding in Uganda*. London: International Alert.

⁷⁰ KIIs with two Colombia consortium members, 05 October 2021 and three DRC consortium members, 14 September 2021.

institution to engaging with civil society on WPS issues was the work with the Ministry of the Interior in Iraq, where ‘the key was to make the Ministry part of change.’⁷¹

The Covid-19 pandemic in the final year of implementation and the subsequent need to suspend many of the in-person activities had an impact on the gender norm change work of the programmes as well. KIIs with implementing partners indicate that there is also a risk that some of the gains made in this respect prior to the pandemic have been rolled back, as has happened elsewhere as well as women’s multiple burdening has increased in its wake.⁷²

Resistances and unexpected impacts



The gender norm change work encountered resistance in a number of countries, and in Libya it was deemed unsafe to push for it in some locations and on some issues in Yemen.⁷³ Programmes often struggled with culturally and religiously embedded notions of appropriate gender roles, norms and power dynamics, and gender relations were often seen as a zero-sum game.⁷⁴ Resistance to the work came from communities in at least Colombia, DRC, Libya, South Sudan, and Yemen, including from husbands of intended beneficiary women but also from other family and community members. In South Sudan, the programme sought to counter this by engaging more actively with male community leaders.⁷⁵ In Colombia and DRC, men participating in positive masculinities-programmes reported facing backlash and ridicule from peers, family and community members. In some cases, such as with justice sector actors in Iraq, the programme was also faced with an unwillingness of the intended audience to engage with social norm change.

⁷¹ KII with Iraq consortium member, 04 October 2021.

⁷² KIIs with Colombia consortium member, 05 October 2021 and with South Sudan consortium member, 28 September 2020

⁷³ KIIs with Libya consortium member, 16 September 2021 and Yemen consortium member, 06 September 2021.

⁷⁴ Namely where any gains by women are automatically a loss to men, and vice-versa (Interview with consortium partner, 28 September 2020).

Working on shifting gender norms, and what this entailed, also required conflict management within the consortia in at least three cases, as some implementing partners had more radical views than others.

There were however also unexpected positive impacts, such as in Colombia and DRC where the gender norm work attracted more people who wanted to participate than had been expected. There is, however, a risk that the norm change work, especially on positive masculinities, tends to attract mainly those who are already open to change, rather than those who are most invested into maintaining harmful norms.

Assumptions in approaches



Emerging evidence on successful norm change approaches points at a need to reverse the common assumption that simply increasing knowledge about unequal norms leads to a change of attitudes and thereby a transformation of practices and norms.⁷⁶ This kind of a linear approach to change was evident in some of the approaches, which relied on, for example, raising-awareness on gender norms without much follow-up. Some of the approaches in Colombia, DRC, Libya, and South Sudan did, however, take a more comprehensive approach which is more likely to lead to sustained change. Work on norm change has also highlighted the need to have individual change processes supported by broader societal change and by efforts to reduce factors which counteract positive change.⁷⁷ While some of the programmes, notably Colombia, DRC and South Sudan, did work in ways in which individual change messages were reinforced through other approaches, there was little in the way of work on structural factors, such as for

⁷⁵ Plan International Nederland 2020.

⁷⁶ Kerr-Wilson et al. 2020; Flood, Michael (2019). *Engaging Men and Boys in Violence Prevention*. London: Palgrave Macmillan.

⁷⁷ *Ibid.* These factors can, for example, be economic stress factors and other forms of structural violence which can contribute to frustrations, negative coping mechanisms and of seeking refuge in strict gender norms.

example integrating economic empowerment into the work.⁷⁸

One of the unstated assumptions around gender norms which was tested in some of the consortia was that there was a joint vision on which gender norms were indeed harmful and necessitated change. Another unspoken assumption, which is also evident in the formulation of the ToC, is that men and boys need to be 'brought into' the work on gender equality. This ignores and renders invisible the fact that men and boys are *already* playing central roles in defining and policing gender norms in a given society, and defining the limits of how far work on gender equality is allowed to progress. A final assumption which had to be revisited in several programmes was that of willingness of state actors – and in part of international actors – to engage in a positive way on WPS issues. As discussed above in the section on flexibility, this necessitated a shift in the approaches of the programmes.

Key recommendations

Some of the key recommendations arising from the evaluation of the gender norm change objective are:

- iv. There is a need to clearly define which norms a programme wants to work on and how, and find consensus on this within the consortium.
- v. Understand what is possible in terms of norm change in a given context and ensure that there is a mitigation of risks or backlash to intended beneficiaries and implementers.
- vi. Work more on addressing root causes, structural factors and supporting individual change through broader societal change.
- vii. Engage directly with resistances to change from women and men, and engage with issues of masculinities more thoroughly.
- viii. Understand 'missing middles' better so as to be able to better trace pathways of change, improve programming and increase positive impact.

⁷⁸ In Libya and Colombia, this did however happen at the local level, in part through the initiative of beneficiaries themselves,

interviews with Colombia consortium members, 05 October 2021 and Libya consortium member, 16 September 2021.

Table 1: Social norm change approaches in the country programmes

Country	Increased capacities, skills, knowledge and resources to contribute to gender equality	Increased understanding of gender equality and gender norms	Increased involvement of men and boys in the implementation of laws and regulations that contribute to gender equality	Comments/Examples of some approaches used
Afghanistan	Yes	Yes	In part (engagement with justice and security sector)	Focus on shifting norms and attitudes within security and justice institutions, engaging with gatekeepers.
Colombia	Yes	Yes	Yes	Range of approaches, including several local partners engaging with positive masculinities
DR Congo	Yes	Yes	Yes	CARE working with 'Men Engage' curriculum, but also broader awareness-raising and training
Iraq	Yes	Yes	In part (engagement with Ministry of Interior)	Building capacity of women to engage with justice and security sector, work on changing norms inside security sector institutions
Libya	Yes	Yes	No	Included emphasis of changing norms and attitudes within consortium.
South Sudan	Yes	Yes	Yes	Champions of Change curriculum, with separate curricula for boys/young men and girls/young women; also engaged male powerholders, peace committees and traditional gatekeepers.
Syria	Yes	Yes	No	Included work on norm change within consortium.
Yemen	Yes	Yes	In part (advocacy at governorate level, including working with religious and community leaders)	Main focus was on building CSOs', women's and youth capacities for participation, including using arts- and sports-focused approaches.



Itad is a global organisation. Our strategy, monitoring, evaluation and learning services work to make international development more effective. We generate evidence on important issues – from malnutrition to migration – to support our partners to make informed decisions and improve lives.

itad.com

 [@ItadLtd](https://twitter.com/ItadLtd)

mail@itad.com

Itad Ltd

Preece House, Davigdor
Road, Hove, East Sussex,
BN3 1RE, UK

+44(0)1273 765 250

Itad Inc

1110 Vermont Ave NW,
Suite 500, Washington, D.C
20005, USA

Tel: +1 (301) 814-1492

Annex 2 – Data collection tools, case study approach and sampling

This section describes what data collection tools we used for both primary and secondary data collection.

Evidence assessment framework

The evidence assessment framework is our data capture framework that enables on-going data collation, analysis, and synthesis throughout the evaluation, which in turn supports the iterative process of triangulation, cross-validation, and the weighting of the relative strength evidence collected. The framework organised evidence in relation to the ToC specific objectives, pathways of change, cross-cutting issues and tagged the evidence for each EQ. The headings we used in the evidence assessment framework are depicted in figure 1.

		Evaluation questions 1 & 2: Impact & effectiveness										Evaluation question 3: Coherence			Evaluation question 4: Sustainability		Evaluation question 5: Efficiency		Evaluation question 6: Learning	
		Evidence of effective SO1: Enhanced protection			Evidence of effective SO2: Decrease in harmful gender norms			Evidence of effective SO3: Equal leverage			Pathways to change	Unintended outcomes	How far were NAP projects aligned to WPS policy frameworks?	How far were NAP projects aligned to National & decentralised policy?	How far were NAP projects aligned to target groups?	How far were NAP projects aligned to the changing context?	Evidence that benefits will be sustained	Evidence of the use of policy reforms, attitude/behaviour change of target groups, capacity building/knowledge/sustainability?	Evidence that additional year (2020) contributed to impact improvement to plan? Context analysis used? Quality of M&E data, finances and resource mobilization, staffing next plan	How did project design and implementation affect sustainability impact VIM? Lessons learned
Doc no.	Year	Country	Programme	Increased capacity & awareness in security & justice sector including economic response needs, access to services (gender/primary/psychosocial)	Attitudes & beliefs: increased understanding of GBV including working with and around GBV and protection of survivors	Law & policy: improved implementation of partner op.	Increased capacity and awareness: leadership	Attitudes and beliefs: increased understanding on gender inequality and gender norms	Law and policy: increased involvement of men and boys in law enforcement	Capacity and resources: increased in number of change in GBV risk prevention, AHA, employment	Attitudes and beliefs: increased political will & awareness for inclusive conflict prevention, peace processes, AHA	Law & policy: Local strategies incorporated into policy & regulation, improved inclusion of women in conflict prevention	Increased dialogue between local, national, regional and international	Increased involvement	Evidence of needs assessment, gender/conflict analysis	Evidence of adaptation context	Evidence of gender equality/sustainability	Evidence of gender equality/sustainability	Evidence of sustainability plan & next plan	Lessons learned

Figure 1: Data capture headings against EQ criteria

Evidence strength rating tool

Evidence collated within the evidence assessment framework was assessed and ranked using the evidence strength rating tool. This tool ensures a uniform understanding amongst the team on what constitutes as strong evidence and ensures a consistent approach. The matrix drew from IOB criteria pertaining to evidence strength (e.g. number of independent information sources, triangulation, and bias) and Itad’s experience in conducting evidence strength reviews. Evidence was rated in relation to its strength using a red, amber, green (RAG) scale. Moderate and weak evidenced findings will still be discussed in the report with a caveat on strength of evidence.

Criteria for evaluation strength is presented in Figure 7 below:

Category	Criteria
Strong	<ul style="list-style-type: none"> Data sources (KIIs, FGDs, implementing partner reporting, literature) allows for triangulation either across sources or stakeholders. Data sources are clearly referenced.

<p>Strong evidence & data coverage</p>	<ul style="list-style-type: none"> • Biases are accounted for. • Data coverage is good and covers all aspects under investigation.
<p>Moderate Some evidence & adequate data coverage</p>	<ul style="list-style-type: none"> • Confidence is reduced by shortcomings in relation to triangulation. • Confidence is reduced in relation to concerns around bias, knowledge, position, reflexivity or analytical capacity of informants or reliability of source. • Data sources clearly referenced. • Data coverage is sufficient but patchy across some areas under investigation.
<p>Weak Limited evidence and poor data coverage</p>	<ul style="list-style-type: none"> • Evidence comes from a small number of sources, little to no triangulation. • Major concerns on the position, knowledge, reflexivity or bias of informant or reliability of source. • Data sources not clearly referenced. • Data coverage is limited across areas of investigation.

Figure 2: Evidence strength rating criteria

Key informant interviews (KIIs) and focus group discussions (FGDs)

KII and FGD semi-structured interviews were the main tool for primary data collection for the case studies. They were conducted, where possible, in person by our local expert team members and where not possible remotely. KIIs and FGDs were guided by lines of inquiry that are drafted following the document review. This semi-structured approach brings a number of strengths in terms of covering the range of topics relating to the evaluation matrix, while at the same time allowing the emergence of potential factors or causal pathways for explaining findings.

We conducted 22 KIIs were conducted in total (10 in Colombia and 12 in South Sudan) and one focus group in South Sudan with 8 participants.

Survey on the partnership model

We used an online survey to gather feedback from implementing partners in Colombia and South Sudan on the efficiency and effectiveness of the consortia partnership model. The survey was sent to 14 participants and had 12 responses. The survey was distributed using the online platform survey monkey and asked the following questions:

- 1) What was your role in the WPS NAP III programme?
 - a) Representative of Dutch NGO (NL-based)
 - b) Representative of Dutch NGO (country-based)
 - c) Representative of local NGO
 - d) Other:
- 2) How frequently did the consortium meet?
 - a) Never
 - b) Monthly
 - c) Quarterly
 - d) Annually
 - e) No regular schedule
 - f) Other:
- 3) What were the benefits of working in a consortium to deliver the programme? [open-ended]
- 4) What were the challenges of working in a consortium to deliver the programme? [open-ended]
- 5) Were any solutions found to the challenges? [open-ended]
- 6) Does the consortium still exist? [Y / N]
- 7) Is your organization still a member of the consortium? [Y / N]
- 8) Do you still keep in touch with other members of the consortium? [Y/N]

Case study approach

For the country case studies, the following lines of inquiry were used by local consultants for primary data collection:

- Can you provide an overview of the NAP III project and your role in it?
- Which of the three objectives of the NAP III did it work towards? [enhanced protection / decrease of harmful social norms / equal leverage in peace conflict prevention, resolution, peacebuilding, relief and recovery] Please describe the intervention's contribution. [EQ1, EQ2]
- Which of the three pathways of change did the project utilize to achieve results and how? [capacity & resources / attitudes & belief / law & policy] [EQ1, EQ2]

- What was the impact(s) of the intervention? [EQ1, EQ2]
- How did was the impact(s) measured? [EQ1, EQ2]
- What challenges were met and how were these overcome? [EQ3, EQ5, EQ6]
- What were key underlying assumptions and did these prove correct? [EQ1, EQ2, EQ3, EQ4, EQ5]
- How sustainable and ensuring were the impacts of the intervention? [EQ3, EQ4, EQ5, EQ6]
- Were there any positive or negative unintended consequences? [EQ1, EQ2, EQ5, EQ6]
- Who else would you recommend I speak with for this evaluation?

For the thematic case study the focus was on shifting harmful social norms that underpin gender inequality, which is a specific pillar of NAP III and is closely linked to the cross-cutting element of **increased involvement of men and boys as crucial actors**. Data collection for this case study consisted of KIIs with staff from the Dutch and local consortium partners by Henri Myrntinen, as well as a review of programme documents, especially in those cases where there was a fully formulated programmatic approach to harmful social norm change. The case study will also draw on insights from the critical reflection sessions with Dutch NGOs and local implementing partners.

Lines of inquiry for the thematic case study followed the questions below, with the relevant EQs they will contribute to answering in brackets:

- What approaches to shifting social norms proved the most successful, how and why? [EQ1, EQ2, EQ3, EQ4, EQ5, EQ6]
- What challenges and resistances were met and how were these overcome? [EQ3, EQ5, EQ6]
- What were key underlying assumptions and did these prove correct? [EQ1, EQ2, EQ3, EQ4, EQ5]
- How did the harmful gender norm change components relate to/engage with policy frameworks at different levels? [EQ1, EQ3, EQ4]
- How were impact and changes in social norms measured? [EQ2, EQ5, EQ6]
- How sustainable and enduring were the impacts? [EQ3, EQ4, EQ5, EQ6]
- Were there any positive or negative unintended consequences? [EQ1, EQ2, EQ5, EQ6]

1.1 Sampling

This section describes how we selected the participants and respondents for data collection in order to reduce bias.

Theory of Change workshop & critical reflection session

Participants for the theory of change workshop and critical reflection session were self-selecting. In order to ensure we had the *right person* (the member of staff with most knowledge on the programme) with the *right availability* (were not on leave) we requested that the Dutch consortia leads reached out to their partners to arrange attendance. For the ToC workshop we requested that Dutch consortia partner members attend as they have most knowledge and context in relation to the NAP III ToC. For the critical reflection session we requested at least one Dutch consortia member and one local partner join to include the voices of in-country implementers.

We understand this approach risks self-selection bias, however, the short evaluation timeline, language restrictions (some NGO partner staff do not speak English), staff turnover in the consortia organisations

(in some cases staff working on the programme had left) and reduced availability of participants due to summer holidays required that we be pragmatic.

To counter the self-selection bias, we ensured representation from all country programmes in both the ToC and critical reflection session. Those who were unavailable to participate in the sessions themselves either provided written feedback (such as the DRC & Libya programme for the ToC workshop) or had independent calls with the evaluation team (such as Libya programme for the critical reflection sessions).

Key informant interviews and focus group discussions

Respondents for KIIs and FGDs were also selected through purposive sampling followed by snowball sampling, which is the most practical approach within the scope of the evaluation. A low number of KIIs were conducted, which thereby not conducive to statistical sampling. Purposive and snowball sampling allowed us to target key stakeholders and access their networks in an efficient way.

Survey

Survey participants were again, purposively sampled. As the topic of the survey was the consortia partnership model for the country case studies, we reached out to all consortia partners involved in Colombia & South Sudan NAP III programme. Survey response rate poses the risk of self-selection bias, however, as the response group was relatively small and we were able to mitigate this risk by triangulating with findings from the KIIs.

For a full list of stakeholders consulted at each stage of the evaluation please see annex 7.

Annex 3 – Evaluation matrix

Presented below is the evaluation matrix, consisting of evaluation questions, sub-questions, indicators and judgement criteria, type of analysis, methods, tools, and data sources.

Evaluation questions	Topic	Indicators and judgement criteria	Analysis methods	Methods, tools and data sources
<p>1. What are the NAP programme contributions to the overall objective of WPS framework 2016-2019 & 2020 and what are the programme contributions to other positive and negative outcomes? (intended and unintended and sensitive to 'do no harm' principles)</p>	Impact	<ul style="list-style-type: none"> • Robustness of impact and outcome data • Evidence of change generated by the project that does not fall into the ToC either positive or negative • Clear definition of what constitutes as an 'enabling environment' 	Contribution analysis	<p><i>Module 2</i></p> <p>Meta-evaluation of project evaluations</p> <ul style="list-style-type: none"> • Document review including: Proposal 16-19, inception report, annual plans, annual reports, approval letters, MTR reports and other project documents, which will also include review of evidence of institutional take-up (if this is documented)
<p>2. What are the NAP project contributions to the WPS framework 2016-2019 & 2020 three specific objectives and what change happened along the causal pathway - did assumptions hold?</p>	Effectiveness	<ul style="list-style-type: none"> • Robustness of impact and outcome level evidence against the WPS framework three specific objectives (enhanced protection, decrease of harmful gender norms, and equal leverage) • Robustness of outcome and output level evidence against the three pathways of change (capacity & resources, attitudes and beliefs, laws & policy) • Robustness of evidence against cross cutting themes (increased linkages, cooperation, involvement of men) • Evidence of validation of assumptions or assumptions that did not hold 	Contribution analysis	<ul style="list-style-type: none"> • Evidence strength rating tool • Critical reflection sessions with representatives from lead NAP consortia members and local NGOs, which will also include questions on longer-term impact and institutional take-up <p><i>Module 3</i></p>

<p>3. How far were NAP projects aligned or responsive to:</p> <ul style="list-style-type: none"> • WPS policy frameworks • National & decentral policies • The needs of beneficiaries • The changing context 	<p>Relevance & coherence</p>	<ul style="list-style-type: none"> • Evidence of intentional alignment to the WPS 2016-2019 & 2020 strategy or national & decentral country policies in design and/or reporting • Evidence of context analysis conducted and/or a mapping exercise on national & decentralised policy • Evidence of needs assessment and/or gender-sensitive conflict analysis • Evidence of beneficiary feedback mechanisms • Evidence of adaptation of programming in light of a changing context e.g. Covid-19, conflict escalation, disasters, political/security developments and pushback against WR&GE in a broader sense. 	<p>Contribution analysis</p>	<ul style="list-style-type: none"> • Case studies: for South Sudan and Colombia country programme we will take a deeper look into what impact has been achieved, what learning there is to draw out and use process questions on efficiency and sustainability EQs. • Key informant interviews & focus group discussions with NGO partners of South Sudan consortium, Plan Nederland (consortium lead), PAX, Healthworks, Support Trust for Africa Development (STAD) and local implementing partners and Colombia consortium (ICCO, Mensen met een Missie and Healthnet TPO & 10 local partners. • Survey with Dutch consortia partners for thematic case study & any required follow up KIIs
<p>4. What evidence is there to show the benefits of NAP programmes will be sustained beyond the life of the programme and how far, and in what ways did actions in the additional year (2020) contribute to strengthening sustainability?</p>	<p>Sustainability</p>	<ul style="list-style-type: none"> • Evidence of institutional take up of project by government or other partners • Evidence of policy/regulatory reforms supported by project implementation • Evidence of attitude, belief & behaviour change of target groups • Extent of capacity building, knowledge development and network sustainability amongst CBOs and local orgs • Comparison of outcome & output achievements end of year 2019 and end of year 2020 	<p>Contribution analysis</p>	<ul style="list-style-type: none"> • Case studies: for South Sudan and Colombia country programme we will take a deeper look into what impact has been achieved, what learning there is to draw out and use process questions on efficiency and sustainability EQs. • Key informant interviews & focus group discussions with NGO partners of South Sudan consortium, Plan Nederland (consortium lead), PAX, Healthworks, Support Trust for Africa Development (STAD) and local implementing partners and Colombia consortium (ICCO, Mensen met een Missie and Healthnet TPO & 10 local partners. • Survey with Dutch consortia partners for thematic case study & any required follow up KIIs

<p>5. How did projects design and implementation affect achievements along the causal pathways of change, project sustainability and, value for money?</p>	<p>Efficiency & sustainability</p>	<ul style="list-style-type: none"> • The extent to which the project plan was implemented • Extent to which context analysis data was used • Quality of M&E data and extent to which this informed programming decisions • The extent to which the planned activities were budgeted and actually financed • Challenges to financial mobilization • Challenges in staffing • Relationship management between donor, consortia, local partners, and CBOs. • Level of resources allocated to interventions using evidence and proportionality • Evidence of sustainability strategy in programme design • Evidence of an exit plan 	<p>Contribution analysis</p> <p>Process evaluation components: to understand whether project activities were implemented as intended</p>	
<p>6. What lessons can be learned from the NAP projects and how far do achievements and lessons learned align with the broader international WPS evidence base?</p>	<p>Learning</p>	<ul style="list-style-type: none"> • Aggregation of 'what works' and key challenges/barriers across countries • Comparison of successful strategies and barriers with international literature on WPS 	<p>Comparative analysis</p>	

Annex 4 – Quick-scan results

This annex shows a breakdown of our ranking of each programme evaluation against the IOB criteria.

General info		Quality control of the evaluation		Description and background of the intervention			Objective and delimitation of the evaluation		Evaluation questions		Evaluation methodology											Results and Conclusions			Usefulness and readability of the evaluation report						
Country programme	Documents used	Pass/fail (23/26 must be 'sufficient' or 'good'?)	A reference group oversees the evaluation independent *	Evaluators are independent *	Description of the intervention	Description of the intervention (ToC) *	Validation of the assumptions underpinning the ToC *	Description of the objective	Delimitation of the evaluation	Choice of OECD-DAC evaluation criteria to be covered	Clear set of evaluation questions	The research design is clearly elaborated and shows how the research results will contribute to answers to the evaluation questions *	The methods are appropriate to evaluate effectiveness: attribution and / or contribution (if effectiveness is an evaluation criterion/question) *	The methods are appropriate to evaluate efficiency (if this is an evaluation criterion/question) *	The indicators or result areas are appropriate to capture the planned results along the different levels in the ToC *	Justified choice of sample, cases and information sources	The analyses are appropriate, given the chosen research design *	Summary of the methodology in an evaluation matrix	Sufficient independent information sources *	Triangulation of results from different information sources	Discussion of bias	Systematic, complete and transparent description of the collection of data and analysis *	Discussion of the limitations of the evaluation *	Conclusion answers research questions *	Conclusion follows logically from the research findings *	Validation of draft conclusions and intended users	Recommendations should be practical, objective and its intended users	The report is well readable, consistent, and includes a clear summary with useful and objective, evaluation questions, conclusions and recommendations			
Iraq	Final evaluation report & ToRs	9	insufficient	insufficient	insufficient	sufficient	insufficient	sufficient	insufficient	good	good	insufficient	insufficient	insufficient	insufficient	insufficient	sufficient	sufficient	good	insufficient	good	sufficient	insufficient	insufficient	insufficient	insufficient	insufficient	insufficient	insufficient	insufficient	insufficient
South Sudan	Final evaluation report	18	insufficient	insufficient	sufficient	sufficient	good	sufficient	good	good	good	insufficient	insufficient	sufficient	insufficient	insufficient	sufficient	good	insufficient	insufficient	good	good	good	good	good	sufficient	good	good	sufficient	good	
Libya	Final evaluation report, annexes and ToRs	14	insufficient	insufficient	sufficient	insufficient	insufficient	sufficient	insufficient	good	good	sufficient	insufficient	N/A	insufficient	sufficient	sufficient	insufficient	insufficient	insufficient	insufficient	insufficient	sufficient	sufficient	good	good	insufficient	sufficient	good	good	
Yemen	Final evaluation report and ToRs	12	sufficient	sufficient	sufficient	insufficient	insufficient	good	good	good	good	sufficient	insufficient	insufficient	insufficient	sufficient	insufficient	insufficient	insufficient	insufficient	insufficient	insufficient	sufficient	sufficient	insufficient	insufficient	insufficient	sufficient	insufficient	insufficient	insufficient
Afghanistan	Final evaluation report	15	sufficient	sufficient	sufficient	insufficient	insufficient	sufficient	good	good	good	sufficient	insufficient	insufficient	insufficient	sufficient	insufficient	insufficient	insufficient	insufficient	insufficient	insufficient	sufficient	sufficient	sufficient	sufficient	insufficient	sufficient	insufficient	insufficient	insufficient
Colombia	Final evaluation reports 2019 and 2020	12	insufficient	sufficient	sufficient	insufficient	insufficient	sufficient	insufficient	good	good	insufficient	insufficient	sufficient	insufficient	insufficient	sufficient	insufficient	insufficient	insufficient	insufficient	insufficient	sufficient	sufficient	sufficient	sufficient	insufficient	insufficient	insufficient	insufficient	insufficient
DRC	Final Evaluation report	13	insufficient	insufficient	sufficient	good	sufficient	sufficient	insufficient	good	good	insufficient	insufficient	insufficient	insufficient	insufficient	sufficient	insufficient	insufficient	insufficient	insufficient	insufficient	sufficient	sufficient	sufficient	insufficient	sufficient	sufficient	sufficient	sufficient	sufficient

Annex 5 – Theory of change & reconstruction

The NAP III Theory of change is depicted below with an impact, vision, overall objective and three specific objectives with the pathways of change towards these objectives expanded below. It also shows cross-cutting issues in the red arrows and types of interventions to achieve the objectives.

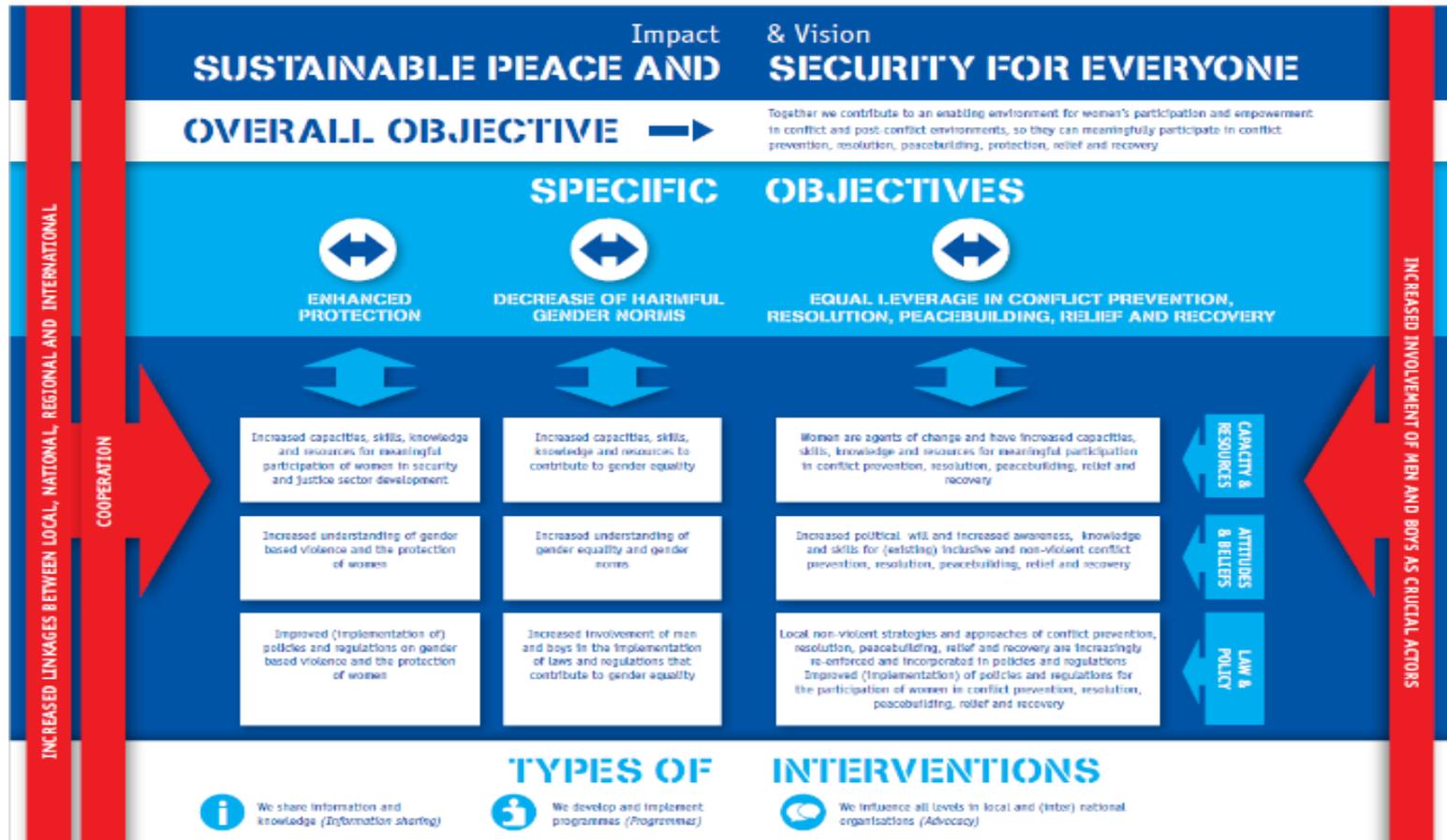
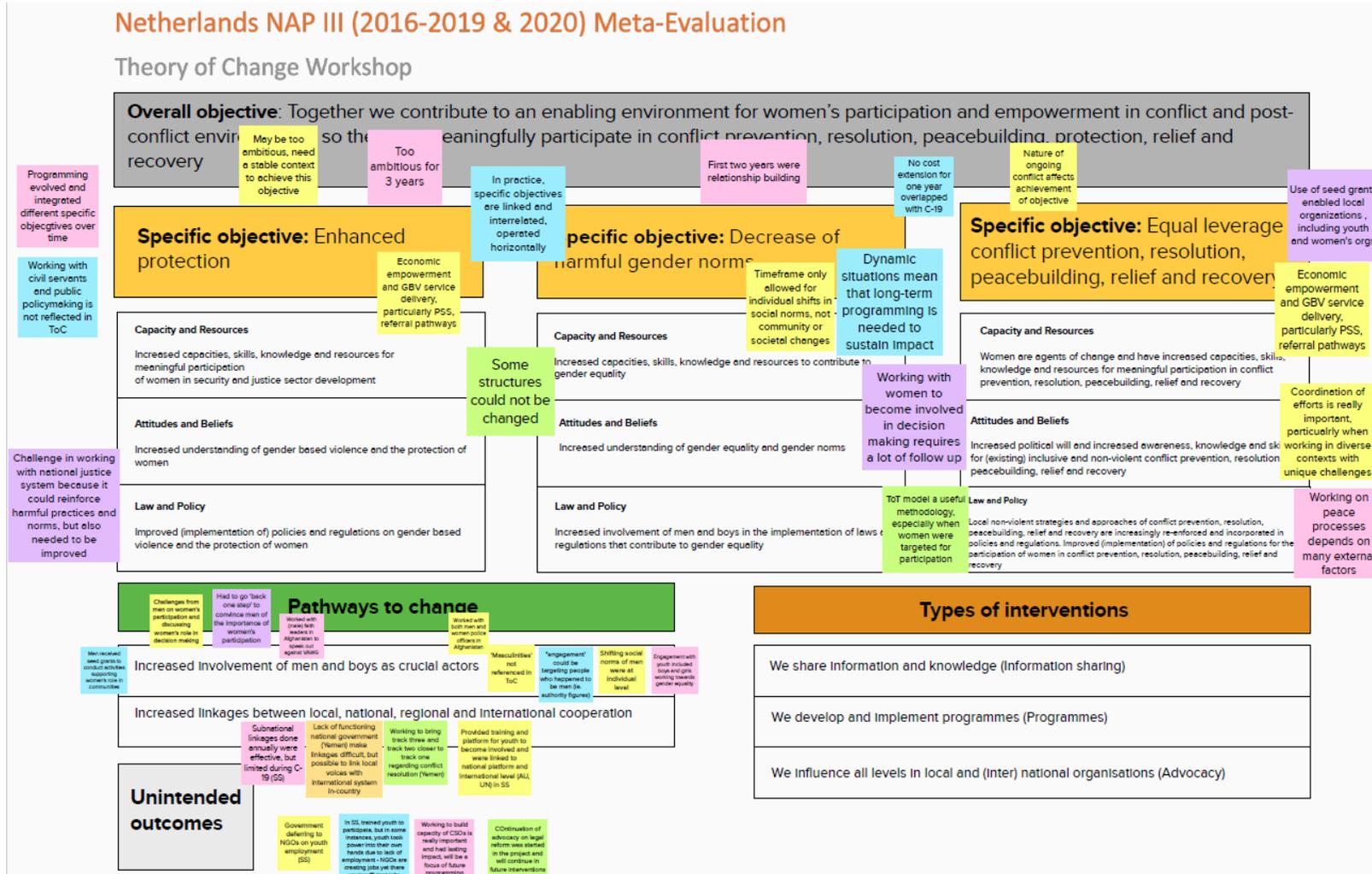


Figure 1: NAP 2016-2019 Theory of change

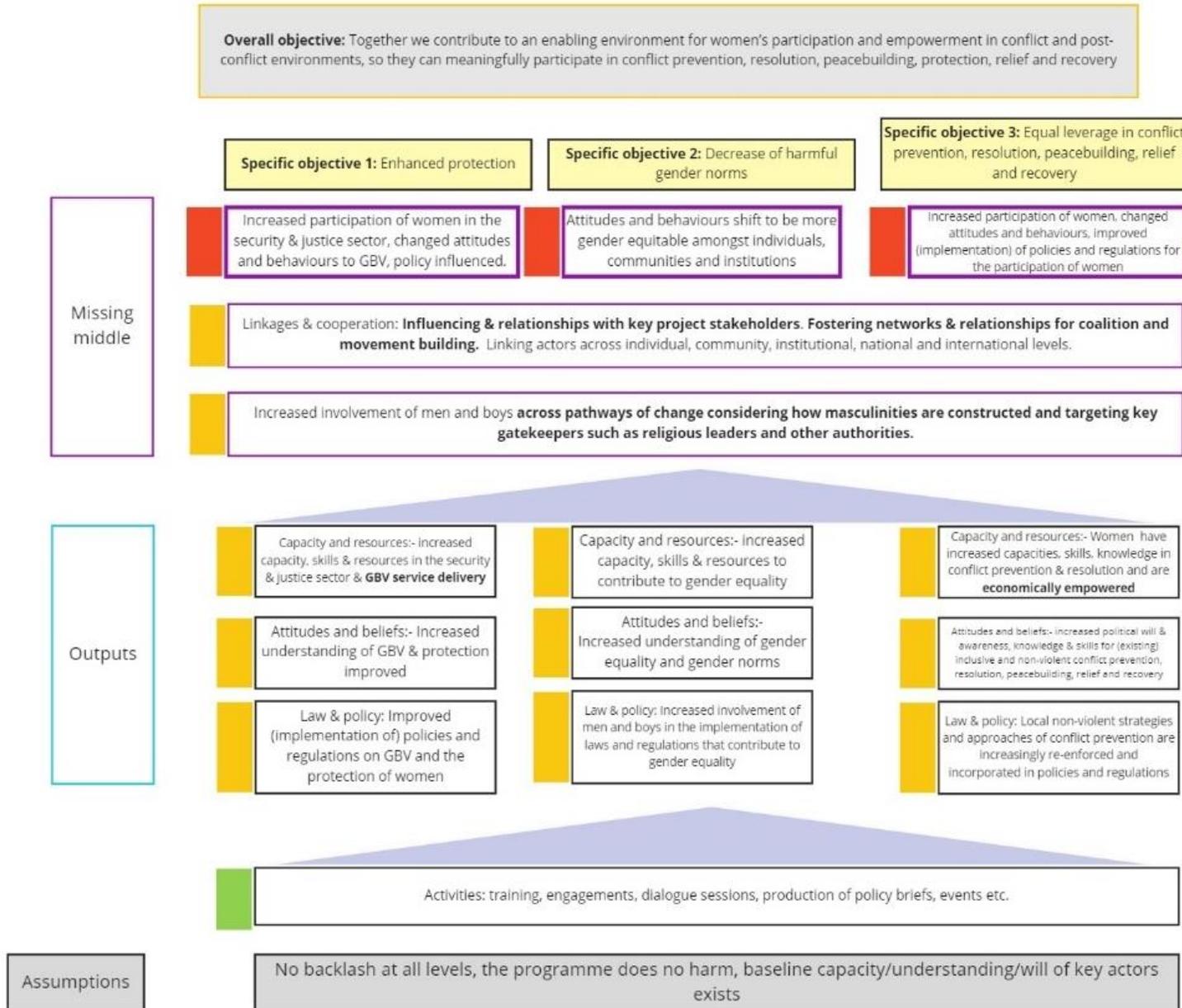
Theory of change workshop interactive whiteboard discussion

The evaluation team conducted a theory of change workshop with a group of Dutch implementing partners to discuss how far the ToC reflected the reality of how change happens across the country programmes. We annotated the ToC using an interactive whiteboard which is shown below.



Reconstructed theory of change

Following the ToC workshop, we reconstructed the ToC to include some of the key elements that were missing. See the diagram below.



Annex 8 – Bibliography

General

Basu, Soumita, Laura J. Shepherd and Paul Kirby (2020) *New Directions in Women, Peace, and Security*. Bristol: Bristol University Press

CDA Collaborative Learning Projects (2015), *From Principle to Practice: A User's Guide to Do No Harm*

El-Bushra, Judy; Myrtilinen, Henri and Naujoks, Jana (2014). *Re-Negotiating the 'Ideal' Society: Gender and Peacebuilding in Uganda*. London: International Alert.

Flood, Michael (2019). *Engaging Men and Boys in Violence Prevention*. London: Palgrave Macmillan.

Hamilton, Caitlin, Nyibeny Naam, and Laura J. Shepherd (2020). *Twenty Years of Women, Peace and Security National Action Plans: Analysis and Lessons Learned*. Sydney: University of Sydney

Jacevic, Mirsad (2019) *WPS, States, and the National Action Plans in The Oxford Handbook of Women, Peace and Security*, edited by Sara E. Davis and Jacqui True, Oxford: Oxford University Press

Kerr-Wilson, Alice; Gibbs, A.; McAslan Fraser E.; Ramsoomar, L.; Parke, A.; Khuwaja, HMA.; and Rachel Jewkes (2020). *A rigorous global evidence review of interventions to prevent violence against women and girls, What Works to prevent violence among women and girls global Programme*, Pretoria, South Africa

Kirby, Paul and Laura Shepherd (2016). 'The Futures Past of the Women, Peace and Security Agenda', *International Affairs* 92, no.2: 373–92.

Myrtilinen, Henri, Laura J Shepherd, and Hannah Wright (2020). *Implementing the Women, Peace and Security Agenda in the OSCE Region*. Vienna: Organisation for Security and Co-Operation in Europe

Myrtilinen, Henri. 2019. "Locating Masculinities in WPS." In *The Oxford Handbook of Women, Peace and Security*, edited by Sara E. Davis and Jacqui True, 88–97. Oxford: Oxford University Press

Trojanowska, Barbara K., Katrina Lee-Koo and Luke Johnson (2018). *National Action Plans on Women, Peace and Security: Eight Countries in Focus*. Canberra/Melbourne: Australian Civil-Military Centre/Monash Gender, Peace and Security Centre

Cross-programme documentation

1325 Dutch NAP partnership (2016) *The Netherlands National Action Plan on Women, Peace and Security 2016-2019*

1325 Dutch NAP partnership (2021) *Fourth Dutch National Action Plan on the implementation of United Nations Security Council Resolution 1325 and successive related resolutions 2021-2025*

Herweijer, Rosien (2019). Mid Term Review National Action Plan on Women Peace and Security 2016 - 2019

MFA (2018) NAP1325 Learning event narrative report, October 9-10 2018

MFA (2021) E-mail from Iris Loonen on no-cost extensions and further funding for programmes in 2020
13/08/2021

Programme level documentation

Afghanistan

ARM consulting (2021) Final Evaluation of Safh-e-jadid project

Mazars (2017) independent auditors report

MFA (2018) Oxfam Novib Afghanistan Annual Report approval letter

MFA (2019) Oxfam Novib Afghanistan Annual Report approval letter

Oxfam Novib (2016) Oxfam Afghanistan grant application 2016-2019

Oxfam Novib (2016) Oxfam Afghanistan grant application 2016-2019 Activity Appraisal document

Oxfam Novib (2016) Oxfam Afghanistan grant application 2016-2019 Logframe

Oxfam Novib (2016) Oxfam Afghanistan grant application 2016-2019 theory of change

Oxfam Novib (2017) Safhe Jaded Annual narrative report 2017

Oxfam Novib (2017) Safhe Jaded Annual narrative report 2017

Oxfam Novib (2017) Safhe Jaded Inception report

Oxfam Novib (2017) Safhe Jaded Inception report Annexes A-H

Oxfam Novib (2018) Annual plans 2018

Oxfam Novib (2018) Safhe Jaded Annual narrative report 2018

Oxfam Novib (2018) Safhe Jaded Annual narrative report 2018 cover letter

Oxfam Novib (2019) Annual plans 2019

Oxfam Novib (2019) Safhe Jaded Annual narrative report 2019

Colombia

ICCO (2016) M&E Plan

ICCO (2016) Partnership Agreement

ICCO (2016) Programme Proposal

ICCO (2017) Annual Narrative Report

ICCO (2017) Annual Operating Plan

ICCO (2017) Consolidated Financial Report
ICCO (2018) Annual Narrative Report
ICCO (2018) Annual Operating Plan
ICCO (2018) Consolidated Financial Report
ICCO (2018) Letter to MFA accompanying Report
ICCO (2019) Annual Budget Plan
ICCO (2020) Final Programme Report
ICCO (2020) Final Programme Report Annex A - Global Results
ICCO (2020) Final Programme Report Annex B - Infographic
ICCO (2020) Final Programme Report Annex C - Planning
ICCO (undated) Annex B - Baseline Report
ICCO (undated) Case Story 1
ICCO (undated) Case Story 2
ICCO (undated) Case Story 3
MFA (2016) Activity Appraisal Document ODA
MFA (2016) Notification Letter to ICCO
MFA (2018) Partner Results Summary
Rogriguez et al (2019) Programme Evaluation
Rogriguez et al (2019) Programme Evaluation Annex B
V&F (2019) Audit Report
V&F (2020) Audit Report
Valencia & Falla (2018) Audit Report

DRC

Cadre de suivi et évaluation «Mwanamke, Amani na Usalama» 2018 – Normes
Cadre de suivi et évaluation «Mwanamke, Amani na Usalama» 2018 – Participation
Cadre de suivi et évaluation «Mwanamke, Amani na Usalama» 2018 – Protection
CARE DRC (2020). “Quels rôles jouent les hommes et les garçons engagés dans la promotion de la participation des femmes et filles dans la prise de décisions autour de paix et sécurité?”
Dutch MFA (2016) Subsidiebeschikking WPS DRC, 31.10.2016
Mensen met een Missie, CARE and Stichting Tosangana (2016) Mwanamke, Amani na Usalama/Women, Peace and Security in DR Congo, Full Proposal.

Mensen met een Missie, CARE and Stichting Tosangana (2017) NAP 1325 DR Congo - Mwanamke, Amani na Usalama (MANU) Inception report, 15 May 2017

Mensen met een Missie, CARE and Stichting Tosangana (2018) Rapport Annuel 2017

Mensen met een Missie, CARE and Stichting Tosangana (2018) Plan Annuel 2019

Mensen met een Missie, CARE and Stichting Tosangana (2019) Rapport Annuel 2018

Mensen met een Missie, CARE and Stichting Tosangana (2019) Plan Annuel 2020

Mensen met een Missie, CARE and Stichting Tosangana (2020) Rapport Annuel 2019

Mensen met een Missie, CARE and Stichting Tosangana (2020). Mwanamke, Amani na Usalama/ Women, Peace and Security – Democratic Republic of Congo (DRC) Final report November 2016 - December 2019.

Mensen met een Missie, CARE and Stichting Tosangana (2021) Rapport Annuel 2020

Mensen met een Missie, CARE and Stichting Tosangana (2021) Women building the path for peace: Experiences from MANU programme Democratic Republic of Congo 2016- 2020 Validation Report

Mensen met een Missie, CARE and Stichting Tosangana (2021) Summary Report 2020

Mawawo, Josée (2018) Evaluation mi-parcours du programme MANU «Mwanamke, Amani na Usalama»

Mawawo, Josée and Nzotsi Paluku (2020) Evaluation externe finale du programme MANU « Mwanamke, Amani na Usalama »

Théorie de changement (undated) – Normes, Participaiton, Protection

Libya

Blommestein, Nele, Caroline van Koot and Fatima El Moussawi (2021) End-term Evaluation of Cordaid and Human Security Collective's Women and Youth as Bridge Builders Programme: Strengthening resilience in Libya 3rd NAP 1325 Women, Peace and Security Grant Instrument 2016-2019 and 2020

Blommestein, Nele, Caroline van Koot and Fatima El Moussawi (2021) Annex 1 – Final Evaluation Matrix

Cordaid and Human Security Collective (2021). Management response 'end of the program evaluation' - Women and Youth as bridge builders, strengthening resilience in Libya (Activity 29410) – issued in April 2021

Theory of Change (undated) – Libya country programme Dutch WPS NAP III

van Marwijk, Anne (2018) Mid-term Review: Women and Youth As Bridge builders programme Libya 2016 – 2019

Women and Youth as Bridgebuilders: Strengthening Resilience in Libya (2017) Inception Report Libya NAP 1325 Program

Women and Youth as Bridgebuilders: Strengthening Resilience in Libya (2017) Year Planning 2018
Women and Youth as Bridgebuilders: Strengthening Resilience in Libya (2019) Narrative Report 2018
Women and Youth as Bridgebuilders: Strengthening Resilience in Libya (2018) Year Planning 2019
Women and Youth as Bridgebuilders: Strengthening Resilience in Libya (2019) Narrative Report 2018
Women and Youth as Bridgebuilders: Strengthening Resilience in Libya (2019) Proposal 2020
Women and Youth as Bridgebuilders: Strengthening Resilience in Libya (2020) Narrative Report 2019
Women and Youth as Bridgebuilders: Strengthening Resilience in Libya (2021) Final Narrative Report 2020 & 2016-2020 (2-IN-1)

Iraq

Elephas consultancy (2021) Final Evaluation NAP Iraq programme
MFA (2017) Approval of PAX narrative and financial report
MFA (2019) Approval of PAX narrative and financial report
PAX (2016) Programme proposal
PAX (2016) Programme proposal, Annexes
PAX (2017) Annual report 2017 response to Minuza questions
PAX (2017) Annual report 2017, Engendering the Transition to Peace and Security in Iraq
PAX (2017) Annual report 2017, Engendering the Transition to Peace and Security in Iraq - Monitoring report planning
PAX (2017) Inception report, Annexes A - E
PAX (2017) Inception report, Engendering the Transition to Peace and Security in Iraq
PAX (2018) Annual report 2018, Engendering the Transition to Peace and Security in Iraq
PAX (2019) Annual report 2019, Engendering the Transition to Peace and Security in Iraq

South Sudan

ELSconsults (2021) End Evaluation
MFA (2016) Activity Appraisal Document ODA
MFA (2019) Letter approving change in audit schedule
MFA (2019) Partner Results Summary
MFA (2020) Activity Appraisal Document ODA
Plan (2016) Gender Sensitive Context, Actors, and Conflict Analysis
Plan (2016) M & E Framework

Plan (2016) Proposal
Plan (2016) Theory of change
Plan (2017) Annual Narrative and Financial Report
Plan (2017) Inception Report
Plan (2018) Annual Narrative and Financial Report
Plan (2018) Annual Plan
Plan (2019) Annex 3 - Detailed Financial Budget 2019
Plan (2019) Annual Narrative and Financial Report
Plan (2019) Detailed overview of Results
Plan (2019) Plan and Budget
Plan (2020) Annual Progress Report 2020 and Final Narrative Report
Plan (2020) Final Financial Report
Plan (2021) Annex 10 - Management response to evaluation
Plan (2021) Annex 11 - Promoting Safeguarding in Nimule and Torit
STAD (2019) Annex 4 - Letter regarding partnership with EWO

Syria

Hivos (2017) Annual narrative progress report, Syrian Women's Participation in Peace and Security, the way forward
Hivos (2017) Annual report, Annex B, D and cover letter
Hivos (2018) Annual narrative progress report, Syrian Women's Participation in Peace and Security, the way forward
Hivos (2018) Annual financial report and cover letter
Hivos (2018) Annual report output table
Hivos (2020) Final Report (December 2016 – December 2019) - Syrian Women's Participation in Peace and Security: The Way Forward
Hivos (2020) Final narrative report 2020, Syrian Women's Participation in Peace and Security: The Way Forward
Hivos (2017) Inception report
Hivos (2017) Inception report Annex B budget
Hivos (2017) Inception report Annex C risk analysis
Hivos (2017) Inception report Annex D MEL plan

Hivos (2015) Programme proposal 2016-2019
Hivos (2015) Programme proposal 2016-2019 Activity appraisal doc
Hivos (2015) Programme proposal 2016-2019 ToC
Hivos (2015) Programme proposal 2016-2019 Budget
Hivos (2020) Programme proposal 2020

Yemen

CARE (2016) CARE Yemen grant application 2016-2019
CARE (2016) Grant application: Activity appraisal document
CARE (2016) Grant application: Final proposal narrative
CARE (2016) Grant application: MEL Plan
CARE (2017) Women Peace and Security in Yemen Annual Report, November 2016 - December 2017
CARE (2017) Women Peace and Security in Yemen Inception report
CARE (2018) Annual plans 2018, Domain 1-4
CARE (2018) Women Peace and Security in Yemen Annual Report, January - December 2018
CARE (2019) Annual plans 2019, Domain 1-4
CARE (2019) CARE Yemen grant application 2020
CARE (2019) CARE Yemen grant application: Activity appraisal document
CARE (2019) CARE Yemen grant application: Annual Plan
CARE (2019) Women Peace and Security in Yemen Annual Report, January - December 2019
CARE (2019) Women Peace and Security in Yemen Annual Report, January - December 2019 Annexes 1 - 14
CARE (2019) Women Peace and Security Program in Yemen NAP 1325: Final report
CARE (2019) Women Peace and Security Program in Yemen NAP 1325: Final report, Annex 1- 13
Insight Source Centre (2021) Final Evaluation NAP Yemen programme
MFA (2017) CARE Yemen Approval letter
MFA (2018) CARE Yemen Annual Report Approval letter
MFA (2019) CARE Yemen Annual Report approval letter

Framework Agreement Evaluations 2020, Lot 2 (Other evaluations)

1 General information

Perceel	DSO
Closing date for proposals	14-05-2021
Assessment criteria	CV's, project proposals and price
Award criteria and weighing	Evaluation team: 40% Methodology (technical proposal): 40% Planning: 10% Budget and price (financial proposal): 10%
Max. amount of CV's	5
Language requirements	English, French
Location	Remotely and in-country for the case studies
Contact persons	

2 Terms of Reference

Consultancy for the external evaluation of WPS 2016-2019 and WPS 2020

Background

On March 8th 2016 the third Dutch National Action Plan 2016-2019¹ (NAP 1325-III) was launched and signed by over 50 civil society organizations, the Dutch National Police, four Ministries (the Ministry of Foreign Affairs, the Ministry of Defence, the Ministry of Security and Justice, and the Ministry of Education, Culture and Science) and four knowledge institutions. Together, the NAP signatories form a platform for cooperation with the joint overall objective of contribution to *"an enabling environment for women's participation and empowerment in conflict and post-conflict environments, so they can meaningfully participate in conflict prevention, resolution, peacebuilding, protection, relief and recovery"*. The third NAP1325 furthermore determined eight focus countries, namely Afghanistan, Colombia, the Democratic Republic of Congo (DRC), Libya, Yemen, Syria, Iraq and South Sudan.

Together, the signatories defined three interlinked goals:

1. Better protect women and girls in conflict and post-conflict situations from violence and violations of their rights;
2. Subvert harmful underlying gender norms, which are obstacles to sustainable peace;
3. Ensure that women have equal leverage in conflict prevention and resolution, peacebuilding, relief and recovery at all levels, and that their efforts are acknowledged and supported.

Since the adoption of UN Security Council Resolution (UNSCR) 1325 much has been achieved, particularly with regards to setting norms, but implementation of the agenda for Women, Peace and Security (WPS) is lagging behind. The ambitions expressed in the various resolutions contrast sharply with reality. The situation of women in conflict areas has hardly improved and in some cases has even deteriorated.

¹ <https://www.nap1325.nl/assets/00-FINAL-NAP-2016-2019.pdf>

In order to improve implementation of the WPS-agenda, the Netherlands has set up two subsequent policy frameworks: Women Peace and Security (Dutch: *Vrouwen, vrede en veiligheid*) 2016-2019², hereafter WPS 2016-2019, and Women Peace and Security (Dutch: *Vrouwen, vrede en veiligheid*) 2020³, hereafter WPS 2020. Through these policy frameworks, the Netherlands financed eight consortia of non-governmental organisations and knowledge institutions to implement programmes⁴ (hereafter: NAP programmes) in the focus countries of the third NAP1325, with a total budget of around EUR 17.4 million:

- Afghanistan, lead Oxfam Novib
- Colombia, lead ICCO
- DRC, lead Stichting CMC – Mensen met een Missie
- Libya, lead Cordaid
- Yemen, lead CARE Nederland
- Syria, lead Hivos
- Iraq, lead PAX
- South Sudan, lead Plan International

The Theories of Change (ToCs) of the NAP programmes are aligned with the ToC of the third NAP1325 of the Netherlands, and therefore generally focus on better protecting women and girls in conflict and post-conflict situations, subverting harmful underlying gender norms, and ensuring equal leverage for women.

Objective of the consultancy

The Taskforce Women's Rights and Gender Equality (TFVG), part of the Department for Social Development (DSO) at the Dutch Ministry of Foreign Affairs (MFA) is requesting an independent external end-term evaluation of the policy frameworks WPS 2016-2019 and WPS 2020. The objective of this end-term evaluation is to gain insight into the extent to which the main goals of these policy frameworks have been achieved, and if so, how.

This objective will be achieved on the basis of:

- A **meta-evaluation** of the programme evaluations that have been externally commissioned by the NAP consortia. Seven of the eight consortia have commissioned an external evaluation of their programmes⁵, of which the evaluation reports have already been submitted to the TFVG or are expected to be submitted in due course.
- **Additional desk research and interviews** to elaborate on and substantiate the conclusions from the programme evaluations that were externally commissioned by NAP consortia.
- Complementing results by conducting two or three **case studies through (digital) missions** to two or three NAP-countries⁶. This aspect is to be discussed and designed further in the inception phase of the consultancy.

² <https://wetten.overheid.nl/BWBR0037916/2016-05-13/>

³ <https://zoek.officielebekendmakingen.nl/stcrt-2019-50466.html>

The policy framework WPS 2020 was set up to cover the gap between WPS 2016-2019 and the new policy framework Women, Peace and Security 2021-2025. It was expected that partners under WPS 2016-2019 would also apply for funding under Women, Peace and Security 2021-2025. A one-year gap between these programmes could have negative impacts on the established networks and the results achieved, whereas an interim-subsidy framework could contribute to consolidation of results, retention of networks, and a smooth transition into new programmes under Women, Peace and Security 2021-2025.

⁴ Five of the eight NAP programmes will continue under the Strengthening Civil Society – Women, Peace and Security partnership Fund, albeit with somewhat different consortia, renewed focus and larger budgets. The NAP programmes in Afghanistan and Libya will not continue, even though some of the activities will be carried on in the context of other programmes. The proposed WPS programme targeting Syria and surrounding countries is still under review.

⁵ The NAP Syria consortium, led by Hivos, has not commissioned an external evaluation for their programme, so their annual (M&E) reports will be used as a basis for the meta-evaluation.

⁶ In case the quality of some of the programme evaluations is low, the case studies would preferably focus on those countries in order to allow for complementation of findings. Furthermore, elements to be taken into

NB: In order to ensure reliability and validity of the outcomes of this independent external end-term evaluation, the TFVG is requesting the team of consultants to conduct a **quickscan** (during the inception phase of the evaluation) of the quality of the material available for the meta-evaluation. The recently updated IOB quality criteria should be used to assess the quality of the available material. Based on the results of this quickscan, the team of evaluators is requested to propose a sound methodology and way forward (e.g. in terms of scope, depth and limitations of the evaluation) for the external evaluation, in which the focus should be on aggregating and complementing the material that is indeed reliable and valid.

In order to inform future policy and programming, main findings, conclusions and recommendations are at least expected on the evaluation criteria, good practices, lessons learned and challenges encountered. Next to building on relevant conclusions from the external evaluations of the NAP programmes, this external end-term evaluation should take into account and build on relevant conclusions of the IOB evaluation on Gender, Peace and Security (2015)⁷ and the Mid Term Review of the third NAP1325⁸.

The principal users of the final evaluation report are the TFVG and all organizations and key stakeholders involved in the eight NAP consortia. The evaluation report will be published at www.Rijksoverheid.nl and www.government.nl. Therefore, other users include a broad (inter)national network of organizations involved in the same sector, as well as the general public.

Evaluation criteria

The evaluation criteria to be used in this evaluation are in accordance with international standards (OECD/DAC⁹) and described as follows:

- **Relevance:** The extent to which the objectives and design of interventions under the policy frameworks WPS 2016-2019 and WPS 2020 respond to beneficiaries', global, regional, country, and partner/institution needs, policies, and priorities, and continued to do so when circumstances changed.
- **Coherence:** The extent to which interventions under WPS 2016-2019 and WPS 2020 are compatible with other interventions in the NAP countries and in the field of WPS.
- **Effectiveness:** The extent to which the objectives of WPS 2016-2019 and WPS 2020 were achieved, or are expected to be achieved, including any differential results across groups.
- **Efficiency:** The extent to which interventions under WPS 2016-2019 and WPS 2020 delivered results in an economic and timely way.
- **Impact:** The extent to which interventions under WPS 2016-2019 and WPS 2020 generated significant positive or negative, intended or unintended, higher-level effects.
- **Sustainability:** The extent to which the net benefits from WPS 2016-2019 and WPS 2020 continue, or are likely to continue.

Guiding questions

Please find below a tentative list of evaluation questions (linked to the relevant evaluation criteria), to be finalised during the inception stage of the consultancy:

- To what extent have NAP programmes contributed to the overall objective of creating an enabling environment for women's participation and empowerment in conflict and post-conflict environments, so they can meaningfully participate in conflict prevention, resolution, peacebuilding, protection, relief and recovery? (**impact**)

account in selecting the countries for the case studies are: 1) preference for covering countries in which new programmes under WPS will be implemented, and 2) adequate geographical spreading.

⁷ <https://www.government.nl/documents/reports/2015/04/01/iob-gender-peace-and-security-evaluation-of-the-netherlands-and-un-security-council-resolution-1325>

⁸ <https://www.government.nl/documents/reports/2020/02/28/mid-term-review-national-action-plan-on-women-peace-and-security-2016---2019#:~:text=The%20goal%20of%20this%20Mid,development%20of%20the%20fourth%20NAP1325.>

⁹ Please see <https://www.oecd.org/dac/evaluation/dacriteriaforevaluatingdevelopmentassistance.htm> and <http://www.oecd.org/dac/evaluation/revised-evaluation-criteria-dec-2019.pdf>.

- To what extent have NAP programmes contributed to the three specific objectives laid down in the third NAP1325 of the Netherlands and the policy frameworks WPS 2016-2019 and WPS 2020? What did pathways of change look like? (**effectiveness**)
- To what extent were NAP programs aligned to the two WPS policy frameworks, national and decentral policies and needs of the target institutions and groups? To what extent did the NAP programs adequately respond to changing contexts? (**relevance**)
- What is the likelihood of the results of the NAP programmes to be maintained in the longer term (post 2020 after NAP programmes have ended)? What steps did consortia undertake to ensure sustainability of programme outcomes? Did, and if so, how did, an additional year of programming (VVV 2020) allow NAP consortia to undertake additional steps to strengthen the sustainability of the results of the NAP programmes? (**sustainability**)
- To what extent did the design and implementation framework (activities, overhead, coordination budget, internal procedures, human resources; monitoring & evaluation, resource mobilisation and budget allocations, e.g. between consortium partners and to local partners) foster or hinder the achievement of the program outputs? What were the most critical factors affecting value for money during programme implementation? (**efficiency**)
- What unintended consequences, both positive and negative (incl. from a Do No Harm perspective), did NAP programmes have and in which ways did they affect the different target groups and stakeholders? (**impact**)
- Which main lessons learned, good practices and challenges can be identified within the NAP programmes?

In answering the evaluation questions, the team of consultants is asked to also include a comparative aspect. What are the stronger and weaker aspects of the NAP programmes, what is the cause of these difference and what are their results? What lessons can be drawn from this comparative analysis?

Lastly, TFIG is also interested in learning more on how the outcomes and lessons learned of the NAP programme compare to the broader (international) evidence base on 'what works' in WPS programming.¹⁰ Are the outcomes of the programmes validated by international research, or do they show a different picture? If so, what elements explain this difference?

Scope

This independent external end-term evaluation will cover the period January 2016 – December 2020 (both WPS 2016-2019 and WPS 2020). Geographically, the scope of this end-term evaluation will encompass all eight NAP countries¹¹. Case studies in two or three NAP countries will contribute to complementing and illustrating evaluation results.

Methodology

The evaluation team is asked to propose a gender-responsive approach to achieve the objectives of this consultancy. Considering the character of the NAP-programmes, it is likely that a mixed method evaluation of both quantitative as well as qualitative research methods¹² is best suited to achieve the objectives of the consultancy. Furthermore, considering the current (travel) restrictions, the team of consultants is encouraged to propose creative and innovative approaches and methodologies in their proposal for this evaluation.

Please find below a tentative list of proposed evaluation activities, which is to be finalised during the proposal and inception stage of the consultancy:

1. Desk review and analysis of programme documents, most importantly the reports of the external evaluations of the NAP programmes, as well as their annual reports, MTR reports, year plans, inception reports, proposals and budgets;
2. Desk review and analysis of internal MFA documents as well as documentation provided by the partners, most importantly assessments of reports and year plans, approval/rejection

¹⁰ Including a qualification of the results by comparing them to systematic reviews and evidence based on lessons learned in broader WPS programming.

¹¹ Afghanistan, Colombia, DRC, Libya, Yemen, Syria, Iraq and South Sudan

¹² Using Outcome Harvesting as a qualitative research method is not preferable to TFIG.

letters, programme appraisal documents, policy dialogue records and programme decisions;

3. Interviews with consortium leads, consortium partners and other persons who have been involved in the NAP programmes (f.e. financial, M&E);
4. Interviews with local partner organisations;
5. Interviews and/or (online) surveys with beneficiaries;
6. Interviews and/or (online) surveys with external parties (e.g. WO=MEN, other bilateral donors, multilateral organizations, other (local) NGOs/CSOs, in-country government actors and institutions, local media and/or journalists and opinion leaders);
7. Interviews with TFIG, DSO Control Unit (CU), and Embassies of the Kingdom of The Netherlands;
8. Where relevant: broader research on WPS programming, for comparative purposes.

NB: Considering the ongoing COVID-19 pandemic, the evaluation activities will in principle be carried out remotely. Should the situation allow it, live interviews and physical missions could be part of the activities and thus tentatively included in the budget.

Ethics to be applied

The consultant team should be dedicated to ensuring that the highest standards of ethics will be applied in this evaluation in line with the UNEG Code of Conduct for Evaluation¹³. The consultant team should include the ethics and safeguarding approach for this evaluation in the inception report.

Deliverables

- A concise **evaluation inception report** should be prepared by the evaluators before starting the evaluation. This report details the evaluators' understanding of what is being evaluated and why, showing how each evaluation question will be answered by way of: proposed methods, proposed sources of data and data collection procedures. The inception report should include an evaluation matrix, data collection tools, a proposed schedule of tasks, activities and deliverables. Considering this assignment concerns the evaluation of programmes that were conducted in fragile settings and/or settings in which civic space is (severely) restricted, the evaluation inception report should also include a risk analysis.
- Furthermore, the evaluation inception report should include the results of a **quickscan**¹⁴ **of the quality of the material** available for the meta-evaluation. The results¹⁵ of this quickscan should inform the proposed methodology and way forward (e.g. in terms of scope, depth and limitations of the evaluation) for the external evaluation, in which the focus should be on aggregating and complementing the material that is indeed reliable and valid.
- The evaluation team is required to submit a **draft evaluation report**¹⁶, in English, of max 40 pages, excluding the executive summary, annexes, case studies and good practices. The evaluation report responds to the evaluation objectives and questions as described above. The reflections on and answers to the evaluation questions should inform and substantiate the conclusions and recommendations.
- The evaluation team is expected to **validate its draft findings** with key stakeholders in a validation / sense-making workshop. If circumstances do not allow for in-person meetings, this workshop will be held online.

¹³ Available on: <http://www.unevaluation.org/uneqcodeofconduct>

¹⁴ The recently updated IOB quality criteria should be used to assess the quality of the available material.

¹⁵ The results of this quickscan should also be shared with the NAP consortia, in order to allow them to provide their feedback and clarify any questions that might have arisen during the quickscan.

¹⁶ The raw data sets must be made available to the client upon request. Furthermore, this evaluation might generate datasets that are potentially useful for other research, later, possibly by other researchers. The Contracting Authority may forward these data (in an anonymized form) to the Data Archiving and Networked Services, as explained in Article 4 of the Framework Agreement Evaluations 2020. The Contracting Authority will indicate this potential 'multiple use' in the Request for the performance of services, on a case by case basis. The Contractor should then explain the potential multiple use of the research data to the participants in surveys or interviews when asking for their consent.

- The evaluation team is expected to **present its draft findings** at the MFA to the evaluation reference group if the circumstances allow for in-person meetings. If this is not the case, the presentation will be held on a videoconferencing platform.
- The **final version of the evaluation report**, integrating feedback from the validation workshop and evaluation reference group should be submitted to the TFGV by 9 July 2021 and will be published on www.rijksoverheid.nl and www.government.nl. Next to a final evaluation report, the evaluation team is expected to, for communication purposes and in consultation with TFGV, develop a **one-pager and infographics** outlining the main results, lessons learned and best practices of five years of programming.
- The evaluation team is also expected to **present the final evaluation findings** at the MFA with a wider audience (f.e. international donors, the wider NGO community, et cetera) if the circumstances allow for in-person meetings. If this is not the case, the presentation will be held on a videoconferencing platform.

The evaluation process will be guided by an evaluation reference group consisting of internal and external members selected by MFA. Its responsibilities are as follows:

- Review and provide comments on the inception report (including the methodology and quickscan on the quality of available material);
- Review and provide comments on the draft evaluation report (in writing and during the presentation of the draft findings).

The reference group will consist of six members: two members from the NAP team of the MFA (one in the capacity of chair of the reference group, one as a member), one MFA colleague specialized in Monitoring, Evaluation and Learning (MEL), one (local) colleague from an Embassy in a NAP-country, and two external WPS experts (preferably from the global south).

Time path and budget

Evaluation timeline¹⁷ (2021):

- Start¹⁸ of the assignment: 15-06-2021
- Submission of inception report: 08-07-2021
- Review of the inception report by the reference group: 09-07-2021 to 15-07-2021
- Data collection and analysis: 16-07-2021 to 10-09-2021
- Submission of first draft evaluation report: 17-09-2021
- Validation of draft findings with key stakeholders: 20-09-2021 – 24-09-2021
- Presentation of draft evaluation findings and recommendations to the reference group: 20-09-2021 – 24-09-2021
- Submission of final evaluation report: 01-10-2021
- Presentation of final evaluation findings and recommendations to a wider audience: tbd

The **maximum budget** available for this evaluation is €100.000 (budgets will be assessed on their cost effectiveness). The Evaluator's proposal should include a detailed breakdown including number of working days per consultant, consultant fees, travel costs and per diems (if applicable and

¹⁷ The evaluation timeline covers the summer holidays, in which it might be more difficult to reach the relevant stakeholders and beneficiaries. The period for data collection and analysis has therefore been extended from an original 3 weeks to 8 weeks, which can be planned part-time and flexible according to availability of consultant and respondents.

¹⁸ The last NAP programme evaluation is expected to be submitted to the MFA at the end of June. It is proposed to start the assignment with the quickscan of six NAP programme evaluations and start developing the evaluation methodology accordingly, in which the quickscan of the seventh NAP programme evaluation would be integrated later.

feasible) VAT/taxes, et cetera. Payments will be based on deliverables as per the schedule above. All costs proposals should be made in EUR.

Evaluation team requirements

TFVG is requesting a team of independent¹⁹ evaluators to perform this external end-term evaluation, possibly with local experts for the (digital) field missions.

The team of evaluators should include consultants with at least the following skills:

- Master's degree in social sciences, international relations or other relevant field;
- Relevant work experience of at least 10 years
- A strong methodological background in quantitative and qualitative research methods and remote research methods, demonstrated by:
 - Considerable knowledge and experience with evaluating multi-country, multi-stakeholder international development cooperation programmes focusing on women, peace and security;
 - Experience in conducting documentation analysis and interviews;
 - Experience with analysis and evaluation of CSO programmes.
- Excellent English writing and communication skills, advanced knowledge of French²⁰. Knowledge of Arabic and/or Spanish is a plus;
- Excellent intercultural and interpersonal communication skills, including coordination, facilitation and presentation;
- Proven ability to interact proactively with clients.

NB: A mixed team including both senior and junior consultants, as well as international and local consultants, is welcomed. Furthermore, there is the option to include a dedicated consultant to conduct the cost analyses to assess efficiency aspects.

Submission of Proposals

Consultants are requested to provide technical and financial proposals (in total maximum 15 pages excluding CVs) to [REDACTED], Taskforce Women's Rights and Gender Equality, no later than midnight **14 May 2021**. Proposals should be submitted electronically to [REDACTED].

¹⁹ Proposed evaluators should have no previous or present involvement in the design or implementation of the programmes or policy under evaluation, nor in the design or evaluation of a preceding programme or policy phase. This includes research and advisory services. The involvement in a previous evaluation will not automatically lead to the exclusion for this evaluation – this depends a.o. on the extent to which previous evaluations have determined the design of the programmes and policy under evaluation. In case of doubt, the Candidate should bring the issue to the attention of the Ministry at the time the expression of interest is submitted. Based on the information provided, the Ministry may decide to accept or reject the Candidate for this evaluation.

²⁰ It is expected that some respondents will speak neither English or French, and that some of the relevant documentation will be in other languages as well (e.g. Spanish or Arabic). The team of consultants is expected to take this into account in the development of the proposal, f.e. by including (local) consultants who speak relevant languages, budget for translation/interpretation, et cetera.